

Singhal, Puja; Hobbs, Andrew

Conference Paper

The Distribution of Energy Efficiency and Regional Inequality

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Singhal, Puja; Hobbs, Andrew (2021) : The Distribution of Energy Efficiency and Regional Inequality, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/242343>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

The Distribution of Energy Efficiency and Regional Inequality*

Andrew Hobbs[†] and Puja Singhal[‡]

April 2021

Abstract

This paper studies the long-term distribution of energy-efficiency outcomes in the German residential sector. To uncover the underlying energy efficiency of buildings, we estimate the causal response of building-level heat energy demand to variability in heating degree days. We examine heterogeneity in temperature response using both panel fixed-effects and causal forests. Our results suggest that the distribution of energy-efficiency is not equitable in the West of Germany, with buildings located in the South attaining the best energy performance standards. Although the housing stock in the East is significantly older and thus less subject to building standards, they perform better than the West counterpart, likely as a result of large investments in retrofitting post-reunification. Finally, we show that the regional distribution of energy-efficiency reflects differences in heating needs – thus, the poorer energy performance of buildings in the North-West should be weighed against the warmer climatic zone.

Keywords: Heat Demand, Energy Efficiency, Targeting, Regional Distribution, Climate Change

JEL Codes: H23, Q52, Q48, Q54

*We would like to thank seminar participants from the SOEP Applied Panel Group at DIW Berlin, MCC Berlin and PIK for helpful comments. The paper used restricted-access data from ista Deutschland GmbH available at the German Institute for Economic Research (DIW Berlin).

[†]University of San Francisco, email: ahobbs@usfca.edu.

[‡]Corresponding author, e-mail: singhal@pik-potsdam.de. Address: Potsdam Institute for Climate Impact Research (PIK), P.O. Box 601203, 14412 Potsdam, Germany. Funding from the Kopernikus-Projekt Ariadne (FKZ 03SFK5A) by the German Federal Ministry of Education and Research is gratefully acknowledged.

1 Introduction

Residential heating continues to be primarily powered with fossil-fuels¹ and accounts for one-fifth of the final energy consumed in Germany (BMWi, 2019). Energy-efficiency has remained a key policy objective to reduce the dependence on fossil fuels in the building sector². Even though energy efficiency measures play an increasing role in achieving GHG emission reductions globally, empirical evaluations of policy instruments such as building standards, retrofit programs, and mandatory energy performance certificates seldom take into account the long-term distribution of energy-savings achieved.

Housing-related policies may have particularly strong distributional effects and thus may have the additional objective of economic equity such that the costs and benefits of environmental policy are allocated progressively across income groups or regions (Bento, 2013). Evidence suggests that investments towards energy efficiency vary significantly by income groups and thus the improvements in energy intensity are likely distributed unequally as well. Borenstein and Davis (2016) and Jacobsen (2019) show that financial incentives such as income tax credits for home weatherization, hybrid or electric cars, solar systems, and energy-efficient appliances are predominantly received by higher-income households in the United States. McCoy and Kotsch (2020) investigate the returns to energy-efficiency investments by income group and find that the poorer households experience lower energy savings both in the short- and long-run in the United Kingdom. Similarly, Bruegge et al. (2019) find that building codes in California led to energy-savings for the lower income groups, but this was due to decrease in square footage rather than improvements in building energy-efficiency attributes.

We add to this literature on the distributional impacts of energy policies in the residential building sector. Here we take a long-term perspective and assess the net outcomes of major policy efforts (energy standards for new construction and renovations, financial subsidies for home retrofits) that have already taken place in Germany and come to a nuanced conclusion. Overall, we find that zip

¹81% of final energy consumed for space heating was produced using natural gas, heating oil, district heating, and coal (AG Energiebilanzen e.V. 2018).

²However, evaluations have shown that actual savings realized from energy-efficiency programs seldom achieve the level of energy savings that are predicted using ex-ante engineering models (Fowle et al. 2018, Giraudet et al. 2018, Christensen et al. 2020)

codes that are associated with the best socio-economic outcomes are home to a disproportionately higher share of energy-efficient buildings. However, the poorest states of Germany, in the East, also benefit from both the use of less carbon-intensive heating fuel type and energy-efficient buildings. A closer look reveals that the spatial distribution of energy-efficiency outcomes is well-explained by regional differences in the intensity of the winter heating season across Germany.

