Short-Term Price Elasticities of Heating Demand: A Statistical Analysis of Energy Billing Data in Germany

- Master Thesis -

In Partial Fulfillment of the Requirements for the Degree of Master of Science (M. Sc.)

Integrated Natural Resource Management (INRM)

Faculty of Life Sciences, Albrecht Daniel Thaer-Institute

Humboldt University of Berlin

Submitted by:

Marc Blauert

(Matriculation number: 605555)

Berlin, August 4, 2022

Submitted to:

1st Supervisor: Prof. Dr. Karsten Neuhoff (DIW Berlin, TU Berlin)

2nd Supervisor: Prof. Dr. Tobias Krüger (HU Berlin)

(Written in Cooperation with the Climate Policy Department at DIW Berlin)

Acknowledgements

I would like to express my sincere gratitude to my two supervisors, Prof. Dr. Karsten Neuhoff and Prof. Dr. Tobias Krüger, for their supervision, valuable guidance, and feedback. Furthermore, I would also like to express special thanks to Dr. Franziska Schütze for the extensive discussions and many practical suggestions provided throughout the course of writing the thesis. This also applies to the lively discussions and feedback from other colleagues in the Climate Policy Department at DIW Berlin – be it in personal conversations or during the joint Brownbag session. I would also like to acknowledge the contribution of Dr. Jan Stede, especially in the phase of exploring the initial research ideas.

Finally, I am truly grateful to my loved ones for their continued motivation and moral support.

Table of Contents

List of Tables

List of Figures

Abstract

Space heating demand of private households accounted for 13.6% of national primary energy demand in Germany in 2020. Due to its predominant generation from fossil fuels, space heating is also responsible for the majority of greenhouse gas emissions (GHG) in the residential buildings sector. In order to achieve ambitious climate targets, decarbonisation of the buildings sector represents a key challenge. Carbon pricing instruments, such as the newly introduced German Fuel Emissions Trading Act 2021, are key instruments to support the transformation of the buildings sector. However, to estimate the effects of any price-related policy instrument, the price responsiveness of residential consumers for the demand of space heating needs to be identified. Therefore, this thesis uses a large-scale energy billing sample of multi-apartment buildings in Germany between 2007 and 2019 to produce estimates of the short-term price elasticity of space heating demand. The results of the analysis show that the demand for space heating is highly inelastic. An overall price elasticity of space heating demand of -0.243 is found relying on the full set of data. Furthermore, when relying on a random stratified subsample, it is found that the different energy carrier groups are associated with varying price elasticities of demand. Buildings with oil heating are associated with the lowest elasticity of -0.16 [-0.10; -0.23] (Mean estimate [95%CI]), gas with an elasticity in the mid-range of -0.35 [-0.29; -0.40] and district heating with a still inelastic but higher elasticity of -0.53 [-0.46; -0.60]. For the wider policy context, the inelastic estimates suggest that to achieve ambitious climate targets, carbon pricing in the buildings sector needs to be embedded in a policy mix to make the transition to a low-carbon buildings sector socially desirable and politically feasible.

Code, GitHub Repository:

`https://github.com/marcblauert/price-elasticities-heating-demand-ext.git`

Chapter 1

Introduction

1.1 Background

The energy demand of residential buildings represents one of the largest energy demand points in Germany and worldwide. In Germany specifically, the building energy demand of private households accounted for 20.2% of the overall primary energy demand in 2020 (AGEB, 2021). While households consume energy in their homes for various purposes, 68.1% of the overall residential energy demand in 2020 was used to generate space heating (RWI, 2021). Space heating, therefore, represents the most significant type of energy end-use in residential buildings in Germany. It far exceeds the energy demand for other types of end-uses such as hot water generation (16.1%), process heat (6.0%), room cooling, information technology, or lighting (all below 5%) (RWI, 2021). Relating the generation of space heating to the overall primary energy demand in Germany, it represents 13.6% of the national primary energy demand.

Also, from an emissions perspective, the two end-uses of space heating and hot water generation are the predominant contributors to GHG emissions in the residential buildings sector. That is because they continue to be produced predominantly from fossil fuels. In 2020, 80.1% of the energy used to generate space heating in private households in Germany came from either natural gas (45.4%), oil (24.5%), or district heating (10.1%)[1] (RWI, 2021). Thus, the generation of space heating is not only the most significant energy demand in residential buildings but also the largest contributor to GHG emissions due to its generation from fossil sources.

[1]This is currently also generated to almost 90 % from fossil fuels (natural gas or coal) or through waste incineration (DESTATIS, 2018).

Over the past decades, multiple trends in the residential buildings sector have affected energy demand and associated GHG emissions. Looking at a more extended period between 1995 and 2020, greenhouse gas emissions in the entire building sector (including commercially used buildings) in Germany have fallen by 34% (ERK, 2020). When, however, only considering a shorter historical period, it becomes apparent that much of the energy efficiency gains in the buildings sector were already realised in the 1990s and 2000s. For the last decade, on the contrary, several independent data sources indicate that the primary energy demand for residential space heating per square meter in Germany has stagnated if effective energy demands are corrected for spatial and temporal variations in climatic conditions (BMWi, 2021; Stede, Schütze, & Wietschel, 2020; Techem, 2019). Furthermore, from a sectoral perspective, energy efficiency gains (better insulation of building stock, more efficient heating technologies) and a lower emission intensity from fuel switches are being partly offset by countervailing effects, including an increase in the number of apartments (smaller average household size) as well as the rise in per capita living space (BMWi, 2021).

In the face of climate change, there is an urgent need to decarbonise the buildings sector (Levesque, Pietzcker, Baumstark, & Luderer, 2021). In response to the Paris Agreement, which aims to limit global warming to well below 2°C by 2100, Germany introduced a national Federal Climate Change Act (Bundes-Klimaschutzgesetz, KSG) to govern the low-carbon transformation. The latest 2021 amendment of the KSG stipulates that national GHG emissions must be reduced by 65% by 2030 (compared to the base year 1990), and net-zero emissions must be reached in 2045. Congruently, the Green Deal of the European Union (EU) defines 2050 as the target year to achieve net-zero in the whole union (European Comission, 2019). One of the most relevant changes in the KSG amendment of 2021 is the introduction of annual sector targets as a concretisation of the overall national emission reduction target. One of the six sectors defined in the KSG is the buildings sector, whose emissions must be reduced from 118 million tonnes of carbon dioxide equivalents (CO_2e) in 2020 to 67 million tonnes of CO_2e in 2030 (-43.2%).[2] To achieve these ambitious reduction targets, a comprehensive transformation of energy use in the buildings sector is required. This includes measures to improve the efficiency of the building stock, the electrification of the remaining energy use on the demand side and the decarbonisation of energy generation on the supply side (Hennes et al., 2021; Levesque et al., 2021).

[2]Federal Climate Change Act of 12 December 2019 (Federal Law Gazette I, p. 2513), as last amended by Article 1 of the Act of 18 August 2021 (Federal Law Gazette I, p. 3905).

1.2 Objective and Relevance

Economic theory relies on the rationale that prices represent the key driver of the level of demand for a good (Pindyck & Rubinfeld, 2018; Stigler, 1987). Therefore, carbon pricing instruments are one of the key instruments to support and drive the decarbonisation of the buildings sector while being supplemented by additional policies (Braungardt, Bürger, & Köhler, 2021). Ultimately, however, modelling the effectiveness and impact of any price-related energy and climate policy instrument strongly depends on the assumed responsiveness of consumer demand (Alberini, Gans, & Velez-Lopez, 2011; Zhu, Li, Zhou, Zhang, & Yang, 2018).

Against this background, this thesis aims to use historical observational data on the space heating demand of private households in Germany to analyse the determinants of demand levels statistically. More specifically, the thesis focuses on estimating short-term own-price elasticities of space heating demand and thereby supplements the existing literature on the responsiveness of the demand response for space heating. The focus is on space heating, as it is the most relevant type of energy end-use in residential buildings and the primary driver of GHG emissions. The empirical basis for the analysis is a sample of energy bills of multi-apartment residential buildings in Germany between 2007 and 2019.

Russia's war of aggression against Ukraine launched in February 2022, and the associated concern about a gas shortage in Europe has given the discussion about reducing energy demand a whole new dimension in recent months. While empirical price elasticities of space heating demand can, in principle, indicate a price-induced reduction in demand that follows from an energy price increase, the current situation must be seen as structurally different from the period studied in this thesis. While the price developments between 2007 and 2019 were relatively gradual, the current situation, in contrast, represents an extreme event with an unprecedented rise in energy prices coupled with media attention and political appeals to save energy. The presence of these structural differences should be considered if the price elasticities presented in this thesis were to be used as input parameters for an analysis of the current crisis.[3]

Evaluation of price-related energy policies

Estimating price elasticities of space heating demand in the residential buildings

[3] First studies that focus explicitly on the period of the recent gas crisis already exist. See for example Ruhnau, Stiewe, Muessel, & Hirth (2022).

sector is relevant for several reasons.

First, empirical evidence on price elasticities is generally important for evaluating policies that affect the price of energy. For any such price-related policy, one needs to understand the relevant determinants which affect the energy demand. Based on elasticity estimates, the economic, distributional, and environmental impacts of policy measures can be better understood and estimated (Alberini et al., 2011; Labandeira, Labeaga, & López-Otero, 2017; Zhu et al., 2018). This also applies to price-related policy instruments that aim to contribute to the attainment of ambitious climate targets.

BEHG and its short- and long-term effects as example

One of the essential policy measures to induce GHG emission reductions in the buildings sector in Germany is the recent introduction of a carbon price for heating fuels, which was rolled out in 2021 based on the Fuel Emissions Trading Act (BEHG).[4] In general, the introduction of carbon pricing aims to correct the market failure of the previously unrecognised external effects of GHG emissions (Aldy, Krupnick, Newell, Parry, & Pizer, 2010; Stiglitz, 2019). It is expected that the consideration of external effects will have a steering effect on demand for fossil fuels through an increased price level. In practice, the recently introduced Fuel Emissions Trading Act specifies a fixed price path until 2025. The price path rises from an initial price of 25 Euros per ton of CO_2e in 2021 to 55 Euros per ton of CO_2e in 2025.

Figure 1.1 illustrates the price effect of the carbon pricing for the buildings sector through the BEHG. For gas (Panel A) and oil (Panel B), the respective left bar (green) represents the average end consumer prices for the five years 2016-2020 from national statistics (BMWi, 2021).[5] The two middle bars (yellow) show the absolute price effects that occur based on the carbon intensity of the energy carrier in 2021 (25 Euros per ton CO_2e) and until 2025 (55 Euros per ton CO_2e). The yellow bars

[4] Fuel Emissions Allowance Trading Act (BEHG) of 12 December 2019 (Federal Law Gazette I p. 2728) as amended by Article 1 of the Act of 3 November 2020 (Federal Law Gazette I p. 2291). Emissions from the buildings sector have not been subject to carbon pricing in the past in Germany, as they do not fall within the scope of the EU Emissions Trading Scheme (ETS). Besides emissions in the buildings sector, the BEHG also covers emissions from the transport sector. From 2026, the fixed price path is planned to be replaced by a national ETS with a defined price corridor.

[5] The five-year period (2016-2020) was chosen due to the volatility of oil prices in recent years. An exclusive consideration of the 2020 average price would have distorted the relative price effects. Furthermore, it should be noted that district heating is not considered in the figure because most district heating power plants were already subject to carbon pricing under the EU ETS. Only smaller plants with a total combustion capacity of less than 20 megawatts (MW) were not already subject to the EU ETS and are now subject to BEHG carbon pricing.

are larger for oil as an energy carrier due to its higher emission intensity.[6] Assuming otherwise constant prices, the initial CO_2e price in 2021 would thus lead to a price increase of +0.50 Cents (+7.4%) per kWh gas and +0.66 Cents (+13.4%) per kWh oil. In 2025, the price effects rise to +1.10 Cents (+16.3%) per kWh gas and +1.46 Cents (+29.6%) per kWh oil.

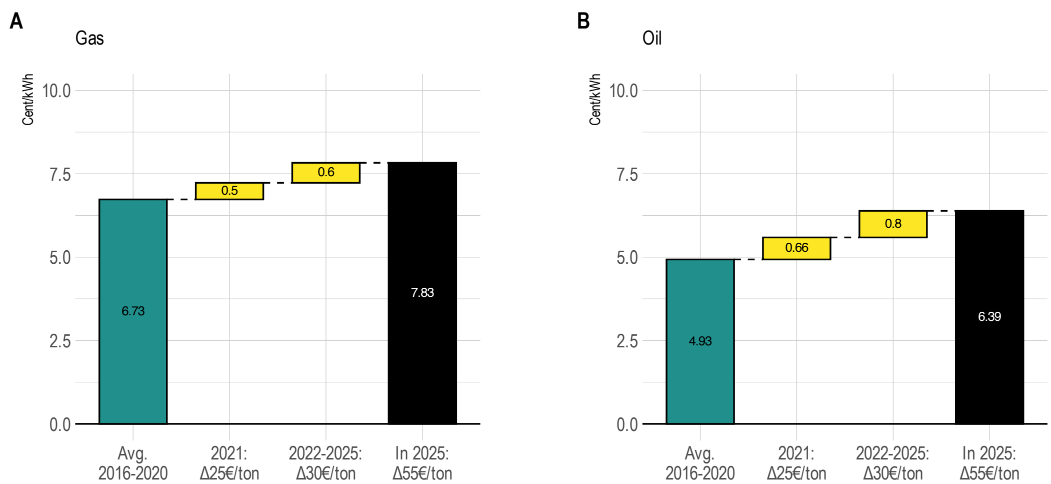

Figure 1.1: Price effects on gas and oil induced by the Fuel Emissions Trading Act (BEHG) until 2025

The relative price increases clearly show that considerable steering effects can already be expected from the BEHG in the next few years. Higher price levels driven by the structural price component of carbon pricing can be expected to have two impact channels. The first impact channel is short-term price reactions. Consumers adjust their demand levels for heating fuels due to higher prices (Alberini et al., 2011). Short-term demand responses are of direct relevance when making predictions on the fuel demand of the buildings sector. The expected magnitude of the effect depends on the demand response, which is reflected in the short-term price elasticity of space heating demand.

The second impact channel is that structurally higher fossil fuel prices will incentivise long-term investment in low-carbon heating technologies and energy efficiency. In the literature, long-term investment decisions, driven by mitigation costs poten-

[6] The emission factors for gas (0.201 kg CO_2e per kWh) and oil (0.266 kg CO_2e per kWh) were taken from UBA (2016).

tials, are modelled using integrated assessment models (IAMs) for the buildings sector (e.g., Bürger, Hesse, Köhler, Palzer, & Engelmann, 2019; Levesque et al., 2021). Depending on the specific model structure, these IAM sector modules are often informed by assumptions on the price elasticities of heating demand as one of their input parameters.[7] Thus, empirical estimates of the short-term price elasticity of space heating demand are also necessary for further analyses of the long-term steering effect concerning investments in energy efficiency and low-carbon technologies in the buildings sector.

Use of billing data as a complementary source of evidence

Another line of argument for the relevance of this study is that the existing empirical evidence on the price elasticity of space heating demand is based on varying types of data sources. But to date, few studies have been conducted based on energy billing data. For the case of Germany, ex-post statistical analyses for the residential buildings sector predominantly rely on panel data from social surveys (e.g., Rehdanz, 2007; Schmitz & Madlener, 2020; Schulte & Heindl, 2017). The studies are built around the survey item *household expenditure for heating.* But actual levels of energy demand and prices are not observed. This means that, for instance, different energy contract conditions between households may obscure accurate demand levels and thus affect the validity of the estimated results.

The data available for this thesis is different since the energy bills provide both energy demand and associated costs on the level of the buildings. To the best of the authors' knowledge, previous studies based on energy billing data have only been conducted in the United States (see Auffhammer & Rubin, 2018).

While using energy billing data offers unique characteristics that social survey data cannot mirror, it should be pointed out that its use also has its weaknesses. Due to the data structure, the billing data does not observe demand responses at an individual household level but at the level of multi-apartment buildings. This aggregation may obscure the demand response of a single household. Furthermore, the use of billing data makes the consideration of socio-economic determinants difficult (see Chapter 2.3). In summary, one can thus argue that the information value of estimates derived from billing data can be seen as complementary to the existing estimates derived from

[7] For a more detailed example of how own-price elasticities of demand are taken into account in IAMs, see the model structure description of the ETSAP TIAM model in Loulou & Labriet (2008). Since IAMs vary in structure, other models may consider price elasticities differently. Additional highly relevant parameters include the cost of capital and substitution elasticities.

social survey data.

Furthermore, the period under investigation in this thesis (2007-2019) is more recent than the period considered in other studies, which further contributes to the relevance of this study.

Research questions

Given the recently tightened climate change mitigation targets for GHG emissions and the newly introduced carbon pricing for the buildings sector, additional evidence on the price elasticities of space heating demand is relevant. Therefore, this work aims to produce estimates of short-term own-price elasticities of space heating demand based on a large-scale energy billing sample in Germany. The guiding research questions for the analysis are:

- How does a change in energy price affect the level of space heating demand (price elasticity estimate)?
- What other determinants affect the level of space heating demand and need to be considered so that their effects are not falsely attributed to energy prices?
- Are there potential factors for heterogeneity in the sample that may not be reflected in the estimation of an aggregated price elasticity of demand?

The thesis is structured into six Chapters. Chapter 1 provided the introduction. Chapter 2 delivers a theoretical background on the price elasticity of demand, presents the prior literature, and develops a conceptual model for this thesis. The data and empirical approach are described in the following Chapter 3. Furthermore, Chapter 3 also provides descriptive statistics. Chapter 4 provides the results of the statistical analysis. In Chapter 5, the results are discussed and integrated into the building sector's policy context. Chapter 6 provides concluding remarks.

Chapter 2

Theory and Literature

The aim of this chapter is to provide the theoretical background on the price elasticity of demand and an overview of the previous literature on the price elasticity of space heating demand in particular. Subsequently, theory and evidence from the literature are used to develop a conceptual model with corresponding rationales for the relevance and expected direction of a variable's effect on space heating demand.

2.1 Price Elasticity of Demand

Elasticities, in general, are one of the key concepts in micro-economic theory. They are used to express the sensibility of one variable to a change in another. The own price elasticity of demand is one type of elasticity. It represents the magnitude of a change in demand of a good following from a change in its own price (Pindyck & Rubinfeld, 2018). Formally, the own price elasticity of demand ϵ can be expressed as:

$$\epsilon = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} = \frac{P \Delta Q}{Q \Delta P} \tag{2.1}$$

where $\Delta P/P$ represents the percentage change in the own price P of a good and $\Delta Q/Q$ the corresponding percentage change in the quantity Q demanded of the same good. Due to budget constraints, consumers tend to consume less of a good when its price increases, which implies that under normal conditions the price elasticity of demand is negative. But the responsiveness of the demand reaction may vary across different goods. A common conception in the econometric literature is that elasticities are a constant parameter (Varian, 2010). Under the assumption of a constant price elasticity of demand, the associated demand function that can be expressed as:

$$Q = AP^{\varepsilon} \tag{2.2}$$

where A represents an arbitrary constant and ϵ again is the price elasticity of demand. Taking the logarithms of the demand function removes ϵ from the exponent and yields:

$$ln(Q) = ln(A) + \epsilon \, ln(P) \tag{2.3}$$

which can be referred to as the elasticity case and will reappear at a later point in this thesis, when constructing the statistical regression models.

The role of varying price elasticities of demand

To make the underlying assumption of a constant elasticity more intuitive, it seems useful to return to the subject of interest for this thesis and show the behavior of demand curves for space heating under varying price elasticities of demand. Generally, the microeconomics literature distinguishes between three different *types* of elasticities. If the demand response for a good is greater than the change in its own price ($\epsilon < -1$), the good is called to be price elastic. If the demand response is exactly equal to the change in price ($\epsilon = -1$), it is said to be unit-elastic. And if the demand response is smaller than the price change ($0 > \epsilon > -1$) – indicating a lower price responsiveness – it is said to be price inelastic (Pindyck & Rubinfeld, 2018).[1]

[1] It should be noted that in the literature the minus of the demand elasticity ϵ is sometimes omitted, as it is assumed that it is generally negative. However, I personally find it more intuitive not to do so – especially when working empirically where positive estimates are a possibility – and will therefore continue to use the negative values in this thesis.

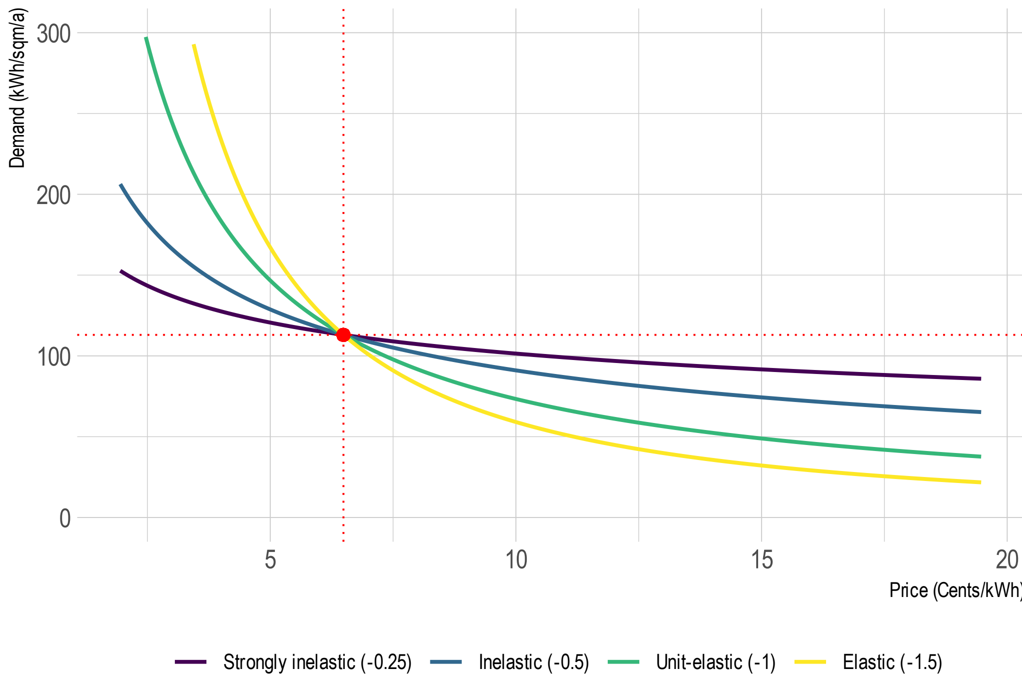

Figure 2.1: Exemplary demand curves associated with varying price elasticities of space heating demand

Figure 2.1 graphically represents demand curves for varying price elasticities of demand for space heating.[2] To construct the demand curves, a common arbitrary intersection point is chosen. The intersection point has a demand value of 113 kilowatt hours per square meter and year (kWh/sqm/y) and a price value of 6.49 Cents/kWh. These demand and price values are chosen because they reflect the average values in the empirical sample (see Chapter 3.4). The graph includes a total of four demand curves with varying price elasticities: A unit-elastic ($\epsilon = -1$) demand curve, an elastic ($\epsilon = -1.5$) demand curve and two inelastic ($\epsilon = -0.25, \epsilon = -0.5$) demand curves.

The special feature of the unit-elastic demand curve (green) is that it gives the same result for total expenditure on heating costs for all combinations of price and demand. This is not the case with the other price elasticities. In the elastic case (yellow curve), total expenditures become lower when the price is higher, which means

[2]Please note that in the graph, energy price as the independent variable is plotted on the horizontal axis and energy demand as the response variable on the vertical axis. I consider this convention – which is common in most of science – to be more intuitive than the traditional Marshallian cross diagram in Economics, where the price is plotted on the vertical axis.

that households can adjust their demand downwards more. In the opposite, inelastic cases (purple and blue curves), on the other hand, households would reduce their demand less and therefore also face higher total expenditure. Conversely, to the left of the intersection point (lower prices), more elastic demand (yellow curve) leads to a higher level of energy demand and thus also to higher total expenditure than in the inelastic cases (purple and blue curves), where demand changes only gradually.

Different types of goods are structurally associated with different price elasticities of demand (Varian, 2010). For the household demand for space heating it is reasonable to assume that demand will be is rather inelastic. There are two arguments that support this assumption. First, heating energy is to be regarded as an essential good that is associated with minimal demand in the winter months in order to ensure the well-being of residents. In general, essential goods tend to follow a more inelastic demand pattern than other types of goods (Gwartney, 1976). Households can only reduce their space heating demand to a certain extent (right side of the intersection), but on the other hand, they do not have a strong rationality for a strong increase in heating demand beyond a comfortable level of warmth (left side of the intersection). The second indication for a rather inelastic demand is that especially tenant households, from which the data for this thesis originate, have only very limited to no practicable options to replace space heating from one source with space heating from another due to their dependence on the existing heating system. With limited ability to substitute comes the necessity for continued demand (Pindyck & Rubinfeld, 2018).

The difference between short- and long-term price elasticities of demand

Another relevant dimension along which one distinguishes different price elasticities of demand is *time*. More precisely, the time that elapses between a price signal and the measurement of the demand response. Here, the literature distinguishes between short-term and long-term price elasticities of demand (also called short- and long-run elasticities) (Pindyck & Rubinfeld, 2018). Short-term elasticities of demand are those that are estimated in temporal proximity to the price change – in this case, that would be the energy price change within a billing year and the associated change in demand in the same period. Short-term price elasticities are mainly driven by behavioral changes such as lowering the average heating temperature or adjusting the ventilation behavior. But may also include minor constructional improvements, such as the installation of door and window sealings. Long-term elasticities, on the other hand, imply that more time has elapsed for consumers to fully adjust to a price change and also to make long-term investment decisions that would change the overall level of demand (Pindyck & Rubinfeld, 2018). For the energy demand of residential buildings, such decisions may include larger investments into the thermal insulation of a building or an exchange of the heating system. Due to these structural adjustments, long-term demand elasticities are usually higher (more elastic) than short-term demand elasticities (Schmitz & Madlener, 2020). Overall, the two types of short- and long term elasticities of demand can be seen as mirroring the two different channels through which a price-related policy instrument such as the BEHG can influence the demand behavior of residents (see Chapter 1.2). Short-term elasticities reflect the immediate demand reaction. And long-term elasticities observe the structural adjustments made based on a new market price and the demand that follows from it.

In the literature, the approaches in the studies differ in the sense that they estimate either short-term and/or long-term price elasticities. For this thesis, I follow the example of Schmitz & Madlener (2020) and only estimate short-term price elasticities of demand. Long-term price elasticities are often estimated using dynamic models, where the elasticity estimated between two periods represents a fraction of the long-term elasticity (e.g., Alberini et al., 2011). However, since the methods vary more widely I consider it as outside of the scope of this thesis.

2.2 Literature Review

In general, the study of the price elasticity of *energy* demand is a broad field of academic research which also has a long history dating back to the mid-twentieth century (e.g., Cutler, 1941; Houthakker, 1951). In the more recent past, as well a large number of academic studies have been published. These studies differ considerably in terms of focus and design. While a subset of studies focus on the demand responsiveness of private households for space heating in particular, most of the literature focuses on energy demand in general or other types of specific energy applications. To organize the comprehensive body of literature, I first focus on meta-studies to provide a first indication on price elasticity for energy in general, and then concentrate on the subset of relevant studies that focus on household space heating demand.

Several conceptual factors can systematically influence price elasticity estimates. The most important of these factors is the type of energy application analysed. The price responsiveness of private households for space heating demand is different form the price responsiveness for other types of energy application, such as the demand for transport fuels. Another relevant factor are different groups of consumers. Private households are likely to react differently to energy price changes than industrial or commercial consumers. Beyond those two dimensions, other factors which may also have an effect on elasticity estimates are the statistical methods used (time-series analysis, panel analysis, cross-sectional analysis), varying sources of data (national accounts, aggregate sector-level data, micro-data), varying spatial and temporal focus, measurement heterogeneity, and price trends (Csereklyei, 2020; Miller & Alberini, 2016).

Meta-estimates as a first indication of price elasticity estimates

To obtain a first indication of price elasticity estimates from previous studies, the meta-analysis by Labandeira et al. (2017) is particularly suitable. It covers a broad spectrum of energy applications and thus goes beyond previous meta-analyses that primarily focus on the elasticity of gasoline demand (e.g., Brons, Nijkamp, Pels, & Rietveld, 2008; Havranek, Irsova, & Janda, 2012). Moreover, it is the most recent comprehensive meta-analysis and thus also partly covers the time period relevant for this thesis.

In their analysis, Labandeira et al. (2017) include a total of 428 papers, all published between 1990 and 2016. For the demand of natural gas, they find an average short-term (long-term) price elasticity of -0.18 (-0.57) based on 230 individual estimates. And for heating oil they find a comparable average short-term (long-term) elasticity of -0.19 (-0.54), which, however, is based on only 44 individual estimates (Labandeira et al., 2017). An average elasticity for space heating – which is less focused on in the literature due to its relatively low prevalence – is not provided.

The above average estimates of the price elasticity of demand for the energy carrier groups gas and oil are drawn from a wide range of consumer groups. This means that the underlying studies include not only those that examine demand response patterns for residential demand, but also others that examine these patterns for the commercial building stock and/or the industrial sector. Labandeira et al. (2017) also provide average price elasticities stratified by consumer groups, but these are, on the other hand, not clustered by the type of energy carrier. When differentiated along the dimension of consumer groups, they report that the average estimates are slightly higher in the residential household sector (short-term: -0.22; long-term: -0.62) than in the industrial sector (short-term: -0.17; long-term: -0.51), which seems reasonable given that households likely have more ability to change their level of demand than the industrial sector with more fixed production patterns and the ability to forward costs. Overall, however, the average price elasticity of demand estimates are all in a similar range and together convey the message that demand for energy is strongly inelastic in the short-term. This evidence is in line with the theoretical arguments that energy is an essential good and involves limited substitutability in many applications (see Chapter 2.1).

The meta-estimates are useful in providing a high-level overview. However, as already mentioned, they include estimates from a wide range of energy applications that not necessarily reflect the demand response for the specific application of space heating in residential buildings. Furthermore, they are average estimates which may have the effect of masking discrepancies and variations between studies.

Subset of studies focusing on the price elasticity of space heating demand in the residential buildings sector

The studies focusing on space heating demand of the residential sector represent a subset of the stream of literature estimating price elasticities of energy demand. The subset of individual studies presented in the following were selected based on the twin criteria that they make specific estimates for the energy application of space heating in the residential buildings sector and are of good quality (only peer-reviewed journal articles). An overview of the selected studies is provided in Table 2.1.

There are three earlier studies that have a spatial focus on Germany. The first study was conducted by Rehdanz (2007), who estimates short-term price elasticities for different energy carrier groups for residential space heating demand. She uses social survey data on household-level heating expenditures from the Socio-Economic Panel (SOEP) in the years 1998 and 2003. Methodologically, a cross-sectional statistical analysis with dummy variables is conducted for the two years. In contrast to the meta-estimates presented in the previous part of the Chapter, the study finds that demand for heating oil is highly elastic with estimates ranging between -1.68 and -2.03, depending on the model specification. For gas, the results indicate an inelastic elastic demand between -0.44 and -0.63 (depending on the model specification) which is nevertheless significantly more elastic than the meta-estimates presented previously. Therefore, the study concludes that the type of energy carrier can have a strong influence on the price sensitivity of private households (Rehdanz, 2007). In a more recent study, as well based on the SOEP social survey data but covering a more comprehensive time period between 1996 and 2014, Schmitz & Madlener (2020) find a price elasticity of space heating and hot water demand of -0.31 to -0.43, depending on the model specifications. They do not differentiate the elasticities by energy carrier group. The price elasticities are derived from household expenditure elasticities,

Table 2.1: Individual studies focusing on the price elasticity of space heating demand

Study	Type of data used	Energy carrier	Short-term price elasticities
Studies with spatial focus on Germany			
Rehdanz (2007)	Social survey, cross-sectional data (SOEP), all types of buildings, 1998 and 2003	Gas Oil	-0.44 to -0.63 -1.68 to -2.03
Schmitz and Madlener (2020)	Social survey, panel data (SOEP), all types of buildings, 1996-2014	(All)	-0.31 to -0.43
Schulte and Heindl (2017)	Social survey, panel data (EVS), all types of buildings, 1993-2008	(All)	-0.50
Studies with focus on other countries and regions			
Alberini et al. (2011)	US, household-level social survey, 50 metropolitan areas, panel data, 1995-2007, only single-family homes and duplexes	Gas	-0.57 to -0.69
Auffhammer and Rubin (2018)	US, energy billing data, household-level, only California, panel data, 2003-2014	Gas	-0.17 to -0.23
Leth-Petersen and Togeby (2001)	Denmark, apartment-block level (>1,500 sqm), panel data, 1984-1995	Oil District heating	-0.08 -0.02
Meier and Rehdanz (2010)	UK, household-level social survey, panel data 1991-2005	Gas Oil	-0.34 to -0.56 -0.40 to -0.49
Metcalf and Hassett (1999)	US, household-level, panel data, 1984, 1987 and 1990	Gas	-0.48 to -0.71

as demand is not directly observed in the SOEP social survey data. In contrast to Rehdanz (2007), the study uses a panel structure where observations are clustered by households and years using fixed-effects regression models. They produce much lower price elasticity estimates, which are more consistent with the meta-estimates presented previously. Moreover, the study discovers that price elasticity is heterogeneous across different socio-economic groups. Given the household-level information at their disposal, they find that higher-income households are less sensitive to energy price changes than lower-income households. Likewise, homeowners show less sensitivity than tenant households (Schmitz & Madlener, 2020).

Also with a focus on Germany, Schulte & Heindl (2017) investigate the price and expenditure elasticities of private energy demand in Germany between 1993 and 2008. For their analysis, they use survey data from the *Einkommens- und Verbrauchsstichprobe (EVS)* and analyse them with a demand system approach. More precisely, they estimate a quadratic expenditure system. For space heating demand, they find an own price elasticity of -0.50 across all households. They further observe that the price elasticity changes with the level of total household expenditure, with households in higher expenditure strata reacting more strongly to price changes (Schulte & Heindl, 2017). Put differently, this means that following an increase in price level low income households tend to lower energy demand to a lesser extent as compared to households with a higher income. Thus, regarding the effect of household income, their results are contradictory to those of Schmitz & Madlener (2020).

Besides the studies focusing on Germany, Table 2.1 also reports the findings from five other international studies. Alberini et al. (2011) examine household demand for gas (and electricity) in single-family homes and duplexes using household-level social survey data in 50 metropolitan areas in the United States (US) over the period 1995-2007. As a modelling approach, they use static (short-term elasticities) and dynamic (long-term elasticities) panel models. For the static models, reflecting the approach also taken in this thesis, they find an own price elasticity of private gas demand between -0.57 and -0.69, depending on the model specification. These estimates are in their magnitude comparable to the results of an earlier study by Metcalf & Hassett (1999) who use the 1984, 1987 and 1990 waves of the US Residential Energy Consumption Survey (RECS) to examine homeowners' investments in energy efficiency measures and as part of their study find price elasticities of residential gas demand range between -0.48 and -0.71.

In contrast to these estimates, a recent study done by Auffhammer & Rubin (2018) finds that the short-term price elasticity of residential gas demand is even more inelastic. Instead of covering the whole US the study focuses on California in particular. The estimates for the price elasticity of gas demand range between -0.17 and -0.23. The study by Auffhammer & Rubin (2018) is especially interesting because it is the only previous study that also relies on data from energy bills. The data analysed covers the period 2003-2014 and consists of monthly energy bills at the household-level. Although the data is similar in principle, there are two structural differences with the energy bills used in this thesis. First, gas tariffs in the US can change monthly rather than annually. Thus, the dataset includes more frequent observations, but may also suffer from more simultaneity of price and demand due to less regulation of the market in the US. Furthermore, another difference is that the data from Auffhammer & Rubin (2018) corresponds to individual households and not to the aggregate building-level.

In addition, the study by Leth-Petersen & Togeby (2001) provides evidence from Denmark. They rely on a panel data approach and analyse data from 1984-1995 on apartment buildings with a floor area of more than 1,500 square meters. Similar to this study they also use the building-level as the level of analysis. As an empirical strategy, they rely on fixed effects models and find highly inelastic demand responses for space heating. For oil, the short-term price elasticity is found to be -0.08 and for district heating -0.02, meaning almost no demand response to a changing price at all.

Another study by Meier & Rehdanz (2010) focuses on the United Kingdom (UK) and analyses a household-level social survey panel for the 15-year period 1991-2005. For their model they use a log-linear approach. They find that the short-term price elasticity of demand for gas ranges from -0.34 to -0.56, depending on the specification. For residential oil demand, the results are in the same order of magnitude, but narrower, between -0.40 and -0.49. In addition, they find that the type of occupant is a relevant dimension for a different price responsiveness. While renters show a higher responsiveness to price changes, homeowners showed a more inelastic demand response.

Synthesis of price elasticity estimates from the literature

When summarizing the literature on the price elasticity of residential space heating demand presented previously, commonalities and differences emerge. The most important commonality is that the price elasticity of energy demand in general and household space heating demand in particular is inelastic in almost all empirical estimations. However, there are differences in the magnitude of the estimates between the individual studies. While the meta-results by Labandeira et al. (2017) as well as a number of individual studies indicate a strongly inelastic demand response (e.g. Auffhammer & Rubin, 2018; Leth-Petersen & Togeby, 2001), other individual studies still indicate an inelastic but higher demand response (e.g., Alberini et al., 2011; Meier & Rehdanz, 2010; Metcalf & Hassett, 1999; Rehdanz, 2007; Schmitz & Madlener, 2020; Schulte & Heindl, 2017). These differences in magnitude may be partly due to the different local and temporal conditions of the studies or due to differences in the data and methodology.

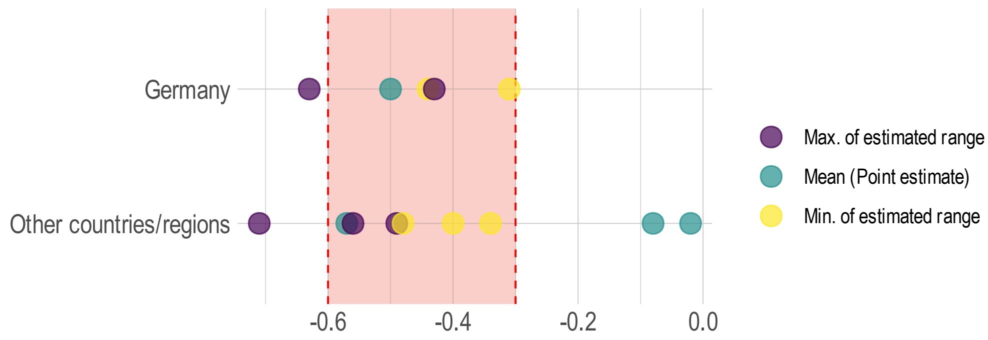

Figure 2.2: Price elasticity estimates for space heat demand from individual studies

Figure 2.2 graphically summarizes the short-term elasticity estimates from the individual studies on space heating demand. All studies that focused on Germany found somewhat more responsive price elasticities ranging between -0.3 and -0.6.[3] And also the majority of estimates from other countries or regions is reflected by that range for the price elasticity.

Another common feature that links all the individual studies presented is that they do not focus on one-off extreme price shocks as an identification strategy, but on gradual price developments. They are mostly built around a panel data set covering a longer period of time. The period examined in this study also covers more than a decade with rather gradual price developments. This should therefore be considered an important condition that facilitates the comparability of the results of this study with the estimates presented in the literature.

2.3 Conceptual Model

The previous two parts of the Chapter presented the theoretical foundations on the price elasticity of demand (see chapter 2.1) and the previous evidence from the literature on household demand response for space heating to energy price changes (see

[3] Note that the extreme results of Rehdanz (2007) for oil as the carrier type are omitted here. They rely on a cross-sectional regression analysis with only two isolated time periods considered. Therefore, the estimates must be considered more prone to extreme results than is the case with a longer panel.

chapter 2.2). To synthesize this basis, this final part of the Chapter is dedicated to developing a conceptual model relying on both theory and approaches from previous studies. The aim is to identify relevant determinants for the level of space heating demand of private households and to justify their relevance. The conceptual model will then be transferred into a statistical model and empirically examined in the further course of this thesis.

Price elasticity of demand

The starting point for the development of the conceptual model is the relationship between energy price and space heating demand, the assumption from the theory of price and demand being that the demand for space heating is influenced by the changes in the energy price. If the energy price goes up, energy demand moves down along the demand curve and vice versa (see Figure 2.1). Expressed in a statistical terminology, this implies that space heat demand is the response variable of a model and energy price the main explanatory variable. In order to support the development of the conceptual model visually, Figure 2.3 depicts it as a directed acyclic graph (DAG). At the center of the DAG, the response variable space heat demand (yellow bubble) is shown, the variance of which is to be explained. The effect of energy price (purple bubble) on space heating demand represents the *price elasticity of demand* and is therefore highlighted by the red arrow. For the rationale of how the energy price influences energy demand, the temporal dimension is also relevant. More precisely, the question of whether it is the energy price in the current billing period that determines demand levels or the price of the previous billing period, which is communicated to households through the delivery of the previous year's energy bill. In this thesis, I follow the approach used in the majority of previous studies and employ the energy price in the same period as the standard approach. However, I will also present results for the case where the energy price of the previous billing period is used as the the basis for demand decisions in order to get a more complete picture.

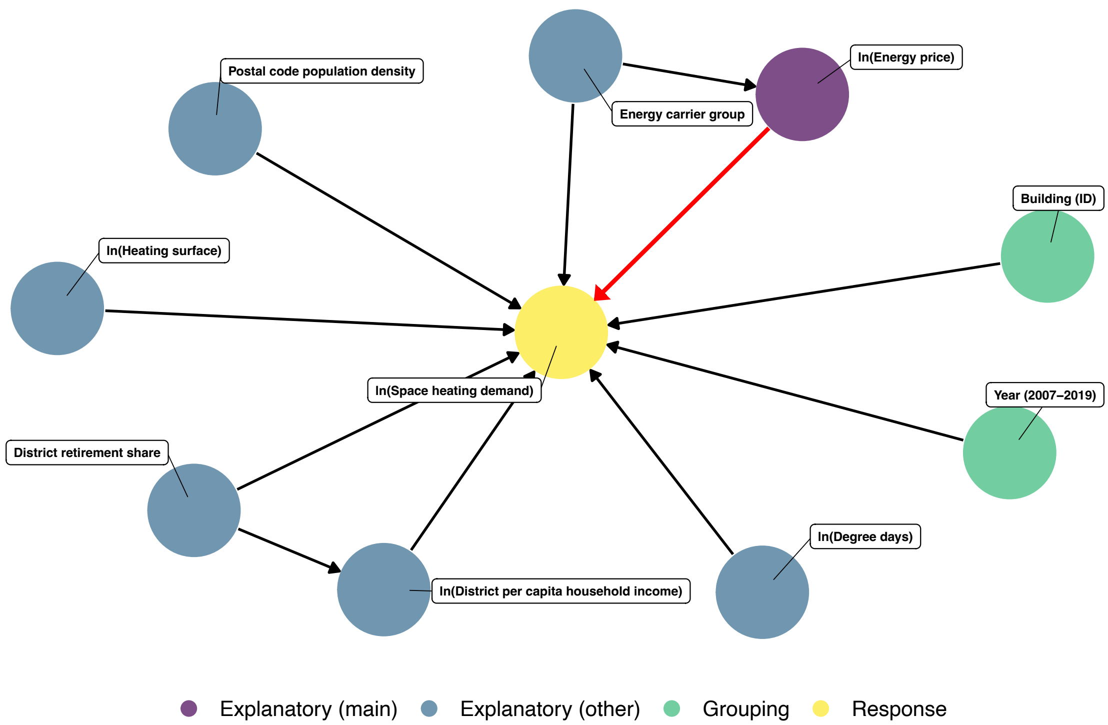

Figure 2.3: Conceptual model as directed acyclic graph

In addition to energy price, there are an array of other determinants that may influence the level of space heating demand. These are shown in the DAG as additional explanatory variables (blue bubbles) and as grouping variables (green bubbles). To determine the set of additional relevant variables, a structured approach was used. First, the recent study by Schmitz & Madlener (2020) was considered as a suitable starting point to create a list of variables. Second, the other individual studies presented previously (see Chapter 2.2) were then used to cross-validate the initial list of variables as well as to screen for further ones. This structured approach led to the variables presented and reasoned for in the following.

Climatic conditions

First, the consideration of the local climatic conditions in a billing period is of importance, as they can have a strong effect on space heating demand (Heße, 2020). Due to the high relevance of climatic conditions, all individual studies presented in Chapter 2.2, except Schulte & Heindl (2017), consider and control for them. For the practical implementation, climatic conditions are usually approximated by the outdoor temperature over a defined period of time and aggregated to periodical (heating) degree days. The basic intuition for the influence of varying climatic conditions is that with lower outdoor temperatures, transmission heat losses increase and thus the demand for space heating increases to compensate for indoor heat losses. In line with VDI (2013), degree days in this study are defined as the temperature difference between a room temperature of 20°C and the daily average outdoor temperature, provided it is below a heating limit of 15°C.[4] Lower mean outdoor temperatures are associated with a higher aggregated number of degree days in a period. Consequently, as a rationale for the direction of the effect of degree days, one would expect a higher space heating demand when degree days are higher in an annual heating period and vice versa. Varying degree days occur on a temporal (one period is warmer/cooler than another) and on a spatial scale (places at higher altitudes or in the interior of the continent have structurally lower temperatures).

Building-level characteristics

In addition, two building-level characteristics are included in the model as additional explanatory variables. First, the size of a building can influence the heat demand, as in larger buildings the ratio between the external surface of the building and the heating surface decreases. Thus, the relative share of transmission heat losses per unit of heating surface is lower correspondingly. Following this rationale, it must be assumed that in buildings with a larger heating surface – which approximates the size of a building – the space heating demand per unit is lower if all other factors are to remain constant.

Second, the type of energy carrier used for the heating system is another building-level characteristic, which is considered in almost all previous studies (see Chapter 2.2). The intuition for the inclusion of the energy carrier as a categorical control variable (oil, gas, district heating) is that different types of energy carriers are associated with different heating technologies, which may lead to structural differences in energy use and efficiency. This effect may also be further amplified by the fact that certain energy carriers have been more commonly introduced in certain time periods in the past. This means that, for example, an average oil heating system may be older and thus less energy efficient than an average gas heating system due to technological

[4] It should be noted that the definition of degree days used here differs from the internationally frequently used definition of heating degree days (HDD), which is calculated by difference between the average outdoor temperature below a heating threshold over time – regardless of an additionally defined room temperature. In this thesis, the VDI based definition of *degree days* was chosen, as this is the common methodology in Germany. It is, for example, also used to create a comparable scoring for building energy performance certificates (see Halbig & Namyslo (2014)).

progress.

District socio-economic variables

Besides building-related characteristics, socio-economic variables may also affect the level of space heating demand. In the literature that draws its data from social surveys, household-level socio-economic information such as income, number of inhabitants, age, education level or employment status is usually part of the core survey data and therefore considered as control variables in the analysis (e.g., Meier & Rehdanz, 2010; Rehdanz, 2007; Schmitz & Madlener, 2020). This study, in contrast, relies on energy billing data at the building-level, which can provide a more accurate picture of actual energy demand and prices than social surveys, but in return does not provide socio-economic information, neither at the household nor at the aggregated building-level. However, as previous studies provide good arguments that the inclusion of socio-economic variables is relevant, an attempt is made to account for their relevance by using the most accurate data available at the district or postal code level. The approach here is that even if they cannot capture household-level effects, they might nonetheless capture overlapping socio-economic effects at a higher level of aggregation.

The first of the three socio-economic variables included is the per capita household income at the district-level. In line with most of the literature, one would expect lower income to lead to lower heating demand as households have fewer funds at their disposal. However, it could also be that income at the district level is more of an approximation of the general affluence of a district and therefore effects such as lower economic resources available for investment in building efficiency outweigh the budget perspective, possibly leading to an opposite outcome. Secondly, retirement share is considered as an additional socio-economic variable at the district-level. Here, the intuition is that with high retirement rates in a district more people spend more of their time at home which may lead to a rise in space heat demand. Lastly, the population density within a postal code area is considered to observe if there may be structural differences in heating demand between densely populated metropolitan areas on the one hand and sparsely populated rural areas on the other.

Chapter 3

Methods

3.1 Data Description and Pre-Processing

To translate the conceptual model set up in Chapter 2.3 into a statistical model that can be used to empirically investigate the price elasticity of space heating demand, data from multiple sources were combined. In the following the variables and the pre-processing of the data is presented. An overview of all variables and their respective data source is given in Table 3.1. First, the focus is on the energy billing dataset and its variables and then on variables retrieved from supplementary data sources.

Energy billing dataset

The key dataset is a large-scale building-level sample of energy bills made available though the Climate Policy Department at DIW Berlin. The data originates from the energy- and billing service provider ista Deutschland GmbH and was provided for scientific use.[1]

The sample contains information on almost 4.5 million annual billing observations for multi-apartment residential buildings in Germany between 2003 and 2019. Since energy prices are first included in the sample in the year 2007, the period under investigation ranges between 2007 and 2019. The smallest buildings observed in the sample are buildings with two apartments. Single-family houses are not observed. The observations in the dataset represent annual heating bills at the building-level. Most importantly, the billing dataset provides the two key variables for the analysis,

[1] The billing data is classified as sensitive. Access was exclusively via DIW Berlin's internal servers. For data protection reasons, it is not possible to make the data available to external parties for the purpose of reproducing the results.

namely space heating demand and energy prices. In addition, the sample also provides other variables relevant for the analysis, such as the building heating surface, the type of energy carrier used within a billing period, as well as a building ID and the billing dates, which are suitable for creating a panel structure.[2]

Space heat demand (kWh/sqm/y): The demand data in the billing dataset is provided as total energy consumed per building and per annual billing period in kilowatt hours (kWh). In order to isolate the share of energy consumed for space heating, the share of energy consumed for hot water generation is deduced for those buildings in which hot water is also provided by the central heating system. In a second step, the total space heating demand at building level is divided by the building heating surface in square meters (sqm) to obtain a comparable metric for the differently sized buildings in the sample. This results in the annual space heat demand per square meter as the demand variable (kWh/sqm/y).