The contribution of this paper is three-fold.

First, we propose a methodological improvement over ratings on energy performance certificates to measure energy-efficiency of buildings. We combine heating bills for a large sample of the existing building stock in Germany with daily temperature data from weather stations to underpin the causal response heat energy demand to short-run fluctuations in the total annual sum of heating degree days. Using this building-level response of heat energy consumption to temperature as an overall measure of energy-efficiency, we are able to simultaneously investigate the regional and socio-economic distribution of outcomes. We thus demonstrate how access to building-specific billing data allowed us to accurately assess building-level performance, which can then be used to identify buildings most in need for energy-efficiency measures.

Second, we document that temperature change is an increasingly dominant factor in explaining the observed decline in heat energy use by the German housing sector in the last decade, and discuss potential implications for energy efficiency investment incentives. Global surface temperatures have been rising steadily relative to historical average temperatures (NASA). Germany, in particular, experienced an unwavering increase in average temperatures in the past decade from 2010 to 2019³. Unsurprisingly, the heating sector is particularly affected by global warming because the demand for energy to heat homes falls, *ceteris paribus*⁴. We quantify the fall in energy demand for space heating due to climate change (8 percent from 2014 to 2018), and thus the weakening economic incentives for homeowners to invest in thermal-efficiency of the existing building stock.

Third, we provide new evidence on the distribution of realized energy efficiency outcomes. We in-

³The German Weather Service (DWD) reported 2018 to be the warmest year in its 138-year temperature records, and 2019 tied with 2014 were the second-warmest years in German history.

⁴Heating degree days fell by between 11 to 7 percent (depending on the region) from 2013 to 2018 relative to the previous decade, 2003 to 2012. See Appendix Figure C1.

investigate comprehensively the main drivers of the heterogeneous response to temperature shocks, allowing us to gain insight into the mechanisms driving estimated differences in energy-efficiency. To this end, we employ machine learning to explore all observable dimensions in a systematic manner. We shed light on the heterogeneity in outcomes by socioeconomic disparities, in particular the East-West divide in Germany. By focusing on the existing building stock with average year of construction in 1967, we also circumvent the problem of measuring energy-efficiency outcomes in a short-time frame as highlighted by [Kotchen \(2017\)](#). We show that although the newer and more energy-efficient buildings are located in the South of Germany, buildings in the poorest East regions of Germany are surprisingly energy-efficient, owing to larger buildings, renovation efforts, and efficient construction that took place post-reunification in the between 1990 and 2001. For instance, we document that the older building stock (built before 1975) in the East of Germany is significantly more energy-efficient than the comparable group in the West. At the same time, East buildings are on average larger than those in the West of Germany and this is further associated with less energy per square meter of living space for each heating degree day.

The next section discusses the unique data used in this paper. Section [3](#) reports on the underlying trends in heat demand and heating degree days. Section [4](#) explains the identification strategy and estimates the average response of energy demand to temperature. Section [5](#) considers the heterogeneous responses to temperature. Section [6](#) offers a forest-based machine learning method to estimate building-level marginal effects to further understand the sources of heterogeneity. Section [7](#) discusses what may explain the regional distribution of energy-efficiency and Section [8](#) concludes.

2 Data

The analysis in this paper is based on data combined from three sources: (1) data on building-level heating bills and energy performance certificates from a leading energy-metering company, (2) weather station data from the German Weather Service (Deutscher Wetterdienst), and (3) socio-demographic data from RWI-GEO-GRID ([Breidenbach and Eilers, 2018](#)).

2.1 Heating Bills

The primary data used come from a large panel of building-level heating bills for 420,573 residential buildings (3,215,800 bills) in Germany, with 12-month billing cycles that start during January 2008 to June 2018. 12-month billing means that all heating bills are for either 365 or 366 days, but the billing start and end dates vary.

The billing dataset contains information on the actual (metered) units of energy consumed for space heating and water heating, along with yearly costs incurred. The billing data also contains important building characteristics that help determine the energy requirements of buildings: living space (in square meters), building size (in number of apartments), location by zip code, and main heating fuel type.