Energy prices (Cents/kWh): The structure of the price data is similar to that of demand. The energy price data is provided as total annual costs for the heating system at the building-level in Euros. Again, relying on the supplementary information on the information on relative shares of energy use for the two purposes of space heating and hot water generation in buildings where hot water is generated by the central heating system, the cost share for space heating can be isolated. In a second step, the metric is then translated into average per-unit costs by dividing the building-level expenditures for space heating by the respective demand for space heating in kWh. This yields a per-unit energy price variable (Cents/kWh).

As energy prices are provided as nominal values, they are deflated using the German annual consumer price index (CPI) with 2015 as the index year (DESTATIS, 2021b). The use of the CPI to deflate energy prices represents the standard procedure in the literature, which is also used by Schmitz & Madlener (2020), for example. The deflated real energy prices represents the variable used in the analysis.

Energy contracts for residential buildings in Germany heated with gas and district heating usually involve a fixed cost component (flat-rate basic charge) and a variable

[2] It should be noted that DIW Berlin also has an additional dataset with information from Energy Performance Certificates (EPC) on a subset of the buildings in the billing dataset. In the preparatory phase, the idea was to use additional information from the EPC dataset. However, since the data was only available for about 15% of the observations in the billing dataset and the data systematically favored larger buildings, this idea was ultimately abandoned. However, the pre-processing steps on the EPC data are preserved in repository on GitHub and may be used by the DIW Climate Policy Department for possible further analyses.

cost component (demand-based price per kWh).[3] Buildings with oil heating also incur some fixed costs, for example for heating system maintenance, but these are lower than for gas and district heating. Yet, the total cost variable provided in the billing data does not contain a breakdown of fixed and variable costs. There is much debate in the literature about whether the demand response of households depends on the average or on the marginal price for energy.I would argue that the use of marginal prices is compelling when the variable price component is dynamically adjusted by the utility provider based on the quantity consumed (*block pricing*) or at a higher frequency based on market prices, as is the case with many residential contracts e.g. in the US (see Auffhammer & Rubin, 2018). But since those factors do not apply to residential energy contracts in Germany, I assume the use of the average per-unit energy price to be a suitable approach. Even though the relative shares for the fixed and the variable cost component remain unknown. Use of the average costs is also the common approach in most of the literature (e.g., Metcalf & Hassett, 1999; Rehdanz, 2007; Schmitz & Madlener, 2020).

(Billing) Year: The billing data also contains the exact dates of the billing period of an individual building. While for most observations the billing period reflects the calendar year, some buildings have billing periods that are split over two years. Therefore, the start and end dates are used to assign these observations to years. The criterion is which of the two years counts the majority of days.

Energy carrier group: There are three main types of energy carriers included in the dataset: Gas, oil, and district heating. To obtain these three carrier groups, energy carrier descriptions from the dataset are grouped into the categories (e.g, *Gas low* and *Gas high* are grouped under *Gas*). In addition, all other heating carriers which have only limited occurrence in the sample (e.g., electricity for heat pumps, coal) are grouped under the category *Others* but later not further considered in the regression analysis due to the small number of observations and difference between carrier types within the group (see Chapter 3.3).

Heating surface (sqm): Information on the heating surface of a building is given in square meters at the building-level. The heating surface does not correspond to the total floor area of a building, as it excludes the unheated surface areas within a building (e.g., corridors or basements). The heating surface area was preferred to the number of apartments within a building in order to approximate the size of

[3]On the basis of exemplary heating bills provided by ista GmbH, it was possible to establish that this was also the case for the buildings in the billing sample.

Table 3.1: Variables and data sources

Variable		Type of variable	Unit	Data source
Demand	Space heating demand	Continuous	ln(kWh/sqm/y)	Ista (billing dataset)
Price	(Average) Energy price, real	Continuous	ln(Cents/kWh)	Ista (billing dataset), DESTATIS (2021b)
Additional billing information	Heating surface	Continuous	ln(sqm)	Ista (billing dataset)
	Energy carrier group	Categorical	Gas / Oil / District heating	Ista (billing dataset)
	Building ID	Categorical	-	Ista (billing dataset)
	Year	Categorical	-	Ista (billing dataset)
Climatic conditions	Degree days	Continuous	ln(Degree days)	IWU (2021)
Socio-economic factors	District per capita household income	Continuous	ln(Euros/y)	Statistische Ämter (2021)
	District retirement share	Continuous	%	DESTATIS (2021a)
	Postal code population density	Continuous	ln(Inhabitants/sq. km)	OSM (2021a), OSM (2021b)

a building. The reason for this is that the heating surface area better reflects the relationship between the indoor and outdoor surface area of a building, as it does not suffer in its accuracy from the different apartment sizes between individual buildings.

Variables from supplementary data sources

Degree days: Data on degree days is retrieved from an Excel tool published by the *Institut für Wohnen und Umwelt (IWU)* at the postal code level (IWU, 2021). The degree days data from the IWU-tool builds on daily temperature data from the 800 weather stations of the German Meteorological Service (DWD) that are aggregated on a monthly basis. In the tool, the settings of a mean room temperature of 20°C and a heating limit of 15°C are chosen to reflect the VDI based definition of degree days (VDI, 2013). Furthermore, the option to assign the postal code area to the three nearest DWD weather stations with importance weighting according to geographical distance is chosen. Relying on more than one station reduces the risk of possible distortions from differences in altitude between a weather station and the centroid of a postal code area. The extracted monthly degree days per postal code area are aggregated to annual periods on a rolling basis and allocated to the energy bills on a same-month basis. The postal code level was chosen because it corresponds to the spatial information on the location of the buildings contained in the billing dataset and thus represents the most accurate allocation possible.

District per capita household income (Euros/y): The per capita disposable income of private households provided by the joint statistics portal of the National Government and the Federal States is used as the income variable at the district-level (Statistische Ämter, 2021). The figure represents the primary income of private households, minus transfers paid and plus transfers received. Disposable income is chosen because it can be considered the most suitable indicator for funds available for households. The district-level is chosen as it is the most granular household income statistics publicly available.

District retirement share (%): To construct the district retirement share variable, population data at the district-level with a segmentation by age groups is used (DESTATIS, 2021a). As an approximation for the actual proportion of retirees, the percentage of persons within a district and year who are 65 years of age and older is calculated. This approach was chosen because no detailed statistics are available on the number of people in retirement at the district-level. The boundary value of 65 was chosen as the age group closest to the current retirement age in Germany and since the same threshold was also chosen in the literature (e.g., Alberini et al. (2011))

Postal code population density (Inhabitants/sq. km): To generate the population density variable, data from Open Street Maps with pre-assigned population figures to postal code areas based on the 2011 Census are used (OSM, 2021b, 2021a). The postal code level is chosen because the higher granularity of data was available and use of population density at the district-level might obscure heterogeneity within districts. Especially when it comes to districts that include both a city and rural areas. To create the population density variable, the number of inhabitants per post code area is divided by the base area in square kilometers (sq. km), which gives the population density as the average number of inhabitants per sq. km in a post code area.

3.2 Empirical Considerations

Causality with observational data

The conceptual DAG presented in Chapter 2.3 and the rationale derived for the cause-and-effect relationship between energy price and space heat demand in an environment where the other explanatory variables also exert an effect on energy demand are all based on the assumption of causal inference. Causal inference can be described as indicating that an observed relationship between two variables is reflected by a causal link and not just mere correlation (Holland, 1986). The analysis in this thesis is build on externally provided observational data. Using observational data to draw causal inference about a treatment effect – in the given case, inference about the price sensitivity of private households for space heating demand – is inherently difficult since the treatment is not controllable and therefore cannot be randomly assigned (Nichols, 2007). Since an experimental research design, which would arguably provide the most unbiased source of evidence, is not feasible for this type of research, it is all the more important to point out potential sources of bias in the estimation process relying on observational data and address those biases where possible.

Potential sources of bias

Reducing potential sources of bias was approached in several steps. First, the set of explanatory variables already described was used to reduce potential bias from omitted variables. This means that by taking into account, for example, degree days or the heating surface of a building, additional relevant effects on space heating demand are included that might otherwise be wrongly attributed to energy price. That being said, there is also a gap in the data related to omitted variables that cannot be fully remedied. Due to the aggregated nature of the billing dataset at the building-level, it is not possible to integrate detailed household-level socio-economic variables (e.g., household income, number of inhabitants, age, education level, or employment status) into the analysis. In previous studies, which build on household-level social surveys, socio-economic variables have been found to be a relevant determinant of energy demand. In particular, household income is a variable for which a relevant effect has been demonstrated (e.g., Meier & Rehdanz, 2010; Rehdanz, 2007; Schmitz & Madlener, 2020). To at least partially fill the gap of household-level socio-economic variables, this thesis employs district or postal code level socio-economic controls to attempt to capture the underlying socio-economic effects at a higher level of aggregation.

Besides omitted variables, errors in the data can be another source of potential bias. In general, the use of observational data where the researcher was not involved in the process of data generation should not be seen as a mechanistic process, but should be based on the application of domain-specific knowledge to critically evaluate and scrutinize the available data. The use of domain-specific knowledge and expert judgement was important in cleaning and processing the energy billing dataset. When assessing the energy price variable from the billing dataset, outliers are found that can be traced back to be placeholders that do not reflect the actual energy demand of a building. If these observations had not been removed from the analysis, they would have distorted the result of the analysis. The process is described in detail later in this Chapter.

In addition, it was assessed if the variables used in the analysis incorporated any systematic bias from their measurement. In principle, several variables are subject to inaccuracies in their measurement. For example, the degree days data is based on a spatial interpolation of the values from the three nearest DWD weather stations to the centroid of a postal code area, which introduces a first layer of inaccuracy. It is then assumed that the degree days assigned to a postal code area apply to every building within that area, even though there may be greater distances and differences in elevation between the centroid and a building's exact location. However, inaccuracies in the measurement – such as those described exemplarily for the degree days variable – do not necessarily present a problem, since they do not introduce a systematic error (bias). Rather, the described inaccuracies indicate a random effect, which is less serious in general and especially given the very large number of observations present in the sample as it can be assumed that they balance out.

Simultaneity problem

Furthermore, when estimating price elasticities of demand, one additional relevant challenge for the identification of a causal effect is the potential simultaneous determination of price and demand. Under the assumption that demand and supply curves shift over time, the observed data on quantities and prices would reflect a set of equilibrium points on both curves (Angrist & Krueger, 2001). The common approach in the economic literature to address the endogeneity steaming from simultaneity is to pursue an instrumental variable approach (Angrist & Krueger, 2001; Burtless & Hausman, 1978). Using an instrumental variable approach was also explored for this thesis. In line with the approach used by Auffhammer & Rubin (2018), the Trading Hub Europe (THE) Gas forward prices were considered as an instrument. However, use of an instrumental variable approach was ultimately discarded because the identifying assumption that the instrument needs to be highly correlated with the energy prices in the sample was not met. This is mainly because the prices observed in the large-scale sample are more scattered than average energy prices, where instruments have been mainly used in the past (e.g., Csereklyei (2020)). However, unlike in the case of Auffhammer & Rubin (2018) residential energy contracts in Germany are more regulated than in the US where, for example, the rates faced by consumers change on a monthly basis—updating as a function of gas wholesale prices paid by the retail utilities. So while not being able to pursue an instrumental variable approach remains an issue to potentially producing biased estimates, it appears less of an issue with the data available for this thesis.

3.3 Workflow and Empirical Approach

For the regression analysis, I use several statistical methods. I first conduct an analysis of the full sample to estimate an overall price elasticity of demand. Given the large number of observations in the sample, I rely on classical statistical methods to conduct the analysis of the full sample. Subsequently, I deepen the analysis based on a stratified random subsample. Here, I rely on Bayesian regression analysis to better communicate the uncertainty inherent to the estimates as well as to investigate potential factors for heterogeneity in the overall elasticity estimates. Generally, I follow an approach where I begin with constructing a simple model Specification and then extend it to a more elaborate statistical model. The aim is to arrive at a model that has a good fit, but also makes the way to get to it transparent.

Processing of billing sample and matching with supplementary data

However, before the regression analysis can be performed, the billing dataset must be cleaned in several steps, which are shown in the upper part of Figure 3.1. After data cleaning, 2,719,270 of the original 4,494,943 annual billing observations (60.5%) remain for the analysis of the full sample. The criteria used in the cleaning process are ordered by the number of observations removed.

The most important reason for exclusion from the sample is the absence of price data. In the period 2003-2006, price data are not available for any observation. This leads to the exclusion of 909,458 observations. In addition, price data is also missing for a part of the observations in the period 2009-2019. Due to these gaps, a further 543,544 observations are excluded from the dataset. Another 233,888 observations are removed from the sample because they are associated with buildings that occur only once in the sample. The decision to apply the criterion that buildings must occur at least twice in the sample is mainly due to the fact that singular building IDs occur mainly in the period 2007-2008, suggesting that IDs from this period are not always linked with observations in later years. In addition, observing a building at least twice allowed cross-validation through matching building attributes and thus better reliability of the data. Another 53,674 observations are removed from the sample because the buildings are heated with an energy carrier other than gas, oil, or district heating. Due to the small number of observations compared to the total sample size and due to the heterogeneity of the group (e.g., heat pumps but also coal heating), it was decided to exclude them. In addition, another 1,722 observations are excluded as duplicates (based on matching building ID and year) and 65 observations due to a deviating length of the billing period (deviation of more than +/- 10 days).

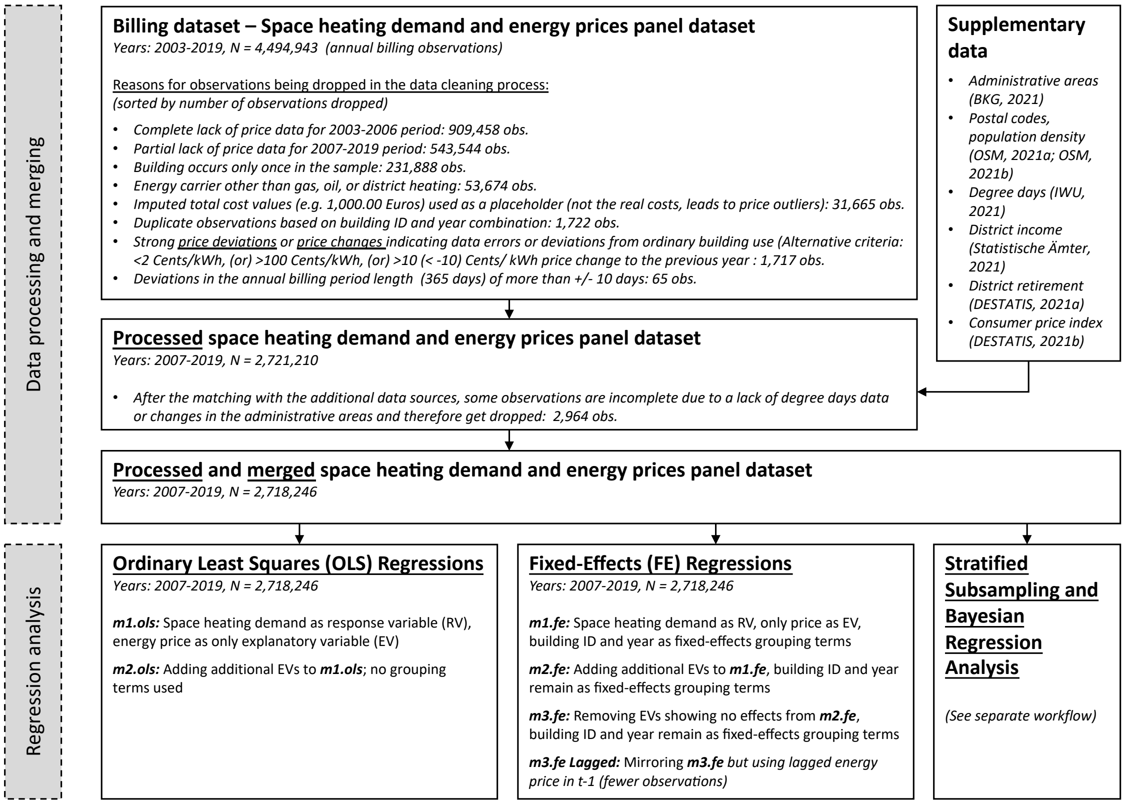

Figure 3.1: Workflow for data processing and regression analysis based on full sample

Removal of outliers representing data errors in the energy price variable

In addition to the factors described above, there are two other reasons for excluding observations from the sample. Both are related to outliers in the energy price data. Conducting the analysis with and without these two cleaning steps for the price data showed that they can have a strong influence on the price elasticity estimates and are therefore particularly important.

First, it was determined through exploratory descriptive analysis of the price variable and confirmed by ista GmbH that the price data contains fictitious cost values (e.g., 1,000.00 Euros for the entire building) that are used as placeholders and do not reflect actual costs.[4] These fictitious cost values arise because ista, as a billing service provider, does not always have information about the costs incurred by the

[4]The internal practice of ista was confirmed on the basis of an email exchange for the Wärmemonitor 2019.

building owners and therefore uses the placeholders for internal technical reasons. The actual costs are only entered later by the building owners and therefore do not appear in the provided data. This practice leads to the formation of outliers in the energy price variable, as the fictitious cost values do not necessarily reflect the size and characteristics of a building. A total of 31,665 observations are removed from the sample due to fictitious cost values identified by searching for round cost rates without cent amounts that occur unusually often in the sample.

Second, despite the removal of the fictitious cost values, some spurious outliers remained in the price data. An exploratory review of individual cases revealed that the remaining price outliers were likely due to other types of data errors (e.g., unrealistic heating surfaces), which is not unusual given the very large number of observations in the sample. Although a relatively small number, the price outliers can significantly interfere with the results of the analysis. Therefore, it was decided to also exclude the remaining observations that showed unrealistic deviations from the reasonable price level expected during the time period under investigation (Real prices: <2 Cents/kWh or >100 Cents/kWh) or large price changes from one period to another (Change in real prices: >10 or < -10 Cents/kWh compared to the previous billing period). The exclusion thresholds are based on expert judgment and are chosen to only exclude highly implausible observations. Use of the described thresholds for absolute prices and relative price changes resulted in the exclusion of an additional 1,717 observations from the sample.

After scrutinizing the reliability of the price variable, the processed billing dataset is then matched with the external data from supplementary sources (see Figure 3.1). Due to missing degree days data and changes in the administrative areas, a further 2,964 observations are removed from the sample in this step, which ultimately leads to 2,718,246 observations remaining for the regression analysis.

Empirical approach for analysis of full sample

Given the large number of observations in the sample, I rely on classical statistical methods to conduct the analysis of the full sample. To begin with, I formulate a simple cross-sectional multiple linear regression (MLR) model based on ordinary least squares

(OLS) (Bailey, 2017; Wooldridge, 2013):

$$\begin{aligned} ln(Demand) = \alpha + \beta_1 \cdot ln(Price) + \beta_2 \cdot ln(Degree.days) + \\ \beta_3 \cdot ln(Heating.surface) + \beta_4 \cdot Carrier.group.oil + \\ \beta_5 \cdot Carrier.group.district.heating + \beta_6 \cdot ln(District.income) + \\ \beta_7 \cdot District.retire + \beta_8 \cdot ln(Pop.density) + \varepsilon \end{aligned}$$

where the response variable $ln(Demand)$ denotes the natural logarithm of the annual space heating demand per square meter and the main explanatory variable $ln(Price)$ denotes the natural logarithm of the average energy price (Cents/kWh) in the same period. The additional terms on the right hand side of the equation denote the additional explanatory variables (degree days, building characteristics, and district/postal code socio-economic variables). ε denotes the error term of the model. Referring back to Equation (2.3) in the theory section of this thesis, the ln-form of both sides of the equation transforms the model into the elasticity case, meaning that the coefficient β_1 of $ln(Price)$ represents the price elasticity of space heating demand. For the cross-sectional OLS regression, I first run the model with $ln(Price)$ as a sole predictor and then with the additional explanatory variables included. All model predictors except the energy carrier group, which is a categorical variable, and the share of district heating supply, which is already a ratio, are included in the ln-transformed form. This is done firstly to facilitate the interpretation of the results in the elasticity case, but secondly also to reduce the impact of possible outliers in the distribution of the variables. The use of ln-transformed continuous predictors is also the standard approach in the previous literature (e.g., Auffhammer & Rubin, 2018; Meier & Rehdanz, 2010; Schmitz & Madlener, 2020).

In the billing dataset, it is possible to assign several observations to the same building (building ID) and year (billing period). Thus, the data structure allows to go beyond cross-sectional OLS and apply panel data estimation techniques. In the OLS estimation shown above, observations from various buildings and years are treated to be systematically no different from each other. Panel data methods, on the contrary, are rooted in the assumption that there may be systematic and unobserved differences between units that may be correlated with observed predictors whose effects on the response are to be measured (Wooldridge, 2013). Thus, panel data methods are considered a powerful tool for observational data where controlling for all relevant factors is inherently difficult (Bailey, 2017). By acknowledging and addressing that there may be unobserved inter-individual differences between the units (buildings) and also an intra-individual dynamic over time (years), use of panel data methods can provide a more accurate picture for the predictors observed in the model.

For the subject at hand, it is likely that the demand for space heating observed is impacted by various unobserved building-specific constant or semi-constant factors, such as a building's energetic condition or its usage properties. Thus, inferences drawn from cross-sectional data are likely to be invalid since building-specific effects are falsely attributed the observed model predictors – including energy price. The same applies for temporal trends. For example, new buildings with higher energetic standards could be added to the sample over the course of the period under investigation and old buildings might drop out of the sample due to demolition. Such factors, which affect the structural composition of the sample over time, may result in space heating demand in the later periods being systematically lower than that in the earlier periods. In the empirical literature, switching from a cross-sectional model to the use of unit-level fixed effects with panel data resulted in a more inelastic estimates for the price elasticity of space heating demand (Miller & Alberini, 2016).

In addition, the choice of panel estimation method should provide more reliable results even when strict exogeneity fails to hold. Therefore, I adopt a two-way fixed effects (FE) model, which is considered one of the more robust estimation methods and is also a commonly used approach in the prior literature on the price elasticity of space heating demand (e.g., Lange, Moro, & Traynor, 2014; Meier & Rehdanz, 2010; Schmitz & Madlener, 2020). The FE model takes the following form:

$$\begin{aligned} ln(Demand_{i,t}) = \alpha + \beta_1 \cdot ln(Price_{i,t}) + \beta_2 \cdot ln(Degree.days_{i,t}) + \\ \beta_3 \cdot ln(Heating.surface_{i,t}) + \beta_4 \cdot Carrier.group.oil_{i,t} + \\ \beta_5 \cdot Carrier.group.district.heating_{i,t} + \beta_6 \cdot ln(District.income_{i,t}) + \\ \beta_7 \cdot District.retire_{i,t} + \beta_8 \cdot ln(Pop.density_{i,t}) + \\ \gamma_{building[i]} + \delta_{year[t]} + \varepsilon_{i,t} \end{aligned}$$

The FE model structure extends the cross-sectional OLS model. The response variable $ln(Demand_{i,t})$ again denotes the natural logarithm of the annual space heating demand per square meter but extended by the indices on building i and the annual time period t. The same applies to $ln(Price_{i,t})$ as the average energy price (Cents/kWh) and the set of additional explanatory variables. The newly added term $\gamma_{building[i]}$ denotes the time-invariant building fixed-effects and $\delta_{year[t]}$ the annual time fixed-effects. $\varepsilon_{i,t}$ again denotes the error term.

In line with the approach for the OLS cross-sectional model, I first run the model without the additional explanatory variables to obtain an estimate for the case where only the energy price is used to explain demand, and then include the set of additional explanatory variables in a second specification. Additionally, for the FE model, I also drop variables with limited explanatory power to arrive at a more condensed and relevant model specification. Furthermore, I also run a model where $ln(Price_{i,t)}$ is substituted by $ln(Price_{i,t-1})$, so taking the lag of the price variable in $t-1$ instead of the price in the same period to explain the space heating demand in period t.

All the models described represent the conditional demand. It therefore has to be inferred that the results reflect the short-term price elasticities of demand. The use of other dynamic estimation approaches to estimate long-term elasticities, as for example Alberini et al. (2011) or Csereklyei (2020) have done, would go beyond the scope of this thesis and are therefore not pursued. Additionally, the short-term price elasticities estimated here may also be viewed as lower bound estimates of the price responsiveness compared to the usually higher long-term estimates. As they are a lower bound estimate, they can also be considered as a conservative estimate for any kind of modelling or policy advice for which they might be used.

To estimate the OLS models, I rely on the lm-function in R. To estimate the FE models, I apply the felm-function from the lfe-package (Gaure, 2013).

Stratified subsampling and Bayesian regression analysis

After using classical statistical estimation techniques on the full sample, the analysis moves to the application of Bayesian regression analysis. In general, use of Bayesian inference has the advantage of allowing for the propagation of uncertainty in the modelling process (McElreath, 2020). Furthermore, the Bayesian framework allows for a more intuitive interpretation of modelling results being the actual chance of an event happening rather than the probability that the same outcome would occur if one were to replicate the data (McElreath, 2020). A further advantage of Bayesian inference is the logic of updating beliefs based on evidence (data) and the potential integration of prior knowledge (Gelman, Hill, & Vehtari, 2021). Here, I refrain from making strong assumptions on the prior distribution of model parameters and instead rely on weakly informed priors because enough data available. Additionally, it should be pointed out that Bayesian inference is more computationally demanding than the classical statistical methods presented above. This is the reason why

Bayesian inference is only used on a subsample and not for the analysis of the full sample.

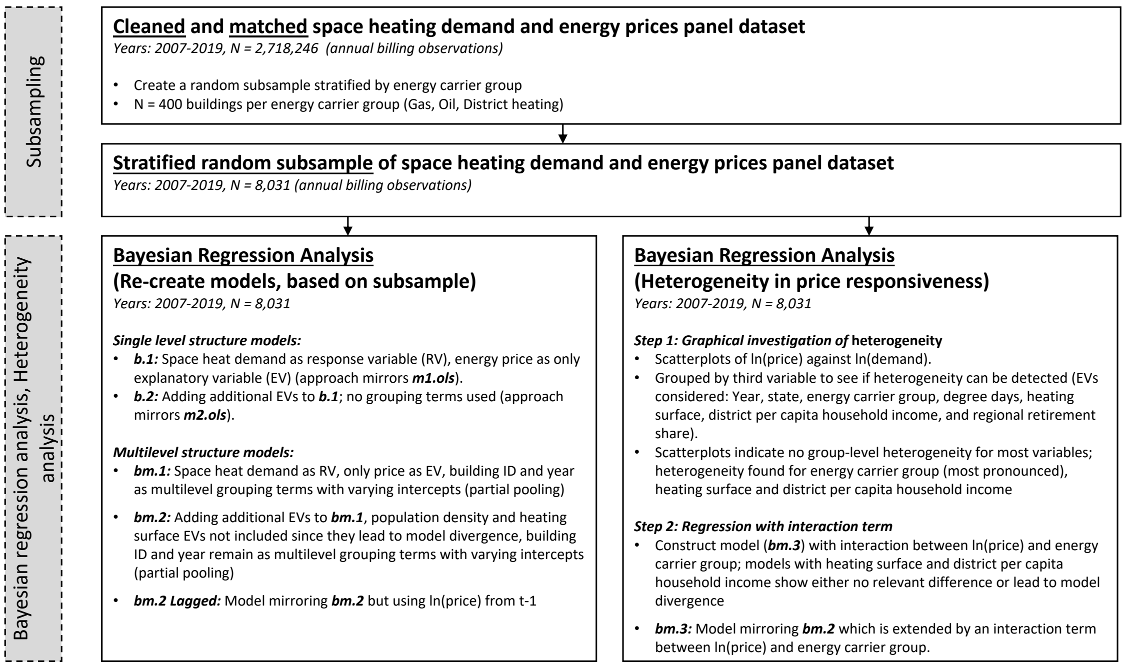

Figure 3.2: Workflow for stratified subsampling and Bayesian regression analysis based on subsample

In Figure 3.2 the second part of the workflow is shown. The subsample is set up as a stratified random selection of 400 buildings per energy carrier group (Gas, Oil, District heating) from the full sample. The number of 400 buildings per energy carrier group is chosen because the sample will remain large enough to draw reliable inference and because running the more computationally demanding Bayesian models with the associated amount of observations is possible with a reasonable time needed for model computation. In delimitation to the full sample analysis, the subsample analysis intends to serve two purposes. First, as mentioned in the previous paragraph, use of fewer data points makes the application of Bayesian methods possible which are well-suited to propagate uncertainty. Second, the subsample is used to investigate potential factors of heterogeneity in price responsiveness in the data first resorting to exploitative visual analysis and then moving to construct an interaction term model. The stratified sampling approach, where all three energy carrier groups are represented equally, allows the investigation of the carrier type as an potential additional source of heterogeneity. Furthermore, having a stronger representation of district heating than in the main sample also allows for a glimpse into the future, where gas and oil heating systems will become less viable given the current gas energy crisis and rising carbon prices, and will therefore be replaced by district heating (or heat pumps [5]).

For the analysis of the subsample, I first run simpler models that mirror the OLS Specifications for the full sample but in a Bayesian setting. When moving to the models integrating the grouping variables building ID and years, I resort to the use of a multilevel partial pooling model with varying intercepts (McElreath, 2020). In contrast to the FE models used for the full sample analysis, where only information from the same units (buildings and years) is considered, multilevel partial pooling models allow for information sharing between units (McElreath, 2020). Use of clusters without information sharing may lead to less reliable estimates, especially if only a few observations are available for a building. Together with the weakly informed priors mentioned earlier, the main model Specification with the multilevel partial pooling structure takes the following form:

$$
\begin{aligned}
ln(Demand_{i,t}) &\sim \text{Normal}(\mu_{i,t}, \sigma) \\
\mu_{i,t} &\sim \bar{\alpha} + \gamma_{building[i]} + \delta_{year[t]} + \beta_1 \cdot ln(Price_{i,t}) + \\
&\quad \beta_2 \cdot ln(Degree.days_{i,t}) + \beta_3 \cdot ln(Heating.surface_{i,t}) + \\
&\quad \beta_4 \cdot Carrier.group.oil_{i,t} + \beta_5 \cdot Carrier.group.district.heating_{i,t} + \\
&\quad \beta_6 \cdot ln(District.income_{i,t}) + \beta_7 \cdot District.retire_{i,t} + \beta_8 \cdot ln(Pop.density_{i,t}) \\
\bar{\alpha} &\sim \text{Normal}(0, 1) \\
\gamma_j &\sim \text{Normal}(0, \sigma_\gamma) \\
\delta_k &\sim \text{Normal}(0, \sigma_\delta) \\
\beta_{1-8} &\sim \text{Normal}(0, 0.5) \\
\sigma, \sigma_\gamma, \sigma_\delta &\sim \text{Exponential}(1)
\end{aligned}
$$

Also for the Bayesian models, the response variable $ln(Demand_{i,t})$ denotes the natural logarithm of the annual space heating demand per square meter in building i in the year t and the main explanatory variable $ln(Price_{i,t})$ denotes the natural logarithm of the average energy price (Cents/kWh) in the same period. The intercept term ($\bar{\alpha}$) denotes the varying intercepts along the two grouping variables buildings ($\gamma_{building[i]}$) and years ($\delta_{year[t]}$). For the model I assume that all parameters follow a Gaussian distribution centered on 0. The intercept term $\bar{\alpha}$ is assumed to have a standard

[5] Heat pumps are not considered here due to the still very limited data available on heat pump use in the billing sample.

deviation of 1. The distribution parameter σ as well as the parameters σ_γ and σ_δ for the varying intercepts parameters were assumed to follow an $Exp(1)$ distribution so that they are limited to the positive values required for the standard deviation.

Investigation of heterogeneity in price responsiveness

The subsample analysis is also concerned with identifying potential sources of heterogeneity in the price response that may remain hidden in an estimation of the overall price elasticity of demand. To this end, a two-step approach is adopted (see Figure 3.2). First, scatter plots between energy price and demand which are grouped by a third variable are used to investigate graphically whether relevant differences for price responsiveness can be found within a set of explanatory variables. In a second step, the heterogeneity found for the energy carrier group variable is then formalized by extending the model Specification shown earlier to include an interaction term between energy price and energy carrier group. The interaction model has the following form:

$$
\begin{aligned}
ln(Demand_{i,t}) &\sim \text{Normal}(\mu_{i,t}, \sigma)\\
\mu_{i,t} &\sim \bar{\alpha} + \gamma_{building[i]} + \delta_{year[t]} + \beta_1 \cdot ln(Price_{i,t}) +\\
&\quad \beta_2 \cdot ln(Degree.days_{i,t}) + \beta_3 \cdot ln(Heating.surface_{i,t}) +\\
&\quad \beta_4 \cdot Carrier.group.oil_{i,t} + \beta_5 \cdot Carrier.group.district.heating_{i,t} +\\
&\quad \beta_6 \cdot ln(District.income_{i,t}) + \beta_7 \cdot District.retire_{i,t} + \beta_8 \cdot ln(Pop.density_{i,t}) +\\
&\quad \beta_9 \cdot ln(Price_{i,t}) \cdot Carrier.group.oil_{i,t} +\\
&\quad \beta_{10} \cdot ln(Price_{i,t}) \cdot Carrier.group.district.heating_{i,t}\\
\bar{\alpha} &\sim \text{Normal}(0, 1)\\
\gamma_j &\sim \text{Normal}(0, \sigma_\gamma)\\
\delta_k &\sim \text{Normal}(0, \sigma_\delta)\\
\beta_{1-10} &\sim \text{Normal}(0, 0.5)\\
\sigma, \sigma_\gamma, \sigma_\delta &\sim \text{Exponential}(1)
\end{aligned}
$$

Since the model represents an extension to the model shown previously, all notations remain the same. Only the additional interaction term between main explanatory variable $ln(Price_{i,t})$ and the categorical energy carrier group is added. Since gas as carrier type serves as the reference category, terms for oil and district heating are shown.

The Bayesian models are estimated using the brms-package by Bürkner (2017) in R. The brms-package serves as an interface to the probabilistic programming language and inference engine Stan (Stan Development Team, 2022).

3.4 Descriptive Statistics

After having established the workflow and empirical approach, the following part of the Chapter presents descriptive statistics on the full sample to obtain a better understanding of the data.

Unbalanced occurrence of buildings

How often an individual building is observed in the processed billing data (full sample) is shown in the histogram in Figure 3.3. Buildings are observed on average 6.77 times during the period under investigation. The histogram shows that not all buildings are observed throughout the whole period, leading to the panel being unbalanced. While a relatively large number of buildings are observed only twice, the distribution exhibits a second peak at nine and ten times observed. This means that information on a relevant proportion of the buildings in the sample is available almost throughout the entire period.

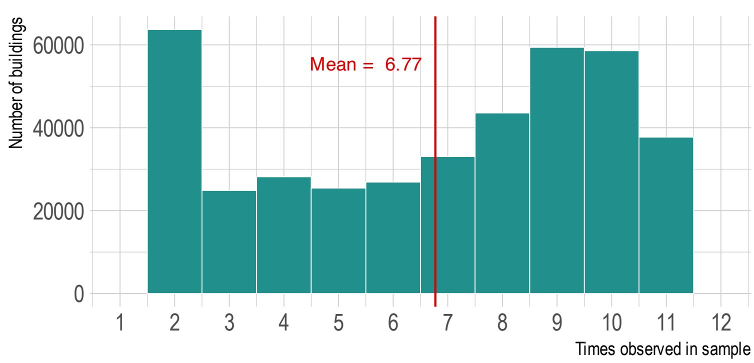

Figure 3.3: Histogram of the occurrence of buildings in the full sample

Spatial and temporal coverage of sample

In addition, also the spatial and temporal distribution of observations in the sample is of relevance. Figure 3.4 graphically depicts the spatial and temporal coverage of the full sample at the district-level.[6] On the spatial dimension, the maps show that the coverage is good. There are very few districts without a building observed (transparent) and only a few districts with less than 10 buildings observed per year (dark blue). For most district-year combinations, more than 100 buildings are observed. In some cities, numbers of more than 10,000 buildings annually are reached (e.g., Berlin). The good spatial coverage of the data indicate that results drawn from the sample have validity for Germany as a whole and are in their explanatory power not limited to certain regions or clusters.

On the temporal dimension, fewer observations are available in 2007 (43,696 observations) and 2008 (43,536 observations), as the energy price data was first included in these years. For the period between 2009 and 2019, an annual minimum of 201,856 and an average of 239,183 buildings are observed. Which means that the explanatory power of the results applies in particular to the period between 2009 and 2019. At the same time, exploratory testing of the data with and without consideration of the years 2007 and 2008 did not lead to a relevant change in the results implying that the results are applicable to the whole period under investigation.

[6] Please note the use of the logarithmic scale in Figure 3.4.

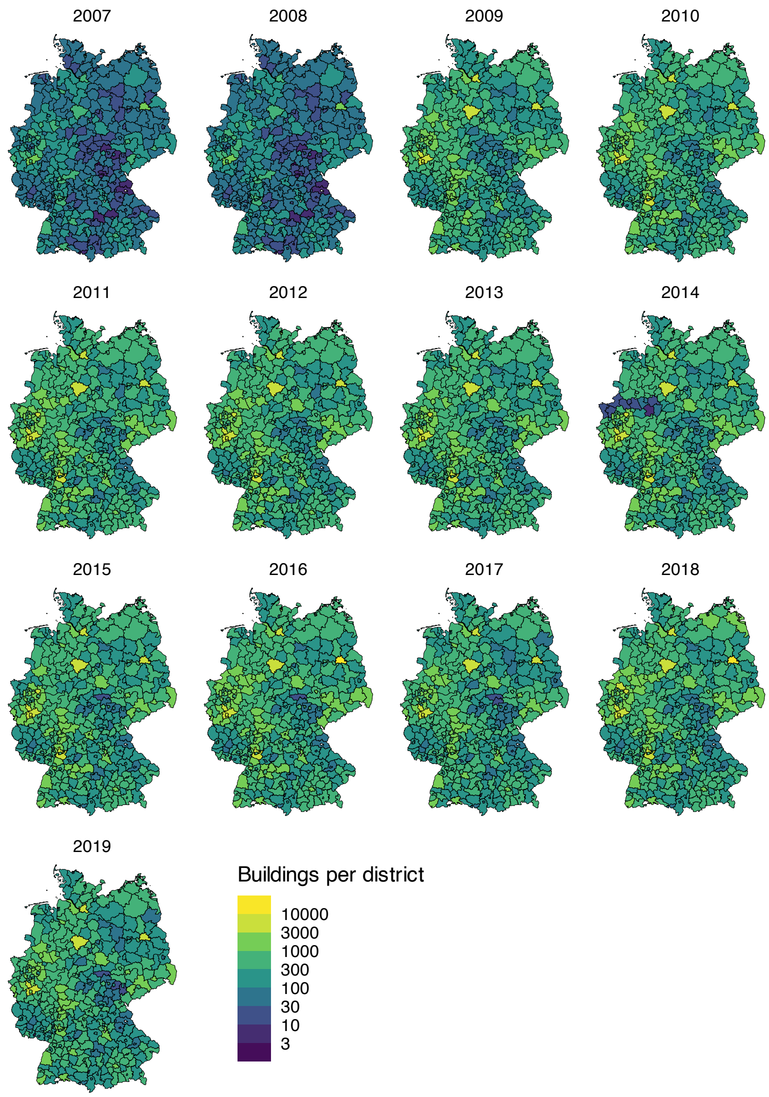

Figure 3.4: Map of the spatial and temporal coverage of the observed buildings

Summary statistics

Table 3.2 provides summary statistics for the full sample reporting the median values and the interquartile range (IQR). In addition to an overall column, separate columns for the three energy carrier groups gas, oil and district heating are also provided. About 60.6% of the observations in the sample are associated with gas as the energy carrier. A further 29.5% with oil and 9.9% with district heating. For space heat demand, as the response variable, the median value (IQR) for the sample overall is 113 (86, 145) kWh/sqm/y and for energy price, the main explanatory variable, the median (IQR) of the sample is 6.49 (5.77, 7.69) Cents/kWh. In terms of energy

Table 3.2: Summary statistics for full sample

Variable [Median (IQR)]	Unit	Overall N = 2,718,246	Gas N = 1,647,563 (60.6%)	Oil N = 802,451 (29.5%)	District heating N = 268,232 (9.9%)
Space heating demand	[kWh/sqm/y]	113 (86, 145)	116 (89, 149)	117 (91, 148)	83 (63, 109)
Energy price, real	[Cents/kWh]	6.49 (5.77, 7.69)	6.18 (5.56, 6.79)	7.06 (6.08, 8.25)	10.12 (8.71, 11.84)
Degree days		3,446 (3,214, 3,733)	3,418 (3,192, 3,697)	3,536 (3,273, 3,826)	3,397 (3,191, 3,642)
Heating surface	[sqm]	404 (260, 707)	424 (271, 701)	305 (226, 466)	1,118 (556, 2,118)
Housing units		6 (3, 10)	6 (3, 10)	4 (3, 6)	16 (8, 32)
District per capita household income	[€/a]	20,695 (18,786, 22,568)	20,658 (18,731, 22,563)	21,098 (19,388, 22,861)	19,217 (17,667, 21,327)
District retirement share	[%]	0.207 (0.193, 0.220)	0.209 (0.194, 0.222)	0.205 (0.193, 0.216)	0.210 (0.192, 0.230)
Postal code population density	[inhabitants/ sq. km]	572 (217, 1,960)	662 (255, 2,121)	320 (149, 839)	2,053 (505, 4,821)

Note: Median (IQR), No missings

demand and price, there are pronounced and relevant differences between the three types of energy carriers. The average demand for gas and oil is higher than for district heating. The prices for gas are the lowest with relatively small fluctuations. Prices for oil are slightly higher, but show greater variation. Prices for district heating are by far the highest and also show the greatest variations. Additionally, buildings with a district heating system installed are three to four times the size of buildings with gas or oil heating installed (cf. heating surface in Table 3.2).

To check for potential multicollinearity issues between the variables Appendix A.2 shows a matrix of Pearson's correlation coefficients. None of the observed correlations between the variables exceed moderate values, so that concerns about multicollinearity can be ruled out. Furthermore, the correlation matrix reveals that energy price is negatively correlated with energy demand, indicating that the expected negative relationship is reflected in the data.

In-depth focus on energy demand and energy prices

Figure 3.5 provides a more detailed visual summary of the demand variable. The aggregated distribution of energy demand (see Panel A) shows that the lowest annual demand in the sample is around 25 kWh/sqm/y. The distribution peaks at just over 100 kWh/sqm/y and then declines more gradually to very high demand values of up to 350 kWh/sqm/y. The differentiated display of the mean demand levels by energy carrier group and over time (see Panel B) reflects the previous findings. While the demands for gas and oil are on an almost similar higher level, demand in buildings with district heating is considerably lower. One significant contributor to this difference is likely that for gas and oil heating systems, conversion losses occur during combustion in the boiler. For district heating those losses already occurred in the heat plant and are therefore already reflected in the price for district heating (Verbraucherzentrale

NRW, 2016). Typical conversion losses of gas and oil boilers are between 10-20% of the primary energy going into the combustion. Furthermore, some of the additional difference observed may be attributable to the strong difference in building size with buildings with a district heating being considerably larger on average.[7]

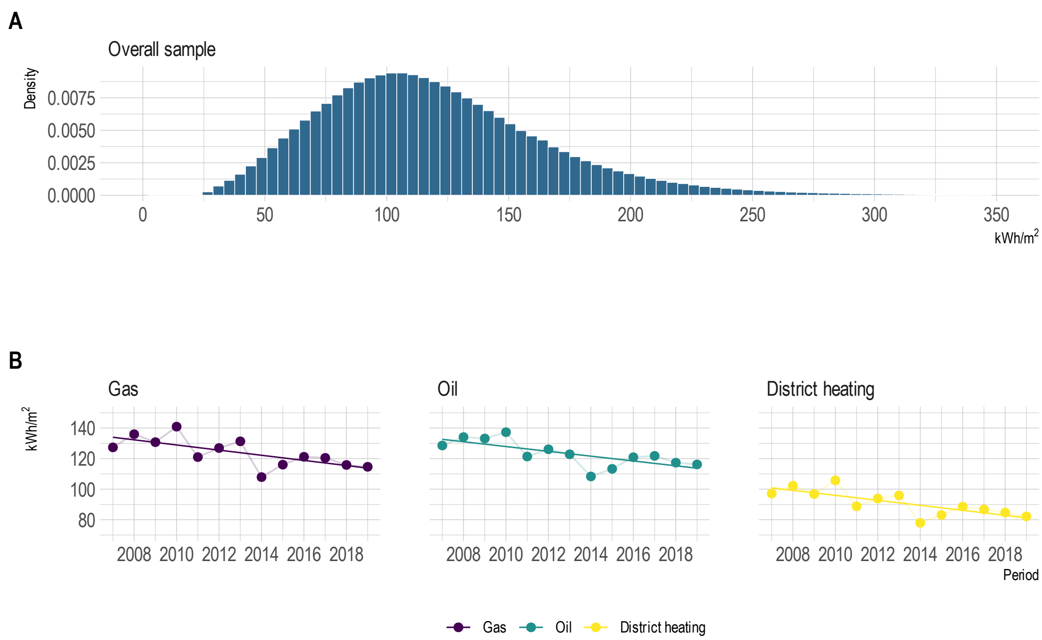

Figure 3.5: Distribution and trend of energy demand

Furthermore, two additional trends can be identified for the demand variable (see Panel B). First, the effective demand decreases over time. While this may be due to changes in the determinants of energy demand, it may also be that newer more energy efficient buildings are added to the sample, while others drop out. Thus, controlling for the building and time dimension appears warranted. Second, while the overall demand level is decreasing, the development for all carrier types follows a similar

[7]Given the significance of the difference in demand, it was further scrutinized if the age of the building and the age of the heating system may deviate between gas and oil buildings on the one and buildings with district heating on the other hand. This information is available for about 15% of the observations from the additional Energy Performance Certificate (EPC) data mentioned earlier, the use of which was otherwise discarded for this thesis. The rationale for looking deeper into these two variables is that newer buildings and more recent heating systems can be expected to go along with lower energy demand. The comparison is provided in Appendix A.1. It does not find differences between the energy carrier groups of a magnitude that would suggest an additional strong argument for varying demand levels.

pattern with year-to-year demand being subject to fluctuations. In Appendix A.1 the distribution of degree days is depicted on a spatial and temporal scale. A comparison of the trends in degree days and energy demand shows a pattern that corroborates the assumption that higher degree days (lower outdoor temperatures) lead to higher energy demand. This illustrates the relevance of considering climatic conditions as a fluctuating determinant of space heating demand.

In Figure 3.6 below, the distribution of nominal (Panel A) and real (Panel B) energy prices as well differentiated by energy carrier group are depicted. The points represent average annual prices. The vertical bars represent one standard deviation (SD). While the energy prices for gas are relatively stable over the period under investigation, prices for oil show stronger volatility. As established previously, prices for district heating are higher than those for gas and oil (cf. Table 3.2). Furthermore, the prices for district heating are not as volatile, but show a much wider range of variation, which is reflected in the width of the vertical bars.

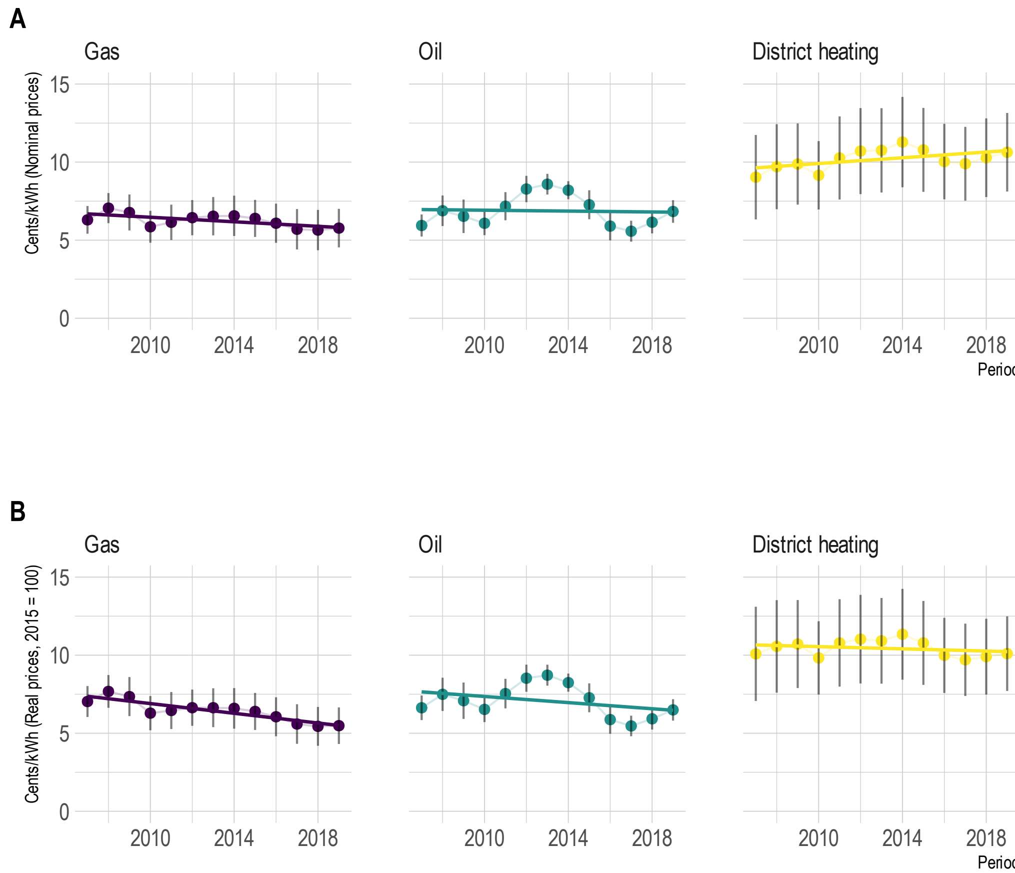

Figure 3.6: Trend of nominal and real energy prices

While nominal prices are relatively stable in the overall trend over time for all three energy carrier groups, the direction of the trend changes after adjusting for inflation. The deflated real prices show declining energy prices over the time period under investigation, with the trend being more pronounced for gas and oil. For the intuition regarding the price elasticity of space heating demand, this means that one would expect an overall increase in demand due to the price effect alone. However, the graph also shows that effects in both directions can be observed if one does not look at the overarching trend, but at the year-to-year movements in energy prices. After having described the underlying data in detail, the following Chapter moves to presenting the results of the regression analysis.

Chapter 4

Results

The results Chapter is structured in two parts. First, the results from the analysis of the full sample are presented (see Figure 3.1). Subsequently the results of the deepened analysis based on the stratified subsample are reported (see Figure 3.2).

4.1 Full Sample Results

Price elasticity estimates

The results of the full sample analysis are shown in Table 4.1. The overarching finding regarding the price elasticity of space heating demand is that the elasticities in all model specifications are *negative* and *relatively inelastic*. On the one hand, the price elasticities being negative confirms that the basic intuition of falling demand with rising prices is reflected in the data. The relative inelastic magnitude of the demand response, on the other hand, validates the assumption derived from the literature that space heating in private households – especially in the short-term – is an inelastic good.

Specifications (m1.ols) and (m2.ols) are the cross-sectional models based on the OLS estimator. Interestingly, the energy price coefficient remains stable when moving from Specification (m1.ols) with price as the only predictor of space heating demand to Specification (m2.ols) where the set of additional control variables is added to the model. The price elasticity coefficient in Specification (m2.ols) is to be interpreted that a 1% increase in energy price leads to a -0.365% reduction in space heating demand. When moving from the OLS estimator to the fixed-effects (FE) estimator

Table 4.1: Regression table for full sample analysis

	Response variable in all model specifications: Ln of space heat demand					
	OLS		Fixed Effects (FE)			
	(m1.ols)	(m2.ols)	(m1.fe)	(m2.fe)	(m3.fe)	(m3.fe, lagged)
(Intercept)	5.399 ***	3.593 ***				
	(0.002)	(0.028)				
Ln of energy price	-0.365 ***	-0.365 ***	-0.251 ***	-0.243 ***	-0.243 ***	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
Ln of lagged energy price (t-1)						-0.071 ***
						(0.001)
Ln of degree days		0.626 ***		0.779 ***	0.779 ***	0.850 ***
		(0.002)		(0.003)	(0.003)	(0.004)
Ln of heating surface		-0.133 ***		-0.405 ***	-0.405 ***	-0.408 ***
		(0.000)		(0.003)	(0.003)	(0.004)
Energy carrier: Oil		0.031 ***		0.097 ***	0.097 ***	0.074 ***
		(0.001)		(0.002)	(0.002)	(0.002)
Energy carrier: District heating		-0.061 ***		-0.020 ***	-0.020 ***	-0.067 ***
		(0.001)		(0.002)	(0.002)	(0.003)
Ln of district per capita household income		-0.277 ***		0.022 **	0.022 **	0,015
		(0.002)		(0.008)	(0.008)	(0.010)
District retirement share		-0.084 ***		0.435 ***	0.435 ***	0.418 ***
		(0.010)		(0.030)	(0.030)	(0.037)
Postal code population density		0.043 ***		0,000		
		(0.000)		(0.005)		
N	2718246	2718246	2718246	2718246	2718246	2020628
R2	0,049	0,156	0,822	0,829	0,829	0,847
logLik	-1322263,852	-1159565,133				
AIC	2644533,703	2319150,267				

*** p <0.001; ** p <0.01; * p <0.05.

in Specifications (m1.fe) - (m3.fe), the estimated price elasticities of demand become more inelastic. This effect is in line with what, for example, also Miller & Alberini (2016) find when moving from a cross-sectional model to the use of unit-level fixed effects. The reason for this is probably unobserved factors at the building level that are not fully reflected in the building characteristics controlled for, as well as temporal effects. The assumption is also supported by the change in r-squared as a metric of the variance in the response parameter explained by the model. While r-squared is at about 15.6% in Specification (m2.ols), it increases strongly to 82,2% when moving to Specification (m1.fe). For the price coefficient, the change from Specification (m1.fe), where price is the only predictor of space heating demand, to Specification (m2.fe), which also includes the other predictor variables, becomes slightly more inelastic, but the effect remains small. Specification (m3.fe) drops the population density from the equation since it is found to have no effect in Specification (m2.fe). Thus, Specification (m3.fe) reflects the final model for the full sample analysis. The price elasticity coefficient in Specification (m3.fe) is to be interpreted that a 1% increase in energy price leads to a -0.243% reduction in space heating demand. The results for the price elasticity of demand are of a comparable magnitude to previous evidence in the literature (see Figure 2.2).

Specification (m3.fe, lagged), also shown in the table, shows the results for the case that not the energy price in the same period, but the price in the previous period (communicated through the energy bill of the previous year) influences the demand decision of households. When the lagged energy price is used as a predictor, the magnitude of the demand response is reduced to one third of the effect size: from -0.243% when using the price in t to -0.071% for the price in $t-1$ (contingent on a 1% increase in energy prices). While the demand response is still negative, this represents a major reduction in magnitude of the effect. In the literature, there are varying rationales on whether the price in the same or in previous periods informs demand decisions for space heat. In this study I follow the approach of Schmitz & Madlener (2020) and assume that prices in the same period drive demand levels. However, the results for when using the price in previous billing period are reported to obtain a more comprehensive picture on the price elasticity estimates. Importantly, it should be noted that part of the difference between the estimates may also be due to the fact that in Specification (m3.fe, lagged) about 0.7 million fewer observations are included in the estimation, since a building must occur in the sample in two consecutive periods (energy price in $t-1$ and energy demand in t) for an observation to be included. This exclusion criterion may lead to some type of selection bias, where buildings with certain characteristics that continuously appear in the sample are more heavily represented.

Other estimates and effects

In general, almost all model predictors are found to be statistically significant with a p-value $< 0.1\%$. However, this is not surprising given the very large number of more than 2.7 million observations (Gelman & Stern, 2006; Johnson, 1999). Besides the interpretation of the price elasticities, the signs of the other relevant model predictors also point in the expected direction (see rationales established in Chapter 2.3). As most of the predictors are also included in the model in the ln-form, they can as well be interpreted in terms of percentage point change.

Referring again to the main model Specification (m3.fe), this means that, for example, a rise in degree days (lower outside temperatures) of 1% would be associated with to a higher energy demand of 0.779% and a larger building heating surface of 1% would be associated with a lower energy demand of -0.405%. Furthermore, the results indicate that an oil heating system is associated with a higher heating energy demand than gas as the reference category and a district heating system with a lower demand. While the results in the cross-sectional OLS model indicate that an increasing district per capita household income would be associated with a decreasing energy demand, this effect disappears in the FE model and is even reversed. At the same time, the magnitude of the district per capita household income effect in Specification (m3.fe) is so small that one must assume that it does not have a strong influence in explaining the variance in energy demand at all. For district retirement share the sign of the coefficient as well switches when moving from the OLS to the FE model. The results in Specification (m3.fe) indicate that a higher district level retirement share is associated with higher energy demand.

4.2 Subsample Results

Price elasticity estimates

Figure 4.1 summarizes the posterior distributions for the price elasticities of space heating demand in the Bayesian models (see Figure 3.2 to compare which model structure and predictor is used in the respective Specification). The complete model summaries – going beyond the price coefficients – are reported in Appendix A.2.

Also for the Bayesian models, the estimates for the price elasticity of space heating demand are in the negative and relatively inelastic range. In contrast to the results of the total sample, however, it is noticeable that the price elasticities found in the subsample are more elastic. This could be partly due to the alternative modelling approach. But probably it is mainly the different composition of the sample that causes the difference in the estimates. That the overall findings of the full and subsample analysis agree well is also evident from the fact that the trend towards lower estimates for the price elasticity is the same when additional predictors are included and when controlling for the group-level effects of buildings and time.

Specifications (b.1) and (b.2) represent the simpler model design without varying intercepts. In Specification (b.1), with energy price as the sole predictor of energy demand, the price elasticity is estimated to be -0.56 [-0.59; -0.53] (Mean estimate [95% CI]). When moving to Specification (b.2) where the additional explanatory variables are added, the estimate moves to -0.41 [-0.45; -0.38] and thus becomes wider and less price responsive (inelastic). Again, using the results from Specification (b.2) as an example, the price elasticity coefficients are o be interpreted that a 1% increase in energy price leads to a -0.41% [-0.45; -0.38] reduction in space heating demand.

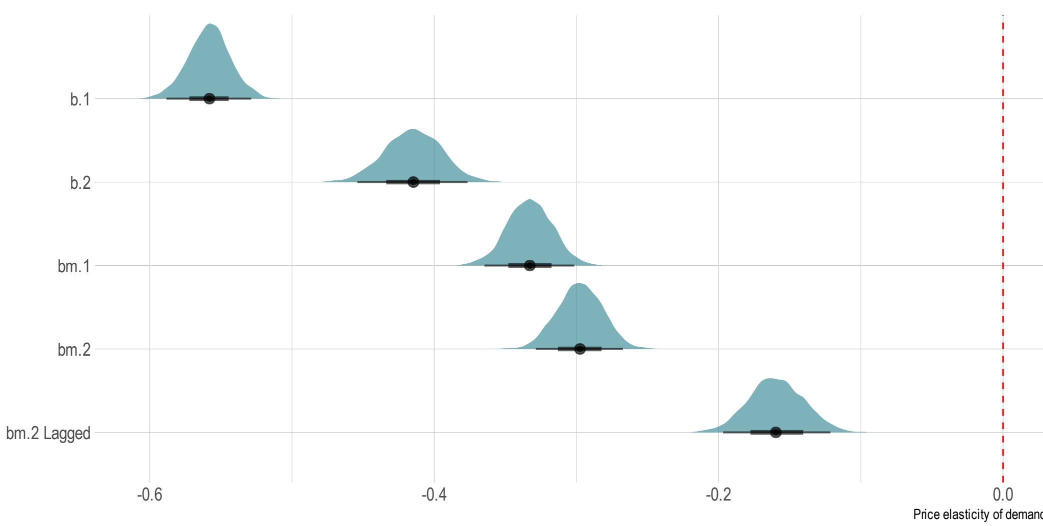

Figure 4.1: Posterior distributions for the price elasticities of space heating demand based on the subsample

In the more sophisticated specifications involving multilevel partial pooling for buildings and years (bm.x models), the price elasticities become lower, similar to the FE models in the full sample analysis. In Specification (bm.1) energy price is used as a sole predictor and group-level terms are included. In Specification (bm.2) the additional explanatory variables are included. The inclusion of the additional variables resulted in divergent transitions caused by the heating area and population density variables. Therefore, the two were excluded from the model to ensure convergence and reliability of results. In Specification (bm.2) the estimate for the price elasticity of space heating demand is -0.30 [-0.33; -0.27], meaning that a 1% increase in energy price would lead to a -0.30% [-0.33; -0.27] reduction in space heating demand. Also similarly to the full sample analysis, the by far lowest estimate is found when using the lagged energy price variable $(t-1)$ instead of the energy price in the same period (t) to explain the variance in space heat demand. For Specification (bm.2 Lagged) the estimate is -0.16 [-0.20; -0.12].

Generally, it should be noted that due to the varying representation of energy carriers between the full- and the subsample, a direct comparison of the resulting estimates is not sensible. In the full sample there is a larger representation of observations relying on a gas heating system and fewer observations with a district heating system. These differences are removed by the stratified sampling approach.

Re-scaling elasticity results to the demand scale and integrating prediction intervals to convey uncertainty

Figures 4.2 and 4.3 illustrate what the differences found in the price elasticity estimates imply when being converted back to the actual scale of demand. Figure 4.2 shows predictions for the basic Specification (b.1) where energy price acts as the sole predictor of space heating demand and group-level differences between buildings and years are not yet considered. Panel A on the left shows the observations in the subsample with model predictions on the ln-scale. Panel B on the right shows the re-scaled values on the scale of actual demand.[1]

[1]Values are re-scaled using $exp()$.

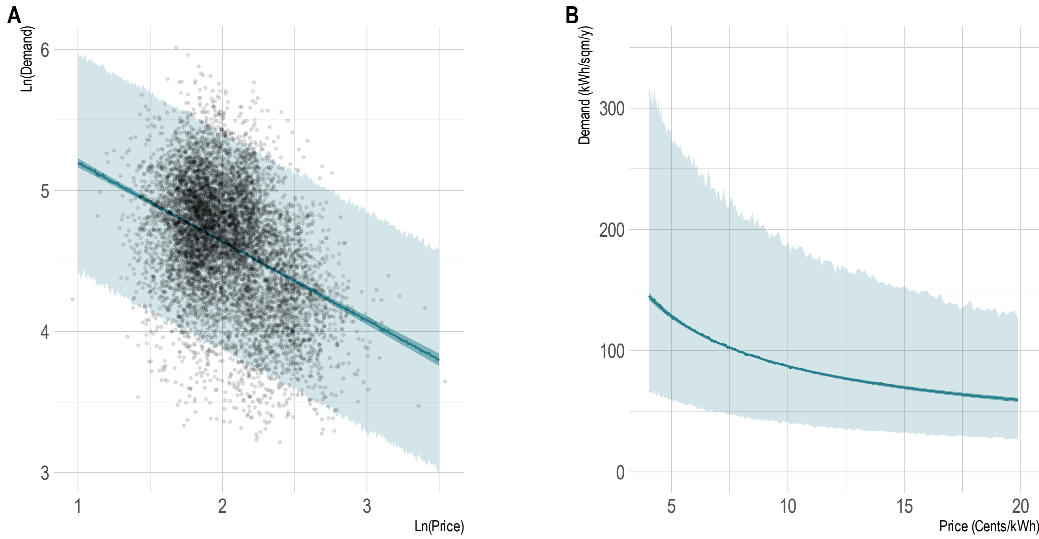

Figure 4.2: Predicted price elasticities of space heating demand based on simple model (Specification b.1)

Methodologically, the line in the center represents the average energy demand at a respective price. The narrower, darker shaded interval around the mean represents the 95% intervals for the posterior distribution of the expected value of the demand parameter μ (see Chapter 3.3). These posterior predictive distributions have a lower variance because only the uncertainty of the mean is included in the draws, but the residual error is ignored. The wider, lighter shaded interval, on the other hand, represents the posterior predictive distributions including the residual error. It is thus to be interpreted as the interval in which the model expects to find 95% of the actual energy demands in the population at each price.[2]

The two types of prediction intervals convey two types of uncertainty: uncertainty in the parameter value (darker shaded interval) and uncertainty in the sampling process (lighter shaded interval). The distribution of simulated outcomes (lighter shaded interval) includes sampling variation and therefore indicates what the future data would look like (McElreath, 2020). In other words, it tells us what demand levels can be expected at different prices in a population of German multi-apartment buildings with equal shares of the three energy carriers gas, oil and district heating. The

[2]Cf. McElreath (2020), Chapter 4 and the brms documentation on the *posterior.epred* and $posterior_predict$ functions in R for the interpretation of the posterior predictive intervals shown in the graph.

lighter shaded intervals are therefore the most important intervals for interpretation. It should be noted, however, that due to the stratified sampling procedure used to create the subsample, the interpretations are only generalisable to a limited extent, as the relative shares of the energy carriers do not reflect the shares in the actual building stock.

For the simple Specification (b.1) in Figure 4.2, the mean space heating demand expected by the model drops from about 150 kWh/sqm/year at a price of 4 cents/kWh to about 60 kWh/sqm/year at a price of 20 cents/kWh.[3] In addition to the mean, the intervals show well the uncertainty associated with the model predictions. For the low end of the chosen price range at 4 cents/kWh, the model expects 95% of the actual energy demand of the population to be between 70 and 315 kWh/sqm/year, which is obviously a very wide range. At the upper end of the price range, the predicted interval narrows. At 20 cents/kWh, the model expects 95% of the actual energy demand to be in the range between 35 and 130 kWh/sqm/year. Overall, the predictions of the simple model thus show the negative relationship between price and demand, but are still subject to large uncertainty due to the residual error of the model.

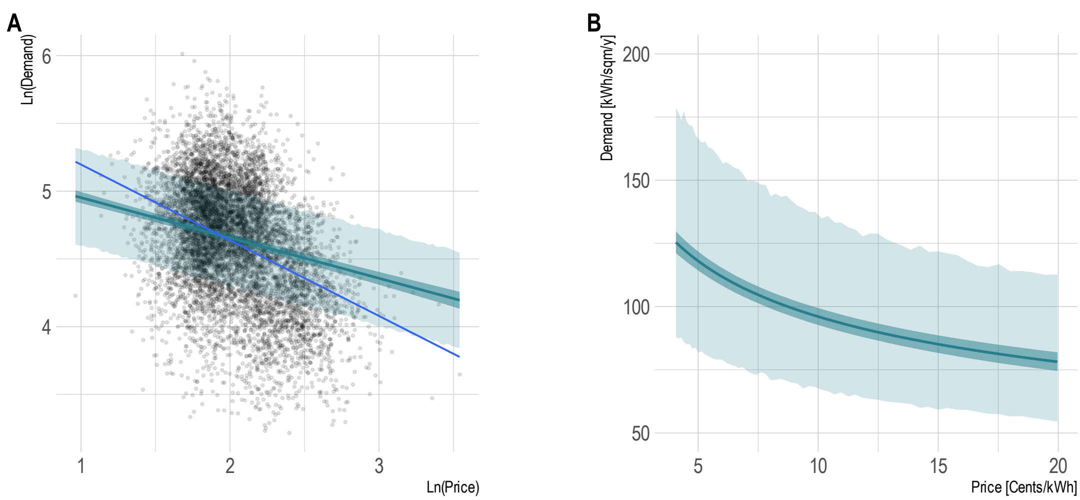

Figure 4.3: Predicted price elasticities of space heating demand based on multilevel partial pooling model (Specification bm.2)

[3]Note that the price range shown in the graph were chosen arbitrarily to some extent. They were chosen to represent the lowest reasonable price at the low end and to reflect about 2-3 times the historical prices observed in the sample at the high end.

Figure 4.3 presents the same type of figure, but relies on the multilevel partial pooling model from Specification (bm.2). The additional blue line in Panel A represents the simple linear model between price and demand. The different trajectory between the simple linear model (blue line) and the prediction for the model Specification (bm.2) (green line) shows that the inclusion of the additional explanatory variables and the group-level pooling for buildings and years has a strong impact on the prediction. The predicted elasticity is less elastic and the intervals are much narrower. On the demand scale, again shown in Panel B, the average energy demand predicted by the model drops from about 125 kWh/sqm/year at a price of 4 cents/kWh to about 80 kWh/sqm/year at a price of 20 cents/kWh. Moving to the simulated results (lighter shaded intervals), the model expects 95% of the actual energy demand in the population to be between 85 and 175 kWh/sqm/year for the lower end of the price range shown at 4 cents/kWh. At the higher end, at a price of 20 cents/kWh, the interval narrows again to 55 to 110 kWh/sqm/year. Thus, the intervals have become much narrower, indicating that the prediction of Specification (bm.2) provides a more realistic picture on the interplay between energy price and demand as compared to the prediction from the simple model in Specification (b.1).

Overall, the figures provide a good indication of what the reported price elasticities of space heating demand would mean in practice, given a range of realistic price levels.

Investigation of potential heterogeneity in price responsiveness

While the results presented thus far for the subsample largely mirror the procedure used for the full sample analysis, but based on Bayesian inference and using a different sample composition, the final part of the subsample analysis aims to go further by examining potential sources of heterogeneity in the price responsiveness. This is done in two steps. First, the variables used in the analysis are examined for potential heterogeneity relying on graphical representations. Second, the heterogeneity found visually for the different energy carrier groups is then being formalized in a regression analysis by means of extending Specification (bm.2) by a interaction term between energy price and energy carrier groups.

For the first part, the visual examination of potential heterogeneity in price responsiveness, the following variables are considered: year, federal state, energy carrier group, degree days, heating surface, district per capita household income, and district retirement share.[4] For the categorical variables years and federal states the graphs

[4]Note that postal code population density was not considered because it was neither included in the final model for the analysis of the full sample (Specifications (m2.fe)) nor in the final model for

are shown in Appendix A.4. The observations in the subsample are grouped by year or federal state and presented in a scatter plot between energy price and demand. The lines represent simple linear models for each grouped category of year or federal state. For the years variable, the graph well indicates that the energy demand in the sample is lower overall in the later periods. With regard to potential heterogeneity in price elasticity between the years, however, no significant differences can be identified. The lines for all years follow more or less the same slope. The same is true for the federal state variable. While there are some differences between the individual states, there is no pattern of strongly diverging price elasticities of space heating demand.

Similarly, Appendix A.5 presents the visual examinations for the variables degree days, heating surface, district per capita household income, and district retirement share. Since all four variables are continuous, they are each divided into three equally sized groups to be able to investigate possible heterogeneity in the price response visually (e.g., one third of observations with highest number of degree days, one third with medium number of degree days, one third of observations with lowest number of degree days). For the two variables degree days and district retirement share, the three grouped lines are almost completely parallel, suggesting that there are no differences in the price response between the groups. For district per capita household income minor differences can be observed. Districts with a lower per capita household income exhibit a more elastic demand response (steeper slope). For the building heating surface more pronounced differences are found with larger buildings being associated with a more elastic demand response than smaller buildings.

the analysis of the subsample (Specifications (bm.2)).

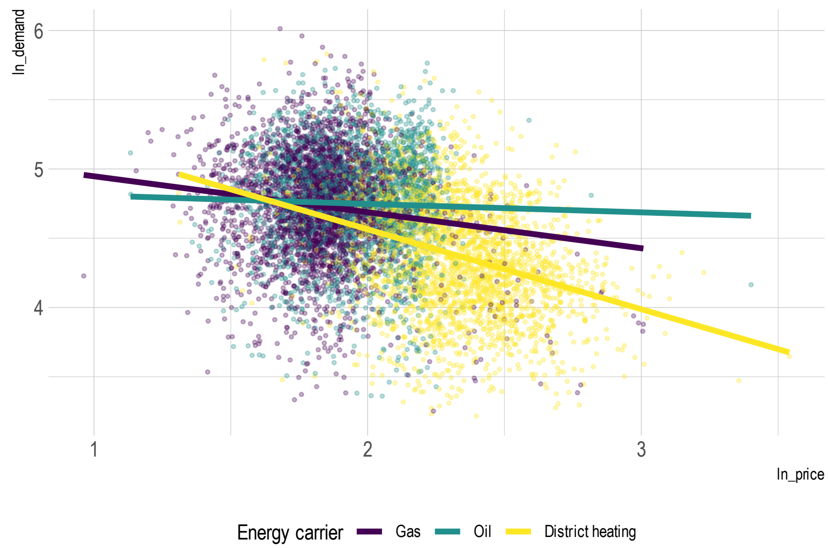

Figure 4.4: Scatter plot for the investigation of heterogeneity between energy source groups

In addition, the energy carrier group was investigated for potential differences between the three carrier types (oil, gas, district heating). The results are shown in Figure 4.4. The differences in the price responsiveness between the energy carrier groups are most pronounced. For buildings with oil heating, demand is almost constant, which indicates that they do not adjust their energy demand to a large extent to a change in the oil price. Buildings with gas heating show a moderate demand response to price changes. And buildings with district heating have by far the most elastic demand response when relying on the simple linear relationship shown in the graph.

Based on the visual examination of the variables, the findings were used to construct a better fitting model which integrates heterogeneity in price responsiveness by adding an interaction term. Since the visual analysis revealed that there were no heterogeneity detected between varying years, federal states, degree days and district retirement shares, these variables were not probed further. The heterogeneity found between the energy carrier groups but also among the district per capita household income and building heating surface variables were assessed further. Exploratory testing of models showed that interacting energy price with district per capita household income showed no relevant difference. Furthermore, interacting energy prices with building heating surface lead to non-convergence of model results and was therefore not further pursued. Interacting energy price with the energy carrier group, however, lead to relevant differences for the price elasticity pared with a converging model.

The posterior distributions for the price elasticities of space heating demand for the additional estimated interaction model Specification (bm.3) are shown in Figure 4.5. The different posterior distributions reflect the results of the visual examination of the simple linear relationship. The most inelastic demand response is found for buildings with oil heating -0.16 [-0.23; -0.10] (Mean estimate [95%CI]). For buildings with gas as the energy carrier type, the estimates are in a medium range of -0.35 [-0.40; -0.29]. For buildings with district heating, demand is most elastic with estimates of -0.53 [-0.60; -0.46]. Also the width of the distributions provides some insights. For gas and district heating as energy carriers the posterior distribution is wider than for oil. This is likely due to the observations for gas and district heating to be more widely spread (see also the previous Figure 4.4). In line with the previous estimates of price elasticity, the coefficients are to be interpreted as a corresponding decrease in energy demand given a 1% increase in energy prices.

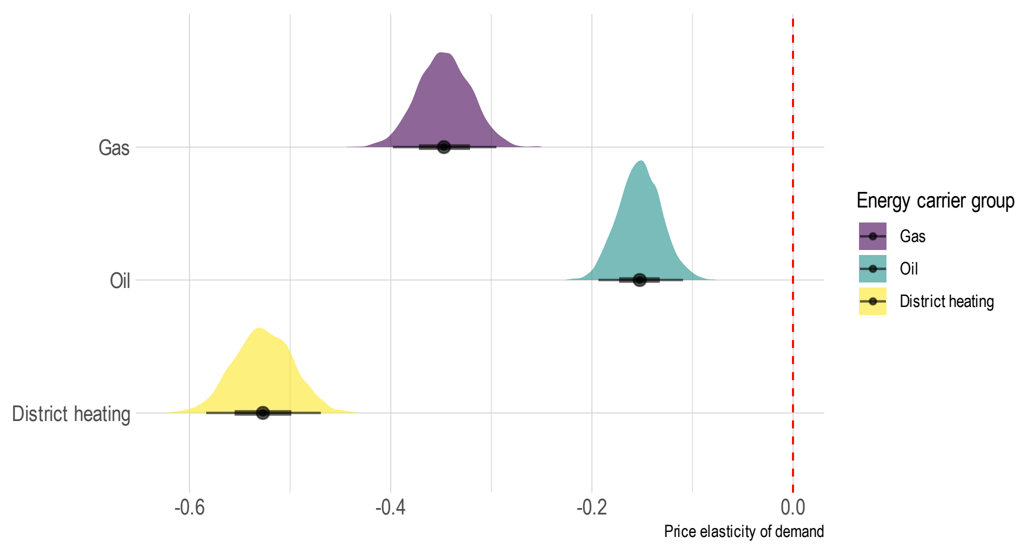

Figure 4.5: Posterior distributions for the price elasticities of space heating demand by energy carrier group based on the subsample

As with the previous models, the results of the model Specification (bm.3), where energy price and energy carrier group are interacted, are shown with prediction intervals in Figure 4.6. Again, the inner darker shaded intervals represent the uncertainty in the parameter value and the lighter shaded intervals represent the uncertainty in the sampling procedure and thus include the residual error of the model (McElreath, 2020).

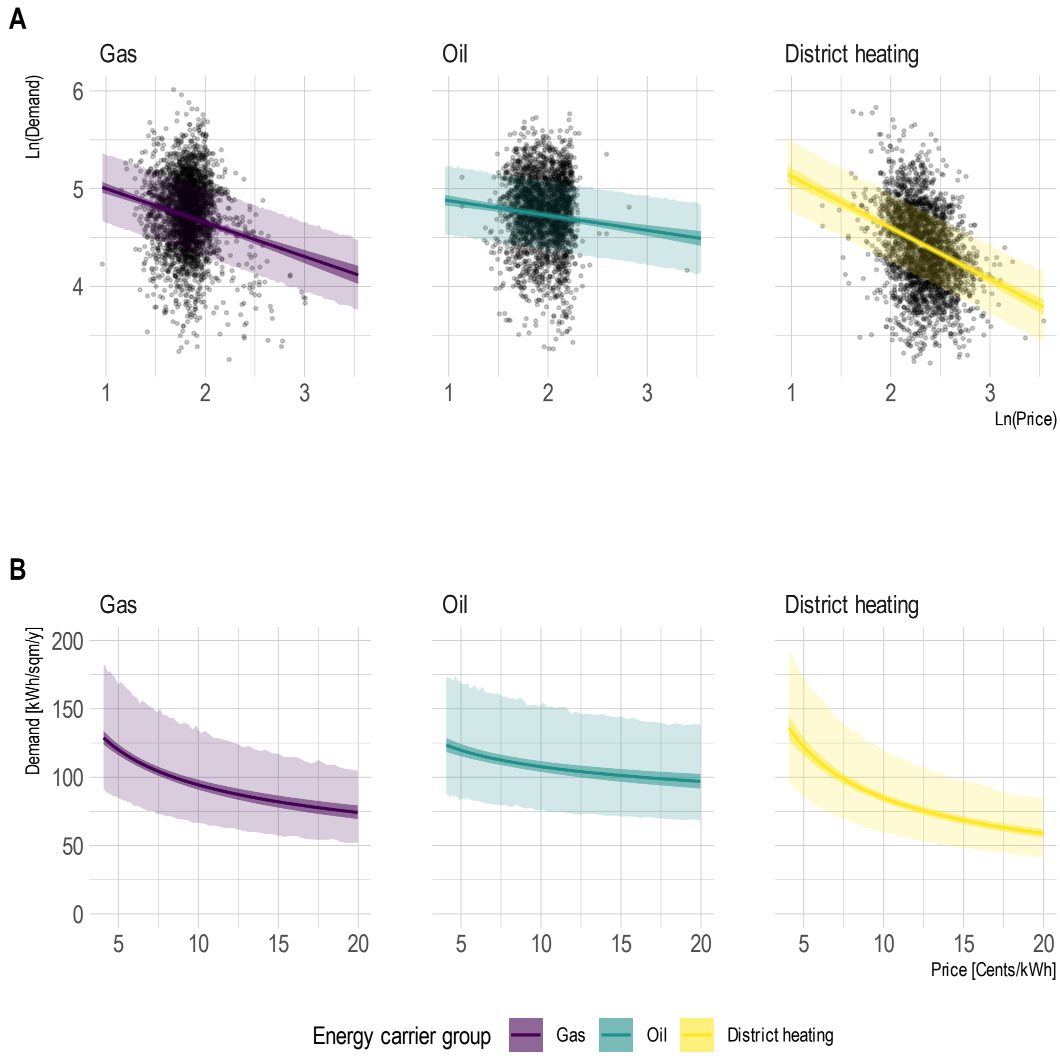

Figure 4.6: Predicted price elasticities of space heating demand by energy carrier group (Specification bm.3)

In Panel A the models are shown on the ln-scale together with the underlying observations for energy price and demand. In Panel B the values from Panel A are rescaled to the demand scale. Mirroring the varying price elasticities of demand for the three energy carriers, the demand curve for district heating is most responsive to rising energy prices and the demand curve for oil the least responsive. Interestingly, due to the exponential relationship, the prediction intervals for gas and district heating become more narrow with rising energy prices, while the prediction interval for oil remains relatively wide, also for higher energy prices. For gas, the results can be interpreted in the way that for the lower end of the shown price range at 4 cents/kWh, the model expects to find 95% of the actual energy demand in the building stock using gas heating in the demand range between about 90 and 180 kWh/sqm/year. For the upper end of the price range shown at 20 cents/kWh, the expected demand values are between about 50 and 105 kWh/sqm/year. For oil, the model expects 95% of the actual energy demand to be between about 85 and 175 kWh/sqm/year for the lower end and between about 70 and 140 kWh/sqm/year for the upper end of the indicated price range of 4 to 20 cents/kWh. Ans lastly, for district heating, the model expects 95% of the actual energy demand to be between about 95 and 190 kWh/sqm/year for the lower end and between about 40 and 85 kWh/sqm/year for the upper end of the indicated price range of 4 to 20 cents/kWh.

Model comparison

In the previous parts of the chapter various model Specifications for the Bayesian subsample analysis are presented, all producing varying results for the price elasticity of space heating demand. Thus, it is important to investigate which model has the more likely fit. For this purpose, Bayesian regression analysis resorts to the principle of of model selection which assumes that many potential models could be generating our data and that based on tools one should investigate which model is doing this more likely than others (Vehtari, Gelman, & Gabry, 2017). To practically implement a model comparison, I rely on the loo-package which uses efficient approximate leave-one-out cross-validation (LOO) for Bayesian models where the approximation uses Pareto smoothed importance sampling (PSIS) (Vehtari, Gabry, Yao, & Gelman, 2018).

The results of the model comparison are shown in Table 4.2. In addition, a visual comparison of model fit is provided in Appendix A.6.[5] In Table 4.2, the models are

[5]Note that the model Specification (bm.2, lagged), which uses the price from the previous period and thus does not have the same number of observations, cannot be compared directly with the others. Therefore, it is not included in Table 4.2. However, because the visual representation of model fit in Appendix A.6 relies on simulated density distributions that can be produced for each model individually, Specification (bm.2, lagged) is included there. The pattern of simulated density distributions indicates a similar model fit to the other Specifications with a multilevel partial pooling approach (bm.x).

Table 4.2: Model comparison based on PSIS-LOO

Model Specification	Difference in ELPD	SE of the Difference
bm.3	0,00	0,00
bm.2	-59.1	14.2
bm.1	-155.6	21.5
b.2	-5102.6	110.7
b.1	-5763.1	116.1

compared in terms of their difference in expected log pointwise predictive density (ELDP) using LOO as a method. The results reveal that the Specification (bm.3) has the highest ELDP and thus the best model fit. The fit of the other models is represented relative to the best fitting model. Since the standard error associated with the ELDP estimate for Specification (bm.3) (14.2) is much smaller than the ELDP difference compared to Specification (bm.2) (59.1) as the model with the second best fit, there is no ambiguity regarding Specification (bm.3) being the model with the best fit. Furthermore, the results in Table 4.2 show that the model fit increases significantly when moving from the simple model Specifications (b.1 & b.2) to the multilevel partial pooling models (bm.x). Hence, the model results using the multilevel partial pooling approach must be considered much more reliable than the results from the simple Specifications. Furthermore, Specification (bm.3) having the best fit indicates that including the interaction term between energy price and energy carrier group increases the reliability of the model and that price elasticities of space heating demand are more precise when being differentiated by energy carrier group.

Chapter 5

Discussion

5.1 Discussion of Empirical Results

Findings on the price elasticity of space heating demand

The first research question which this thesis aimed to answer is: How does a change in the price of energy affect the level of space heating demand? Put differently, the fist research question aims at determining an estimate of the price elasticity of space heating demand.

The results presented previously confirm the assumption that the demand for space heating is a highly inelastic. In the final model Specification in the analysis based on the full sample, the price elasticity was estimated such that the demand for space heating would decrease by -0.243% if the energy price increased by 1% (see Chapter 4.1). This means that the overall price elasticity estimated from the full set of data is lower than estimates from previous studies focusing on Germany, most of which estimate an elasticity between -0.3 and -0.6 (see Chapter 2.2). Interestingly, however, the estimated overall price elasticity mirrors the results by Auffhammer & Rubin (2018), the only other previous study that also relies on energy billing data. This might suggest that the previous studies on Germany which are based on social survey data, and therefore involve no observation-specific information on energy prices, overstate actual demand responses. Furthermore, the fact that the price elasticity estimates observed in the different model Specifications moved downwards when the grouping variables for buildings and years were included either as fixed-effects in the full sample or with a partial pooling approach for the Bayesian regression analysis of the subsample is consistent with findings also made by others in the literature (e.g., Miller & Alberini, 2016). This refers in particular to the results of the study by Rehdanz (2007), which finds very high elasticities compared to this thesis and also compared to other studies, but does not control for unobserved inter-individual differences between buildings and temporal dynamics over time.

Due to the very large sample size of more than 2.7 million observations in the full sample analysis and its inversely proportional relationship to the standard error of a model, the standard error of the price elasticity estimates is very small and does not necessarily reflect the true uncertainty and dispersion about the different combinations of price and demand observed in the data. The uncertainty and dispersion becomes more apparent when looking at the results for the stratified random subsample (see Chapter 4.2). Even though the estimates for the price elasticity of demand are not directly comparable due to the different sample composition, the final model Specification (bm.2), which does not yet include the interaction term, clearly shows that the prediction for price elasticity as a single parameter is rather narrow, but that the prediction interval becomes significantly wider when the residual error term of a model is included (see Figure 4.3).

In addition, it is also necessary to discuss the estimated elasticity that results when the price of the previous period, rather than the price of the same period, is used to determine the price elasticity of space heating demand. The results of the analysis of the full sample show that the estimated elasticity becomes even far more inelastic, dropping from -0.243 when using the same period price to -0.071 when using the previous period price. This means that if one assumes that the price from the previous year informs the demand decision of inhabitants, because they were made aware of the prices through the last energy bill, the demand response is considerably more limited. However, it remains questionable whether a direct comparison of the two elasticity estimates is appropriate. Due to the need to have two consecutive billing observations (energy price in $t-1$ and other variables in t), about 0.7 million observations are removed from the sample in the model with the price lag. In contrast to the dropped observations, buildings that appear consecutively in the sample may have special characteristics (e.g., only larger buildings). This may introduce a bias compared to the full sample used in the other model Specifications. Moreover, it remains questionable whether the information obtained from a previous year's bill is actually more informative for residents about energy prices than their monthly prepayments in the current billing period together with the reporting of energy prices in the media or via other information channels. Nevertheless, to incorporate the estimate, one could use it, for example, as a lower bound of the price responsiveness when doing sensitivity testing with more than one estimate.

In summary, the inelastic results found for the price elasticity of space heating demand indicate that households do adjust their demand to the incremental price changes that occurred during the period under investigation, but only to a rather limited extent. Transferring this finding to the effects of price-related policy instruments, such as the BEHG recently introduced in Germany for the building sector to price carbon externalities, means that the direct demand effect to be expected is relatively limited in the short-term. Importantly, for the case of renter households that are observed here, the short-term elasticity may also reflect the long-term elasticity under the current design of the BEHG. As the carbon price is currently only borne by renters, who generally have limited options to make structural changes to the thermal conditions of their housing and change the heating technology, no additional long-term effect can be expected. However, this aspect is discussed further later in this section (see Chapter 5.2).

In addition, it should be pointed out once again that the historical period examined in this thesis is characterized by a gradual change in energy prices. Accordingly, the estimates for the price elasticity of space heating demand can only be transferred to the current situation of energy and especially natural gas scarcity in Europe to a limited extent. At the same time, the effects of the current strong price increases for residential heating energy on the price elasticity are anything but clear. On the one hand, it can be assumed that the strong public attention to energy prices probably leads to a promotion of energy-efficient behavior, which in turn would result in a more elastic demand. At the same time, the strong price increases have a particularly strong impact on the budget of low- and middle-income households, which is an additional factor likely to contribute to a more elastic demand. On the other hand, the price increase is also so strong that households may not be able to reduce their demand further in order to maintain a minimum level of comfort, which in turn would have the opposite effect on the elasticity. Overall, the price elasticities of space heating demand presented in this thesis can provide an empirical starting point also for an analysis of the current crisis situation. However, the underlying structural difference in price development and the limited transferability that this may entail should be made clear.

Heterogeneity between the energy carrier groups

Furthermore, this thesis also intended to provide evidence on the additional research question whether there any potential factors for heterogeneity in the sample that may not be reflected in the estimation of an overall mean price elasticity of demand?

Through the analysis of the subsample multiple variables were investigated for potential heterogeneity in price responsiveness. It was found that the type of energy carrier used within a building represented a relevant dimension for varying price elasticities of demand. The interaction term model used to quantify this difference showed that buildings with an oil heating system are associated with a price elasticity of -0.16 [-0.10; -0.23] (mean estimate [95%CI]) and thus have the lowest price responsiveness. Buildings with a gas heating system exhibit a price elasticity of -0.35 [-0.29; -0.40], implying a medium price responsiveness, but higher than the elasticity found for the full sample. And buildings with a district heating system show a price elasticity of -0.53 [-0.46; -0.60] and thus represent the group with the highest price responsiveness. Since the model comparison showed that the model interacting energy price and energy carrier group had the best model fit, the elasticities should be considered as more precise than the overall elasticity, when information on the type of energy carrier is available.

The reasons for the different price elasticities of demand between the three energy source groups are not completely evident. However, there are some lines of argumentation that may serve as partial explanations for the observed differences. One possible reason for the very inelastic demand response for oil could be that market prices for oil were much more volatile than for the other two energy carrier groups during the study period. This volatility may have led people to view price volatility as more of a normal and familiar state of events and no longer adjusted their demand based on price movements. A second explanation could be that the oil prices observed in the sample are probably less correlated with the current market prices for oil, since oil, unlike gas and district heating, which are on-demand fuels, is usually stored in the building and the price for new supplies is usually mixed with the price for the remaining heating oil stock. These stock effects might obscure the demand reactions within buildings with an oil heating.

Another reason for the difference in price elasticities could be that buildings with a gas or district heating system usually have a fixed cost component (flat-rate basic charge) as part of their energy contract. Due to the aggregated form of the data on energy costs, which are not broken down into fixed cost component (flat-rate basic charge) and variable cost component (demand-based price per kWh), it could happen that buildings with a low energy demand also have a high energy price per kWh because a larger share of the fixed cost component is attributed to a unit of demand. As a result of this trend, the low demand levels could be incorrectly attributed to the role of energy price in a model, potentially distorting the picture. Although this effect may have some impact, it should be noted that the inclusion of the building ID grouping variable helps to mitigate the effect by determining whether low energy demands are common for a particular building.

If one relates the heterogeneity in the price responsiveness found for the energy carriers groups back to the literature, the results can at first glance be seen as contradictory to the findings of the meta-study by Labandeira et al. (2017), which was presented in detail in chapter 2.2. The meta-results indicate that the average short-term elasticities for gas and oil (district heating was not observed) are similar to each other and in the order of magnitude of about -0.2. Thus, the results of this thesis indicate for gas that the price response for the specific purpose of residential space heating demand is higher than for the total pool of gas energy uses observed by Labandeira et al. (2017), and vice versa for oil. However, since the scope of the underlying individual studies in the meta-study is naturally much larger, it can be assumed that the results of this thesis are more accurate for the specific purpose of space heating demand in the residential buildings sector, although less generalisable beyond this scope.

Besides the heterogeneity found for the energy carrier groups, it was established trough the analysis that there was little to no heterogeneity in the sample for the other variables considered (e.g., households in colder regions in Germany do not systematically deviate in their energy demand response to changing energy prices from households in warmer regions).

Effects of other determinants on space heating demand

The third research question this thesis aimed was answer was: What other determinants do affect the level space heating demand and need to be considered so that their effects are not falsely attributed to energy prices?

The empirical analysis considered a set of additional variables, most of which showed a relevant association with the demand for space heating. The most relevant effect on the model results was clearly related to the inclusion of unobserved inter-individual differences between the units (buildings) and also an intra-individual dynamic over time (years) either trough fixed effects in the full sample analysis or trough the partial pooling approach in the subsample analysis. The model comparison conducted for the subsample analysis showed how significantly the model fit increased when considering the two grouping variables.

Even though the additional consideration of the other variables in most specifications has only a minor impact on the coefficient of the energy price when already controlling for the building and the year (see e.g., Figure 4.1, the results for some of the other relevant variables point in the expected directions. The results show that at lower outside temperatures (higher degree days) space heating demand increases to a relevant extent. It can also be seen that the space heating demand in larger buildings is systematically lower than in smaller buildings. The effects of the additionally considered socio-economic variables at district or postal code level, on the other hand, are not clear and lead to different results with different model specifications.

5.2 Integration into the Policy Context

In addition to the discussion of the empirical results, it is also of relevance to embed the price elasticity of space heating demand estimate into the current policy context. The focus is on the medium- to long-term challenges of decarbonising the residential buildings sector in order to achieve climate targets, and less on the current situation of energy and gas scarcity, which has already been briefly discussed above. The highly inelastic demand response to energy price changes observed in this work illustrates that carbon pricing in the building sector is a relevant tool, but one that needs to be embedded in a broader policy mix to make the transition to a low-carbon buildings sector socially desirable and politically viable (Braungardt et al., 2021). The estimated low elasticity means that households are price takers when it comes to price-related policy instruments for space heating. Since the demand response is much more inelastic than a unit-elastic response, the budget that households have to allocate to meet their desired demand for space heating becomes larger when energy prices rise.

This is particularly problematic for rental households, as they are the sole bearers of the carbon price signal under the current design of the Fuel Emissions Trading Act (BEHG) (see Chapter 1.2 for more details on implementation). Owners of rental buildings, on the contrary, are not burdened by the carbon price in the current design of the instrument and therefore have no financial incentive to improve a building's energy quality (Braungardt et al., 2021). Since the opportunities for renters to improve the energetic quality of their apartments is rather limited, it is to be expected that the long-term price elasticity does not deviate strongly from the short-term elasticity. The elasticity estimate presented here therefore implies that the budgets of rental households are not only more burdened by additional heating expenditures in the short-term, but also in the long-term, and that this effect is amplified by the rising carbon price trajectory. This particularly affects the household budgets of low- and middle-income households, as Kröger, Longmuir, Neuhoff, & Schütze (2022) show in their analysis of the impact of the current sharp increase in household heating fuel prices.

In order to make the impact of increasing household expenditure from space heating more understandable, it is useful to return to the representation of average demand curves shown earlier in this thesis (see Chapter 2.1). The following Figure 5.1 shows the impact of the BEHG-induced carbon price of 55 euros per tonne of CO_2e in 2025 on gas. The baseline chosen is the average gas demand in the full sample (116 kWh/sqm/y) and the price is the average gas price in the five-year period between 2016 and 2020 (6.73 Cents/kWh). Although current market prices for gas are evidently much higher, the graph is intended to serve as an exemplary illustration of the structural effects of the carbon price on household expenditures. For the price elasticity represented by the demand curve, the average estimate of -0.35 for gas from the interaction model, Specification (bm.3), is used.

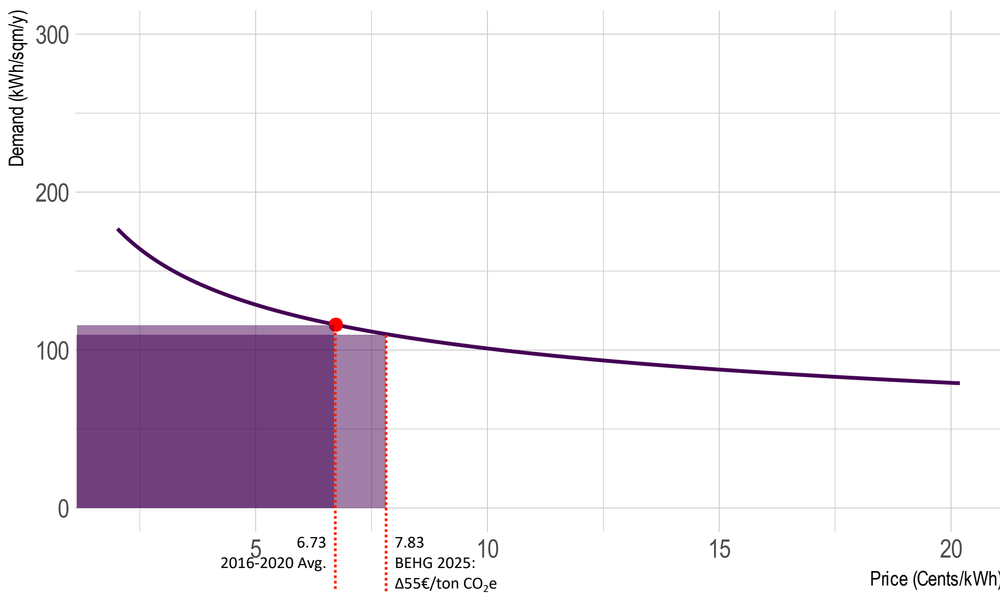

Figure 5.1: Exemplary expenditure effect of inelastic gas demand through the BEHG carbon price in 2025

Under the baseline scenario households would incur expenditures of 7.81 Euros/sqm/y (represented by the rectangle whose corner point lies in the red point in the demand curve) and under the scenario of the BEHG induced price effect in addition to the baseline price households would pay 8.61 Euros/sqm/y (represented by the second rectangle going further to the right).

It is evident that this exemplary effect on household expenditures cannot be borne by households alone in the long-term and needs to be moderated and mitigated by the wider policy mix for the decabonisation of the buildings sector in order to make the transition to a low-carbon buildings sector socially desirable and politically viable. Firstly, it is of overarching importance that the carbon price signal is passed on, at least in part, to the owners of rental buildings in the near future. Doing so would ensure that the long-term price elasticities in the rental segment of the buildings sector would increase. A gradual sharing of the total carbon price between renters and owners depending on the energetic condition of a building, as proposed by the Federal Ministry for Housing, Urban Development and Building earlier this year, is an option where the carbon price signal is partially passed on to building owners (cf. BMWSB, 2022). An alternative approach might be a warm-rent model, as it is implemented in Sweden, for example. In a warm rent model, a surcharge for heat supply is added to the cold rent based on the energy condition of a rental building.

Under this system, the price signal of a carbon price lie entirely with rental property owners.

In addition to passing on the carbon price signal, other policy measures are also important. A recent study for UBA (2022) also concludes that an ambitious mix of instruments is needed to achieve the 2030 emission reduction targets based on a range of possible scenarios. From a regulatory perspective, additional instruments include, above all, the introduction of minimum energy performance standards (MEPS), as proposed by the EU Commission in November 2021 as part of the revision of the Energy Performance of Buildings Directive (cf. EP, 2022). Furthermore, clear phase-out rules for fossil heating systems, especially gas, are an important lever for emissions in the building sector. Both bans on the installation of fossil heating systems and the MEPS could be implemented into national law in Germany by amending the Building Energy Act (GEG) (cf. BMWK/BMWSB, 2022). In addition to regulatory measures, according to the scenarios, fiscal instruments to support the energy-efficient renovation of the building stock, as well as programs to promote renewable heating technologies, are also regarded as highly relevant (UBA, 2022). As well as the expansion of central heat generation via district heating including network planning and the supply-side decabonisation of electricity and district heat generation (UBA, 2022). The carbon price, in turn, strengthens the effects of some instruments in the policy mix such as the use of funding programs. In addition, it can be assumed that the carbon price signal will increase the level of ambition, for example in the case of energy-efficient renovations (UBA, 2022).

Chapter 6

Conclusion

This thesis investigated short-term price elasticities of space heating demand based on a large-scale sample of energy bills in Germany in the period between 2007 and 2019. The estimation of price elasticities of space heating demand is of relevance as it is a key parameter for assessing the expected impact of price-related energy policy instruments, such as the newly introduced Fuel Emissions Trading Act (BEHG) 2021 in Germany. The novelty of this work is that it relies on energy billing data instead of social survey data, which provides a more accurate picture of energy demand and prices than previous studies, which usually use information on household expenditures and average price data from national statistics to derive demand (e.g., Rehdanz, 2007; Schmitz & Madlener, 2020). Furthermore, this thesis relies on a broad range of statistical methods and applies Bayesian regression analysis with a multilevel partial pooling approach on a subsample of the data. Bayesian methods are well suited to propagate and communicate the uncertainty in the model estimates.

According to the theory of price and demand, the demand for a good is mainly determined by its own price and the price of possible substitute goods (Pindyck & Rubinfeld, 2018). Space heating is to be considered as an essential good, which is associated with a minimum demand in the winter months in order to ensure the well-being of residents. Furthermore, the production of space heating depends on the energetic conditions and heating technologies installed in a building, which cannot be easily replaced in the short-term. This is especially true for rental buildings, where the tenants usually have no direct say regarding renovation activities such as building insulation or a renewal of the heating system. Due to these reasons, the price elasticity of space heating demand can be assumed to be rather inelastic – especially for the segment of rental buildings observed in this thesis. The theoretical assumption of a rather inelastic demand reaction is also confirmed by previous evidence from the literature. The majority of studies focusing on the specific energy application of space heating demand in the residential buildings sector arrive at estimates for the price elasticity of space heating demand in the range between -0.3 and -0.6 (see Chapter 2.2).

The results of this thesis confirm the assumption that the demand for space heating is a highly inelastic. In the final model Specification in the analysis based on the full sample, the price elasticity was estimated such that the demand for space heating would decrease by -0.243% if the energy price increased by 1% (see Chapter 4.1). Thus, the estimated elasticity is even lower than in most previous studies. Furthermore, an exploratory analysis of the heterogeneity of the price response based on a subsample of the data and its formalization through the use of an interaction model found heterogeneity in the demand response of the different energy carrier groups. Buildings with oil heating are associated with the lowest demand elasticity of -0.16 [-0.10; -0.23] (Mean estimate [95%CI]), gas with an elasticity in the mid-range of -0.35 [-0.29; -0.40] and district heating with a still inelastic but higher elasticity of -0.53 [-0.46; -0.60] (see Chapter 4.2).

For the broader policy context for the decarbonisation of the residential buildings sector, the estimated highly inelastic demand response implies that households are price takers when it comes to price-related policy instruments for space heating, such as the carbon price signal induced by the BEHG. This entails that households have to spend a larger share of their budget to meet their desired demand for space heating and that this effect becomes even more relevant with increasing energy prices. This effect is particularly relevant for the rental building segment, because under the current design of the BEHG, renters are the sole recipients of the carbon price signal – which in turn means that owners have no incentive to improve the energy conditions of a building or to decarbonise the heat supply. From a policy perspective, there is therefore a need to shift the carbon price signal, at least in part, to building owners to ensure that long-term investment in decarbonising the buildings sector is encouraged alongside short-term demand responses. Furthermore, carbon pricing in the building sector needs to be embedded in a wider policy mix to make the transition to a low-carbon building sector socially desirable and politically feasible. Accompanying and reinforcing instruments include regulatory measures such as the introduction of minimum energy performance standards (MEPS) for buildings and roadmaps for a consecutive phase-out of the installation of new fossil-based heating technologies. Fiscal instruments to promote energy-efficient renovation of the building stock and renewable heating technologies. And the expansion and decarbonisation of district heating supply [SOURCE UBA].

Since the sample data is limited to the rental segment of the buildings sector, the price elasticity estimates presented in this thesis should be considered as elasticities for this specific segment. Due to varying conditions, it can be assumed that the price elasticities for single-family homes, self-owned apartment buildings, and for commercial properties differ from those for rental apartment buildings. By estimating an elasticity for the rental segment specifically, this paper fills a research gap that has been identified in the literature (e.g., Braungardt et al., 2021). Beyond Germany, the elasticity estimates from this work may also be relevant for researchers and policy makers in other countries and regions that have comparable conditions in terms of building stock, climatic conditions and energy demand and also face the challenge of decarbonising the buildings sector in the coming decades.

Limitations

The main limitation related to this work is that it was not possible to control for the socio-economic determinants of space heating demand at the household-level. Previous studies show that household income in particular is a relevant determinant that could not be taken into account together with other relevant factors such as the number of household members, age or education. At the same time, the billing data underlying this work has been shown to also offer an advantage over previous studies based on social surveys, as it provides specific energy demand and prices. Hence, it would be particularly valuable for the future if both specific energy demand and prices on the one hand and socio-economic factors on the other hand were collected at the household level in dedicated household panels, e.g. through the use of smart meters, in order to produce even more reliable estimates of the price elasticity of space heating demand.

Another limitation of this study is related to the risk that the different price elasticities found between the energy carrier groups could be partly attributable to the higher share of fixed heating cost components at lower levels of energy demand. Due to the absence of a breakdown of costs into fixed and variable components, a more detailed investigation was unfortunately not possible.

Further research

Another limitation of the work is that an additional estimation based on an instrumental variable approach was attempted to address a possible simultaneity bias in the data, but was ultimately abandoned because the identifying assumption for a suitable instrument was not met. Further studies could tackle this issue and test other instrumental variables while also using energy billing data, which would potentially eliminate the simultaneity bias.

Another aspect for future research could be to estimate long-term price elasticities of energy demand based on the energy billing data. In this study, due to the limited length of the thesis, only short-term price elasticities of space heating demand were estimated, which represent the conditional demand response. Estimating long-term price elasticities would be an important research question, especially because the specific estimation for the rental segment could verify the assumption made in this thesis that the long-term elasticities for the rental segment are not much higher than the short-term ones and below those of owner-occupied homes.

Appendix A

Supplementary Materials

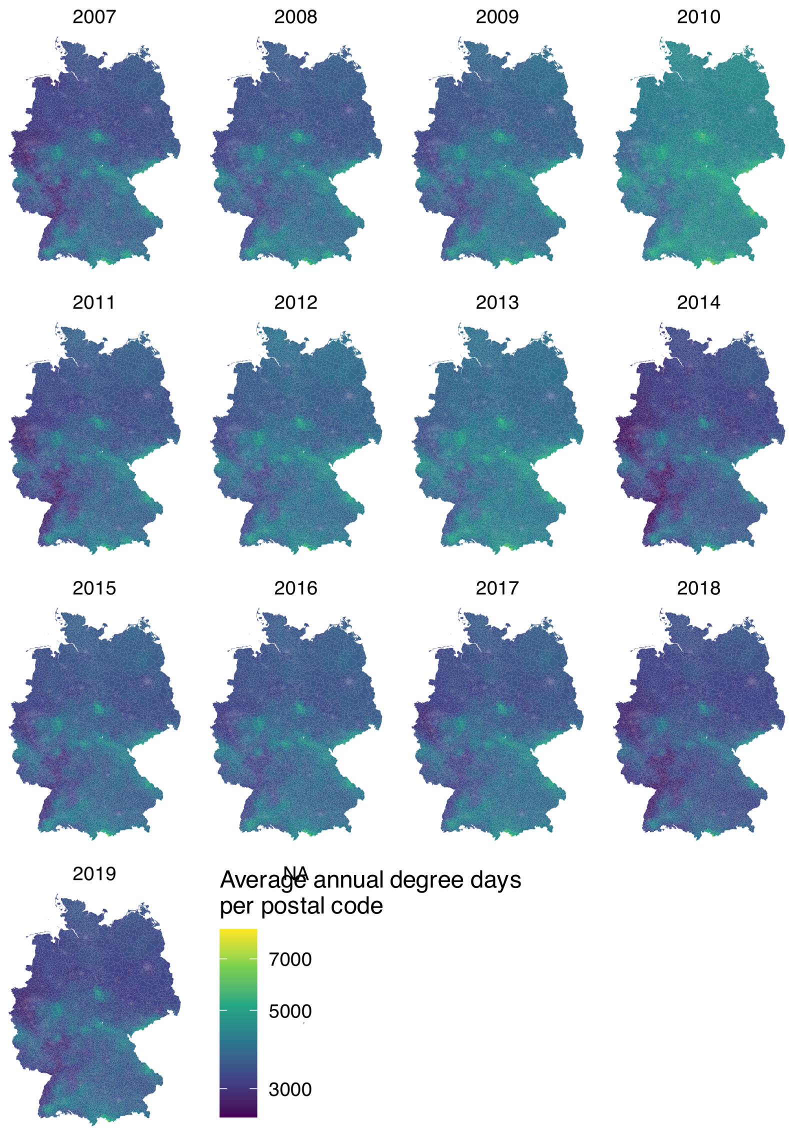

Figure A.1: Spatial and temporal variation in degree days

Figure A.1 shows the spatial and temporal variation of degree days as an operationalisation of outdoor temperature conditions. Degree days data was taken from IWU (2021) and are used as an additional determinant of space heat demand in the regression analysis. Data is shown on the postal code level. Higher annual degree days are associated with lower outdoor temperatures and vice versa. Lower outdoor temperatures are in turn likely to be associated with increased space heating demand. In terms of spatial variation, the maps show colder outdoor temperatures (higher degree-days) mainly in the higher altitude regions, but also in eastern Germany, which is further inland in the European continent. In terms of temporal variation, considerable differences can be observed between the years. Especially 2010, but also 2012 and 2013 were cooler than the average of the years observed. The years 2014, 2018 and 2019, on the other hand, were significantly warmer.

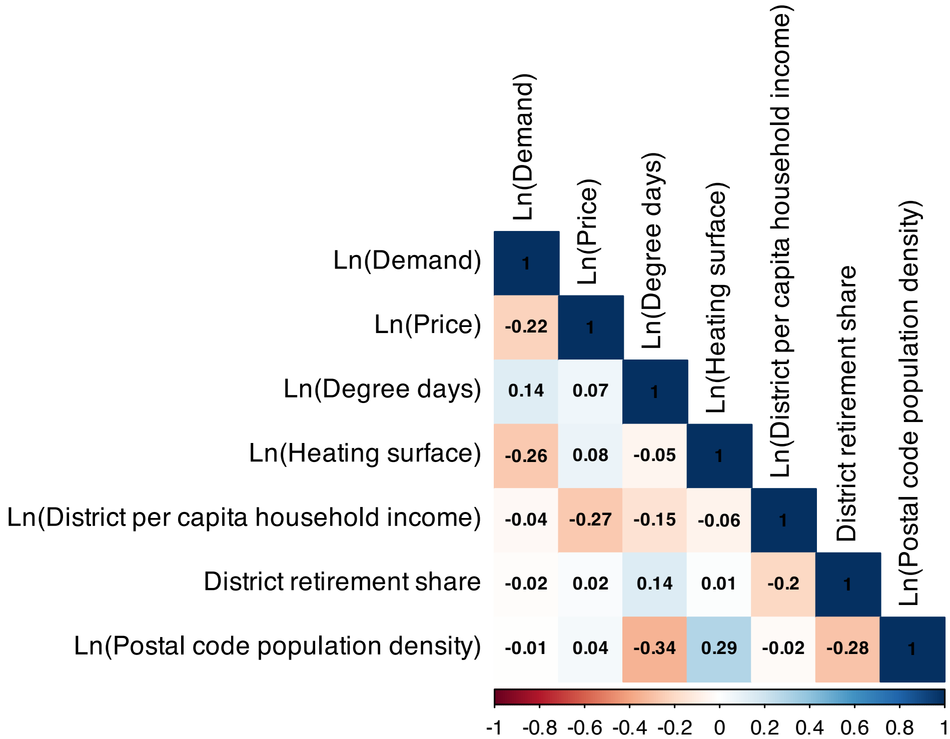

Figure A.2: Matrix with Pearson's correlation coefficients

The values in Figure A.2 represent Pearson's correlation coefficients. Values closer to 1 indicate a stronger positive relationship, values closer to -1 indicate a stronger negative relationship. Values close to 0 indicate no relationship. None of the correlations observed between the variables exceed moderate values, ruling out possible multicollinearity issues. Energy price is negatively correlated with energy demand, indicating that the expected negative effect of price on demand appears to be reflected in the data. Note that all variables except the district retirement share, which is already a ratio, are ln-transformed and reflect the data transformation used in the regression analysis.

Table A.1: Descriptive statistics for age of buildings and heating systems

Variable	Overall N = 2,718,246	Gas N = 1,647,563	Oil N = 802,451	District heating N = 268,232
Building construction year				
Until 1918	34,536 (9.5%)	22,004 (10.0%)	6,622 (6.1%)	5,910 (17%)
1919-1948	22,652 (6.2%)	14,494 (6.6%)	4,742 (4.3%)	3,416 (9.7%)
1949-1978	121,949 (33%)	68,253 (31%)	42,973 (39%)	10,723 (30%)
1979-1995	126,299 (35%)	76,869 (35%)	40,072 (37%)	9,358 (26%)
1996-2009	58,582 (16%)	38,409 (17%)	14,411 (13%)	5,762 (16%)
After 2010	1,172 (0.3%)	738 (0.3%)	288 (0.3%)	146 (0.4%)
(Missing)	2,353,056	1,426,796	693,343	232,917
Heating system installation year				
Until 1978	11,486 (3.5%)	6,360 (3.2%)	4,296 (4.5%)	830 (2.6%)
1979-1995	118,401 (36%)	71,848 (36%)	36,359 (38%)	10,194 (32%)
1996-2009	151,172 (46%)	93,734 (47%)	41,161 (43%)	16,277 (51%)
After 2010	46,556 (14%)	27,564 (14%)	14,532 (15%)	4,460 (14%)
(Missing)	2,390,631	1,448,057	706,103	236,471

Note: n (%)

Table A.1 provides descriptive statistics for the age of buildings and heating systems grouped by the type of energy carrier. The information is taken from the Energy Performance Certificates (EPC) data which is otherwise not used for this thesis, as it is only available for a relatively small share of observations. For those buildings where the information is available, it can be observed from the table above that a relatively larger share of district heating systems are installed in older buildings but at a more recent date. Overall, however, the differences are moderate and likely only a minor determinant for the difference in energy demand between buildings with gas and oil heating on the one side (relatively higher demand) and district heating on the other side (relatively lower demand). Rather, it can be assumed that two other reasons explain most of the observed difference. Firstly, there is no additional loss of efficiency in local combustion in the boiler when using district heating, as is the case with oil and gas. And secondly, the observed buildings with district heating are on average significantly larger than the observed buildings with gas or oil heating (cf. Table 3.2. At the same time, it should be noted that the validity of the age classification of buildings and heating system installation year presented here is limited, as the information is available only for 15.5% and 13.7% of the total number of observations, respectively. This is also the reason why the two variables were not included as variables in the regression analysis.

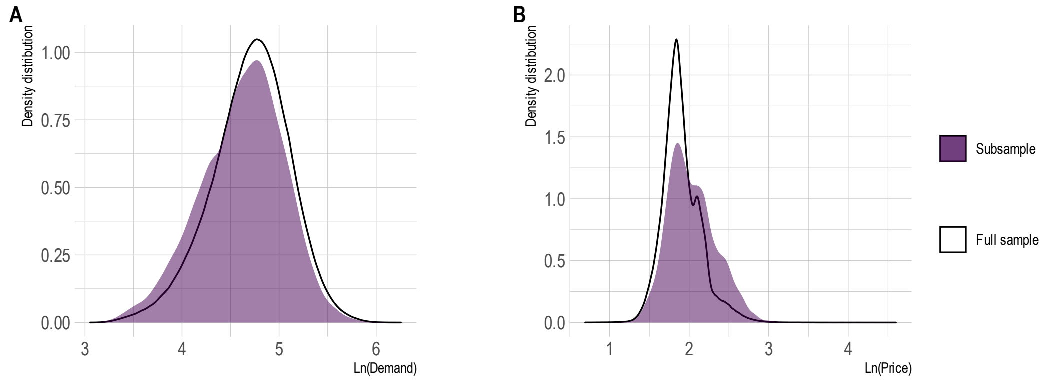

Figure A.3: Comparison of density distribution for energy demand and energy price between full sample and subsample

Figure A.3 compares the density distribution for the two key variables energy demand and energy price between the full sample (transparent) and the subsample (purple area) with 400 randomly selected buildings per energy carrier group (gas, oil and district heating). The visual comparison shows that there are only relatively small deviations in energy demand (Panel A), with the distribution in the subsample being slightly shifted to the left (lower demand). The comparison of distribution in energy prices (Panel B), on the other hand, show more divergence, with the subsample distribution showing a lower peak around $Ln(price) = 2$ and also more observations in the higher price range. However, the deviations in energy demand and energy prices can be well explained by the stratified sampling approach for the subsample. The stratified sample leads primarily to a relative reduction of gas observations and to a relative increase in district heating observations. For the demand variable, more district heating observations lead to more observations with a lower demand. For the price variable, it leads to a less pronounced peak in gas observations at the center (around $Ln(price) = 2$) and more district heating observations with higher prices at the upper end of the distribution.

Table A.2: Complete model summaries for Specifications in subsample analysis

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk ESS	Tail ESS
B.1							
Population-Level-Effects:							
Intercept	5.75	0.03	5.69	5.82	1.00	3240	2312
ln(Price)	-0.56	0.01	-0.59	-0.53	1.00	3240	2432
B.2							
Population-Level-Effects:							
Intercept	4.93	0.49	3.94	5.89	1.00	3423	2799
ln(Price)	-0.41	0.02	-0.45	-0.38	1.00	2740	2644
ln(Degree.days)	0.57	0.04	0.50	0.65	1.00	4404	2986
ln(Heating.surface)	-0.15	0.01	-0.16	-0.14	1.00	3662	3067
Carrier.group.oil	0.04	0.01	0.02	0.06	1.00	3633	2650
Carrier.group.district.heating	-0.06	0.01	-0.09	-0.03	1.00	2588	2890
ln(District.income)	-0.35	0.03	-0.41	-0.29	1.00	4126	2746
ln(Pop.density)	0.05	0.00	0.04	0.05	1.00	3955	3125
District.retire	-0.05	0.17	-0.37	0.28	1.00	4222	3258
BM.1							
Population-Level-Effects:							
Intercept	5.31	0.04	5.23	5.39	1.00	1141	1800
ln(Price)	-0.33	0.02	-0.36	-0.30	1.00	1565	1970
Group-Level-Effects:							
∼id (Number of levels: 1200)							
sd(Intercept)	0.36	0.01	0.34	0.37	1.00	482	990
∼year (Number of levels: 13)							
sd(Intercept)	0.09	0.02	0.06	0.14	1.01	745	1844
BM.2							
Population-Level-Effects:							
Intercept	1.65	0.77	0.12	3.14	1.01	471	1247
ln(Price)	-0.30	0.02	-0.33	-0.27	1.00	2318	2796
ln(Degree.days)	0.69	0.05	0.60	0.79	1.00	1378	1915
Carrier.group.oil	0.06	0.02	0.02	0.10	1.01	443	1067
Carrier.group.district.heating	-0.12	0.02	-0.16	-0.08	1.02	401	930
ln(District.income)	-0.20	0.06	-0.32	-0.08	1.01	364	934
District.retire	-0.18	0.25	-0.67	0.30	1.01	660	1176
Group-Level-Effects:							
∼id (Number of levels: 1200)							
sd(Intercept)	0.35	0.01	0.34	0.37	1.01	461	855
∼year (Number of levels: 13)							
sd(Intercept)	0.03	0.01	0.02	0.05	1.00	1013	1516
BM.2 Lagged							
Population-Level-Effects:							
Intercept	0.88	0.92	-0.97	2.70	1.01	532	975
ln(Price, lagged t-1)	-0.16	0.02	-0.20	-0.12	1.00	2254	2452
ln(Degree.days)	0.72	0.06	0.60	0.83	1.01	1251	1975
Carrier.group.oil	0.04	0.02	0.00	0.09	1.01	364	617
Carrier.group.district.heating	-0.20	0.02	-0.25	-0.16	1.00	460	903
ln(District.income)	-0.17	0.07	-0.31	-0.03	1.01	455	963
District.retire	-0.33	0.28	-0.86	0.22	1.00	586	1219
Group-Level-Effects:							
∼id (Number of levels: 1164)							
sd(Intercept)	0.36	0.01	0.34	0.37	1.00	506	866
∼year (Number of levels: 11)							
sd(Intercept)	0.03	0.01	0.02	0.05	1.00	1116	1753
BM.3							
Population-Level-Effects:							
Intercept	2.13	0.77	0.65	3.71	1.01	557	1248
ln(Price)	-0.35	0.03	-0.40	-0.29	1.00	1309	2241
ln(Degree.days)	0.68	0.05	0.58	0.78	1.01	1106	1872
Carrier.group.oil	-0.32	0.06	-0.44	-0.19	1.01	1175	2149
Carrier.group.district.heating	0.31	0.08	0.15	0.46	1.00	1139	1866
ln(District.income)	-0.23	0.06	-0.35	-0.12	1.02	426	651
District.retire	-0.18	0.27	-0.71	0.34	1.01	370	924
ln(Price):Carrier.group.oil	0.19	0.03	0.13	0.26	1.00	1441	2276
ln(Price):Carrier.group.district.heating	-0.18	0.04	-0.25	-0.11	1.00	1202	2009
Group-Level-Effects:							
∼id (Number of levels: 1200)							
sd(Intercept)	0.35	0.01	0.33	0.36	1.01	386	687
∼year (Number of levels: 13)							
sd(Intercept)	0.03	0.01	0.01	0.04	1.00	912	1752

Table A.2 presents the complete model summaries for the brms models in the subsample analysis. The parameters are summarized using the mean (estimate) and the estimated error of the posterior distribution with the 95% confidence intervals reported. Rhat is a measure of model convergence. Values below 1.05 indicate that chains mixed well and the model has converged. Bulk ESS and Tail ESS are diagnostics of the sampling efficiency.

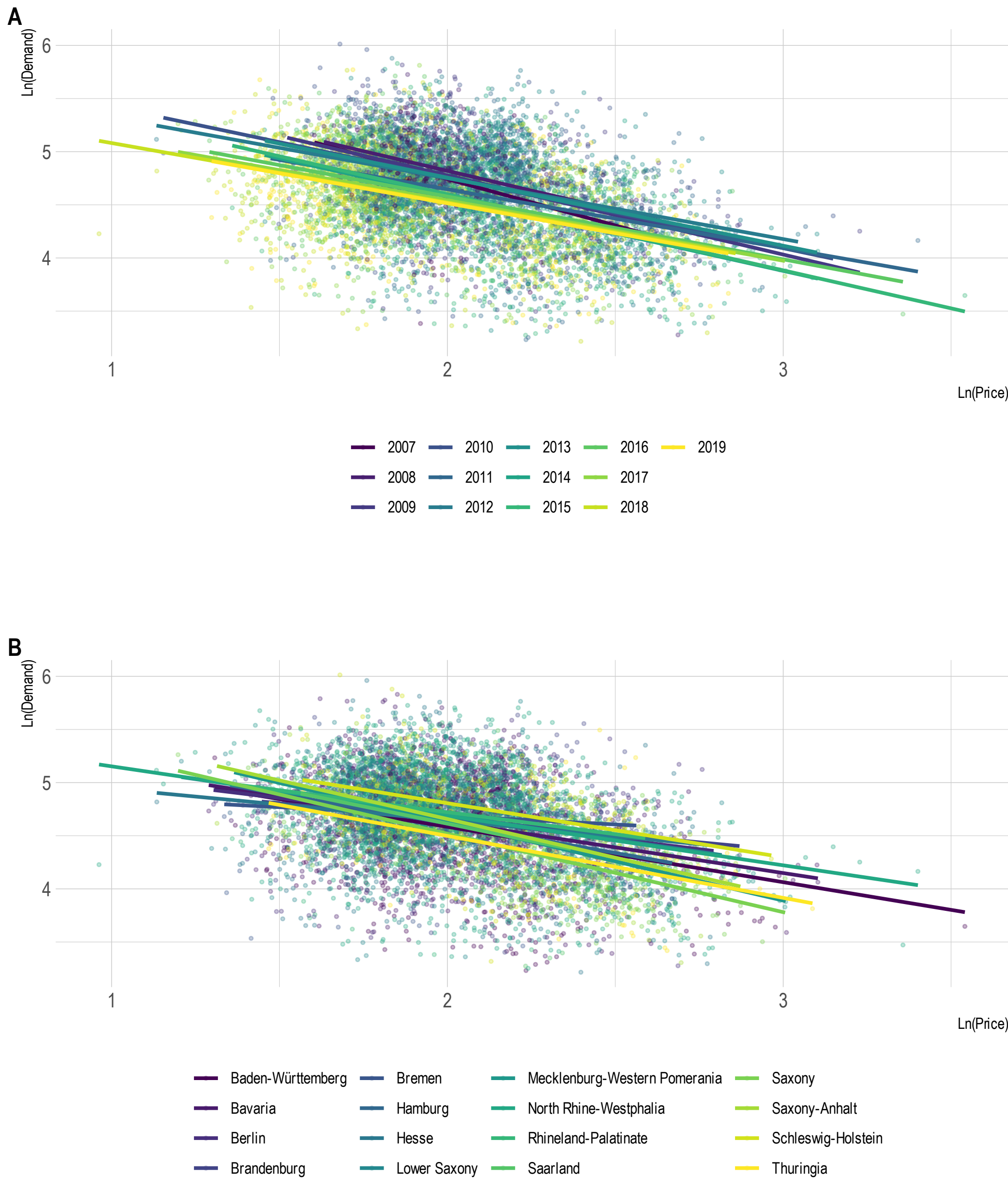

Figure A.4: Scatter plot for the investigation of heterogeneity between years and federal states

Figure A.4 presents a visual analysis designed to detect potential heterogeneity for the price elasticity of space heating demand between years (Panel A) and federal states (Panel B). To construct the graphs, the observations in the subsample are grouped by year or federal state and presented in a scatter plot between energy price and demand. The lines in the diagrams reflect simple linear models for the grouped years or federal states. Between the years observed in the sample, all lines are almost parallel, indicating that there is no relevant difference in price elasticity. In the case of the federal states, the lines diverge a little. But no relevant pattern can be detected (e.g., differences between formerly eastern and western federal states or other regional clusters). Thus, no relevant heterogeneity in price responsiveness of detected for the two variables.

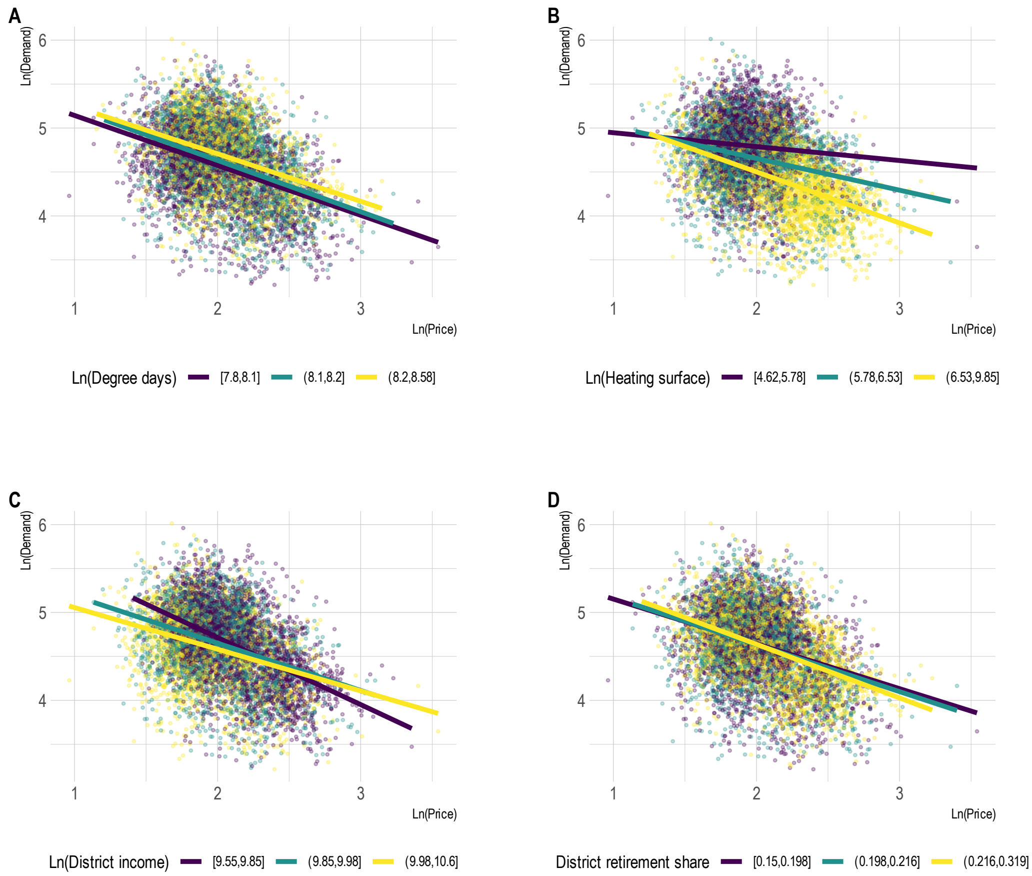

Figure A.5: Scatter plot for the investigation of heterogeneity for the variables degree days, heating surface, district income, and district retirement share

Figure A.5 presents a visual analysis designed to detect potential heterogeneity for the price elasticity of space heating demand for the four variables degree days (Panel A), building heating surface (Panel B), average district income (Panel C), and district retirement share (Panel D). Since all four variables are continuous, the observations in the subsample are grouped into three equally sized groups[1] and presented in a scatter plot between energy price and demand. The lines in the graphs reflect simple linear models for the respective three groups of equal size. For degree days and district retirement share, the lines are almost parallel and therefore indicate no difference in the price elasticity of demand. For district per capita household income minor differences can be observed. For building heating surface a more pronounced difference

[1]Taking the example of degree days, this means that the third of the observations with the lowest degree days (between 7.8 and 8.1), the third of the observations with the medium degree days (between 8.1 and 8.2) and the third of the observations with the highest degree days (between 8.2 and 8.58) are combined as groups.

is observed with larger buildings being associated with a more elastic demand response than smaller buildings. The possible heterogeneity of the price responsiveness for district per capita income and heating surface is further investigated (see Chapter 4.2), but ultimately discarded.

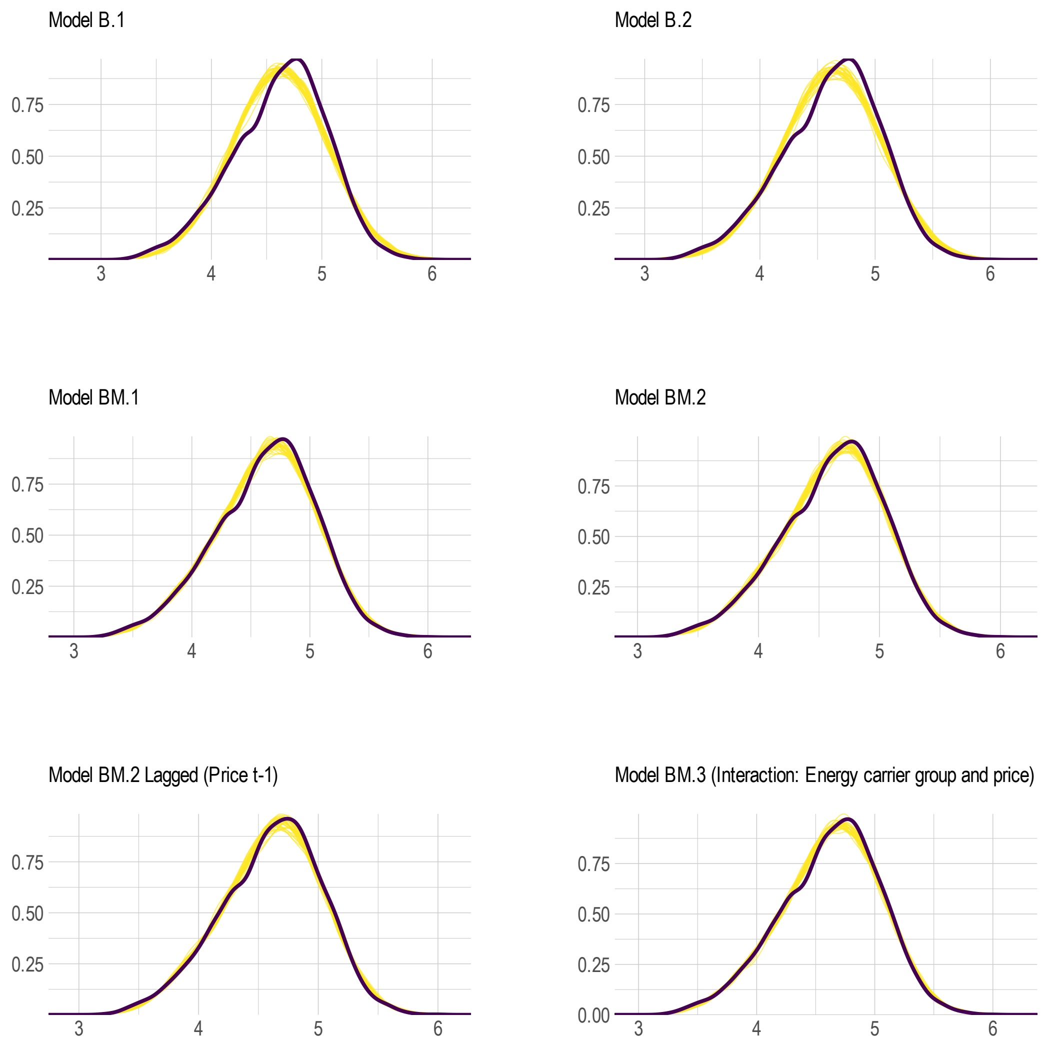

Figure A.6: Graphical model comparison for Specifications in subsample analysis

Figure A.6 presents a visual examination of the model fit for all six models estimated for the subsample analysis. In contrast to the model comparison table presented in Chapter 4.2 here also model Specification (bm.2, Lagged) could be included. One can observe in the graphs, that the predictions (yellow lines) align better with the reference (purple line) once moving to the partial pooling models (bm.x). Among the partial polling models, the differences reported in Table 4.2 are difficult to see visually.

Declaration of Authorship (Eigenständigkeitserklärung)

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst habe und sämtliche Quellen, einschließlich Internetquellen, die unverändert oder abgewandelt wiedergegeben werden, insbesondere Quellen für Texte, Grafiken, Tabellen und Bilder, als solche kenntlich gemacht habe.

Ich versichere, dass ich die vorliegende Abschlussarbeit noch nicht für andere Prüfungen eingereicht habe.

Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung gemäß der fachspezifischen Prüfungsordnung und/oder der Fächerübergreifenden Satzung zur Regelung von Zulassung, Studium und Prüfung der Humboldt-Universität zu Berlin (ZSP-HU) eingeleitet wird.

Berlin, den 04. August 2022

Ort, Datum

Unterschrift

References

AGEB. (2021). *Auswertungstabellen zur Energiebilanz 1990 bis 2020.* Berlin. Retrieved from https://ag-energiebilanzen.de/daten-und-fakten/auswertungstabellen/

Alberini, A., Gans, W., & Velez-Lopez, D. (2011). Residential consumption of gas and electricity in the U.S.: The role of prices and income. *Energy Economics*, *33*(5), 870–881. http://doi.org/10.1016/j.eneco.2011.01.015

Aldy, J. E., Krupnick, A. J., Newell, R. G., Parry, I. W. H., & Pizer, W. A. (2010). Designing Climate Mitigation Policy. *Journal of Economic Literature*, *48*(4), 903–934. http://doi.org/10.1257/jel.48.4.903

Angrist, J. D., & Krueger, A. B. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives*, *15*(4), 69–85. http://doi.org/10.1257/jep.15.4.69

Auffhammer, M., & Rubin, E. (2018). *Natural Gas Price Elasticities and Optimal Cost Recovery Under Consumer Heterogeneity: Evidence from 300 million natural gas bills* (No. w24295) (p. w24295). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w24295.pdf

Bailey, M. A. (2017). *Real econometrics: The right tools to answer important questions.* New York: Oxford University Press.

BMWi. (2021). *Gesamtausgabe der Energiedaten.* BMWi. Retrieved from https://www.bmwi.de/Redaktion/DE/Binaer/Energiedaten/energiedaten-gesamt-xls.xlsx?__blob=publicationFile&v=133

BMWK/BMWSB. (2022). *Sofortprogramm gemäß § 8 Abs. 1 KSG für den Sektor Gebäude 2021.* Retrieved from https://www.bmwsb.bund.de/SharedDocs/downloads/Webs/BMWSB/DE/veroeffentlichungen/bauen/sofortprogramm-sektor-gebaeude.pdf?__blob=publicationFile&v=1

BMWSB. (2022). *Fairness bei den CO2-Kosten.* Retrieved from https://www.bmwsb.bund.de/SharedDocs/pressemitteilungen/Webs/BMWSB/DE/2022/04/co2-preis.html

Braungardt, S., Bürger, V., & Köhler, B. (2021). Carbon Pricing and Complementary Policies—Consistency of the Policy Mix for Decarbonizing Buildings in Germany. *Energies*, *14*(21), 7143. http://doi.org/10.3390/en14217143

Brons, M., Nijkamp, P., Pels, E., & Rietveld, P. (2008). A meta-analysis of the price elasticity of gasoline demand. A SUR approach. *Energy Economics*, *30*(5), 2105–2122. http://doi.org/10.1016/j.eneco.2007.08.004

Bürger, V., Hesse, T., Köhler, B., Palzer, A., & Engelmann, P. (2019). German Energiewende—different visions for a (nearly) climate neutral building sector in 2050. *Energy Efficiency*, *12*(1), 73–87. http://doi.org/10.1007/s12053-018-9660-6

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. http://doi.org/10.18637/jss.v080.i01

Burtless, G., & Hausman, J. A. (1978). The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment. *Journal of Political Economy*, *86*(6), 1103–1130.

Csereklyei, Z. (2020). Price and income elasticities of residential and industrial electricity demand in the European Union. *Energy Policy*, *137*, 111079. http://doi.org/10.1016/j.enpol.2019.111079

Cutler, H. A. (1941). The Elasticity of the Residential Demand for Electricity. *The Journal of Land & Public Utility Economics*, *17*(2), 242–245. http://doi.org/10.2307/3158428

DESTATIS. (2018). Fernwärmeversorgung 2017: Abgegebene Wärmemenge leicht gesunken. Retrieved January 19, 2022, from https://www.destatis.de/DE/Presse/Pressemitteilungen/2018/11/PD18_434_434.html

DESTATIS. (2021a). Bevölkerung: Kreise, Stichtag, Altersgruppen (12411-0017) [Text]. Retrieved January 22, 2022, from https://www-genesis.destatis.de/genesis/online?operation=previous&levelindex=0&step=0&titel=Tabellenaufbau&levelid=1642855066461&acceptscookies=false#abreadcrumb

DESTATIS. (2021b, November 12). Verbraucherpreisindex: Deutschland, Jahre (61111-0001). Retrieved November 12, 2021, from https://www-genesis.destatis.de/genesis//online?operation=table&code=61111-0001&bypass=true&levelindex=1&levelid=1636748679996#abreadcrumb

EP. (2022). *Revision of the Energy Performance of Buildings Directive: Fit for 55 package*. Retrieved from https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/698901/EPRS_BRI(2022)698901_EN.pdf

ERK. (2020). *Bericht zur Vorjahresschätzung der deutschen Treibhausgasemissionen für das Jahr 2020*. Retrieved from https://expertenrat-klima.de/content/uploads/2021/04/210415_Bericht_Expertenrat_Klimafragen_2021-2.pdf

European Comission. (2019). *The European Green Deal* (No. COM(2019) 640 final). Brussels. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52019DC0640&from=EN

Gaure, S. (2013). lfe: Linear group fixed effects. *R J.*, *5*(2), 104.

Gelman, A., Hill, J., & Vehtari, A. (2021). *Regression and other stories*. Cambridge New York, NY Port Melbourne, VIC New Delhi Singapore: Cambridge University Press. http://doi.org/10.1017/9781139161879

Gelman, A., & Stern, H. (2006). The Difference Between "Significant" and "Not Significant" is not Itself Statistically Significant. *The American Statistician*, *60*(4), 328–331. http://doi.org/10.1198/000313006X152649

Gwartney, J. D. (1976). Demand and Consumer Choice. In *Economics Private and Public Choice* (pp. 289–309). Elsevier. http://doi.org/10.1016/B978-0-12-311050-3.50021-8

Halbig, G., & Namyslo, J. (2014). Neue Witterungsbereinigung für Energieausweise auf der Basis des Referenzklimas Potsdam. *EneV aktuell*, (4), 11–13.

Havranek, T., Irsova, Z., & Janda, K. (2012). Demand for gasoline is more price-inelastic than commonly thought. *Energy Economics*, *34*(1), 201–207. http://doi.org/10.1016/j.eneco.2011.09.003

Hennes, O., Jeddi, S., Madlener, R., Schmitz, H., Wagner, J., Wolff, S., & Zinke, J. (2021). Auswirkungen von CO2-Preisen auf den Gebäude-, Verkehrs- und Energiesektor. *Zeitschrift für Energiewirtschaft*, *45*(2), 91–107. http://doi.org/10.1007/s12398-021-00305-0

Heße, W. (2020). *Energieeffiziente Wärmeversorgung von Gebäuden: Tatsächliche Versorgungsverhältnisse und Maßnahmen zur Effizienzsteigerung*. Wiesbaden:

Springer Fachmedien Wiesbaden. http://doi.org/10.1007/978-3-658-27571-6

Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, *81*(396), 945–960. http://doi.org/10.1080/01621459.1986.10478354

Houthakker, H. S. (1951). Some Calculations on Electricity Consumption in Great Britain. *Journal of the Royal Statistical Society. Series A (General)*, *114*(3), 359–371. http://doi.org/10.2307/2980781

IWU. (2021, October 7). IWU-Tool „Gradtagzahlen-Deutschland.xlsx“. Retrieved January 22, 2022, from `https://www.iwu.de/publikationen/fachinformationen/energiebilanzen/`

Johnson, D. H. (1999). The Insignificance of Statistical Significance Testing. *The Journal of Wildlife Management*, *63*(3), 763. http://doi.org/10.2307/3802789

Kröger, M., Longmuir, M., Neuhoff, K., & Schütze, F. (2022). The Costs of Natural Gas Dependency: Price Shocks, Inequality, and Public Policy. *DIW Discussion Paper*, (2010). Retrieved from `https://www.diw.de/documents/publikationen/73/diw_01.c.847604.de/dp2010.pdf`

Labandeira, X., Labeaga, J. M., & López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy Policy*, *102*, 549–568. http://doi.org/10.1016/j.enpol.2017.01.002

Lange, I., Moro, M., & Traynor, L. (2014). Green hypocrisy?: Environmental attitudes and residential space heating expenditure. *Ecological Economics*, *107*, 76–83. http://doi.org/10.1016/j.ecolecon.2014.07.021

Leth-Petersen, S., & Togeby, M. (2001). Demand for space heating in apartment blocks: measuring effects of policy measures aiming at reducing energy consumption. *Energy Economics*, *23*(4), 387–403. http://doi.org/10.1016/S0140-9883(00)00078-5

Levesque, A., Pietzcker, R. C., Baumstark, L., & Luderer, G. (2021). Deep decarbonisation of buildings energy services through demand and supply transformations in a 1.5°C scenario. *Environmental Research Letters*, *16*(5), 054071. http://doi.org/10.1088/1748-9326/abdf07

Loulou, R., & Labriet, M. (2008). ETSAP-TIAM: the TIMES integrated assessment model Part I: Model structure. *Computational Management Science*, *5*(1-2), 7–40. http://doi.org/10.1007/s10287-007-0046-z

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd ed.). Chapman and Hall/CRC. http://doi.org/10.1201/9780429029608

Meier, H., & Rehdanz, K. (2010). Determinants of residential space heating expenditures in Great Britain. *Energy Economics*, *32*(5), 949–959. http://doi.org/10.1016/j.eneco.2009.11.008

Metcalf, G. E., & Hassett, K. A. (1999). Measuring the Energy Savings from Home Improvement Investments: Evidence from Monthly Billing Data. *Review of Economics and Statistics*, *81*(3), 516–528. http://doi.org/10.1162/003465399558274

Miller, M., & Alberini, A. (2016). Sensitivity of price elasticity of demand to aggregation, unobserved heterogeneity, price trends, and price endogeneity: Evidence from U.S. Data. *Energy Policy*, *97*, 235–249. http://doi.org/10.1016/j.enpol.2016.07.031

Nichols, A. (2007). Causal Inference with Observational Data. *The Stata Journal: Promoting Communications on Statistics and Stata*, *7*(4), 507–541. http://doi.org/10.1177/1536867X0800700403

OSM. (2021a). Einwohnerzahl auf PLZ-Gebiete abbilden. Retrieved from `https://blog.suche-postleitzahl.org/post/132153774751/einwohnerzahl-auf-plz-gebiete-abbilden`

OSM. (2021b). Postleitzahlen Deutschland. Retrieved from `https://www.suche-postleitzahl.org/downloads`

Pindyck, R. S., & Rubinfeld, D. L. (2018). *Microeconomics* (9th Edition). New York, NY: Pearson.

Rehdanz, K. (2007). Determinants of residential space heating expenditures in Germany. *Energy Economics*, *29*(2), 167–182. http://doi.org/10.1016/j.eneco.2006.04.002

Ruhnau, O., Stiewe, C., Muessel, J., & Hirth, L. (2022). *Gas demand in times of crisis. The response of German households and industry to the 2021/22 energy crisis*. Kiel, Hamburg: ZBW – Leibniz Information Centre for Economics. Retrieved from `http://hdl.handle.net/10419/261082`

RWI. (2021). *Anwendungsbilanzen 2020 für den Sektor der privaten Haushalte und den Verkehrssektor in Deutschland*. Retrieved from `https://ag-energiebilanzen.de/wp-content/uploads/2020/10/rwi_anwendungsbilanz_2020__priv._hh_und_verkehr__vorl._eb.pdf`

Schmitz, H., & Madlener, R. (2020). Heterogeneity in price responsiveness for residential space heating in Germany. *Empirical Economics*, *59*(5), 2255–2281. http://doi.org/10.1007/s00181-019-01760-y

Schulte, I., & Heindl, P. (2017). Price and income elasticities of residential energy demand in Germany. *Energy Policy*, *102*, 512–528. http://doi.org/10.1016/j.enpol.2016.12.055

Stan Development Team. (2022). Documentation. Retrieved from `https://mc-stan.org/users/documentation/`

Statistische Ämter. (2021). Einkommen der privaten Haushalte in den kreisfreien Städten und Landkreisen der Bundesrepublik Deutschland 1995 bis 2019. Retrieved November 12, 2021, from `http://www.statistikportal.de/de/vgrdl/ergebnisse-kreisebene/einkommen-kreise`

Stede, J., Schütze, F., & Wietschel, J. (2020). Wärmemonitor 2019: Klimaziele bei Wohngebäuden trotz sinkender CO2-Emissionen derzeit außer Reichweite (Version 2.0). *DIW Wochenbericht*. http://doi.org/10.18723/DIW_WB:2020-40-1

Stigler, G. J. (1987). *The theory of price* (4th ed). New York : London: Macmillan ; Collier Macmillan.

Stiglitz, J. (2019). *Addressing Climate Change through Price and Non-Price Interventions* (No. w25939) (p. w25939). Cambridge, MA: National Bureau of Economic Research. Retrieved from `http://www.nber.org/papers/w25939.pdf`

Techem. (2019). *Energiekennwerte 2019*. Eschborn: Techem. Retrieved from `https://www.techem.com/content/dam/techem/newsroom/studien/Techem-Energiekennwerte-Studie-2019.pdf`

UBA. (2016). *CO2-Emissionsfaktoren für fossile Brennstoffe*. Retrieved from `https://www.umweltbundesamt.de/sites/default/files/medien/1968/publikationen/co2-emissionsfaktoren_fur_fossile_brennstoffe_korrektur.pdf`

UBA. (2022). *Klimaschutzbeitrag verschiedener CO2-Preispfade in den BEHG-Sektoren Verkehr, Gebäude und Industrie* (p. 55). Retrieved from `https://www.umweltbundesamt.de/publikationen/klimaschutzbeitrag-verschiedener-co2-preispfade-in`

Varian, H. R. (2010). *Intermediate microeconomics: a modern approach* (8th Edition). New York, NY: Norton.

VDI. (2013). *Richtlinie 3807 Blatt 1: Verbrauchskennwerte für Gebäude*. Retrieved from https://www.vdi.de/richtlinien/details/vdi-3807-blatt-1-verbrauchskennwerte-fuer-gebaeude-grundlagen-1

Vehtari, A., Gabry, J., Yao, Y., & Gelman, A. (2018). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. *R Package Version*, *2*(0), 1003. Retrieved from https://cran.r-project.org/web/packages/loo/index.html

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. http://doi.org/10.1007/s11222-016-9696-4

Verbraucherzentrale NRW. (2016). *Heizen mit Fernwärme: Klimafreundlichkeit, Kosten und Gesetze*. Retrieved from https://www.verbraucherzentrale.de/sites/default/files/2019-03/Fernwaermebroschuere_V06_final.pdf

Wooldridge, J. M. (2013). *Introductory econometrics: a modern approach* (5th edn). Mason: South-Western Cengage Learning.

Zhu, X., Li, L., Zhou, K., Zhang, X., & Yang, S. (2018). A meta-analysis on the price elasticity and income elasticity of residential electricity demand. *Journal of Cleaner Production*, *201*, 169–177. http://doi.org/10.1016/j.jclepro.2018.08.027