The main dependent variable is calculated as the annual quantity of heating energy consumed per square meter of heated living space in a building. This took several steps: first, building-specific consumption values are limited to the amounts of energy used for heating space (excluding warm water). Second, the metered consumption value is multiplied by the net calorific value corresponding to the building's energy fuel type, giving us the absolute heat energy consumption in kilowatt-hours (kWh) for a building during the billing period. Fourth, we divide total kilowatt-hours consumed by the amount of heated living space in the building. The units are therefore, kilowatt-hours per square meter of heated living space per year (kWh/sqm).

We calculate heating costs used for the heating energy consumed by first deducting costs of heating water from the total energy costs reported in each bill. Then dividing total costs for space heating by the amount of heated living space in the building gives us the cost per kWh of heat energy billed.

To create the estimation sample, we only consider heating bills from buildings that use either natural gas, heating oil, or district heating as the main fuel type, which is 98 percent of all buildings observed. We trim the sample further by removing the top and bottom 1% tails from the distribution of heat energy consumption, used as the main dependent variable: such that consumption is above 30 kWh/sqm and below 400 kWh/sqm of heated living space.

Finally, we only consider those buildings observed at least two times in the (unbalanced) panel. After these steps, the full sample consists of 384,223 buildings with a total of 3,030,063 observed heating bills. On average we observe a building 9 times, minimum number of 2 times, and a maximum number of 11 times.

2.2 Energy Performance Certificates

For about 40 percent of the buildings, we observe energy performance certificates issued from 2008 to 2019 that give us important measures for the thermal-efficiency attributes of the buildings, including the energy performance score, construction year of the building, year (or renewal year) of the heating system, roof, loft ceiling, exterior wall, windows, and basement ceiling. For energy performance certificates issued from 2014 to 2019 (about 20% of the sample), I further observe whether individual building components meet thermal insulation standards set out in the national thermal insulation ordinance from 1995 ("Wärmeschutzverordnung 1995 or WSVVO 1995).

2.3 Supplementary Data

We supplement the energy-related data with data from local weather stations collected by the German meteorological service (Deutscher Wetterdienst). To construct variables that capture the number heating degree days, we find the nearest available weather station to 8303 geocoded zip codes of Germany, provided that there is not more than one consecutive daily observation record missing for mean temperature for each weather station from 2003 to 2019. For the few missing values, we impute using the average of mean temperatures recorded for the previous and next day. This procedure amounts to using daily mean temperatures from 204 weather stations scattered across Germany to calculate heating degree days corresponding to each billing month and cycle.

The socio-economic variables used in this paper are the unemployment rate, housing density, and the purchasing power⁵ per household computed using high-resolution grid level (1x1 kilometer cells) data ([RWI and microm 2020](#), [Breidenbach and Eilers 2018](#)), averaged at the zip code level

⁵A measure of disposable income – "[t]he variable purchasing power reflects the household income. It comprises information on labour supply, capital wealth, rental and leasing income minus taxes and social security contributions, including social transfers such as unemployment benefits, child-allowances and pensions." ([Breidenbach and Eilers, 2018](#))

and matched to the billing sample. The data is available for 2005 and 2009 to 2017.

2.4 Descriptive Statistics

In this section, we describe aggregate trends in key variables observed from the data. Table 1 presents summary statistics for the full sample at the yearly level for important determinants of heat energy demand. We report averages for the East of Germany in parentheses.

Based on this table, there is evidence that financial returns for heat energy-efficiency investments have been falling. Along with warmer heating seasons, costs of heating fuel (cents per kWh) have not risen significantly, resulting in declining annual heating expenditures per square meter (energy use x price) since 2008. Meanwhile, information from the sample of energy performance certificates suggests that thermal-efficiency and insulation standards have not increased noticeably for the existing housing stock.

These trends are also true for the East of Germany. However, there is a noteworthy exception. Although the buildings in the East of Germany are on average older, a higher share of buildings (about half) observed with energy performance certificates from 2014 to 2019 met thermal insulation standards set out in WSVO 1995.

To make this distinction clearer, we split up the sample to those buildings built prior to 1995 and those built starting 1995, and then look at the shares that meet the WSVO 1995 standard. Figure 1 shows that a significantly higher share of the East building stock built before 1995 meet thermal insulation standards under WSVO 1995. In order to meet the WSVO 1995 standards, buildings built before 1995 likely underwent higher rates of retrofitting. At the same time, most buildings built after 1995 are in compliance with the regulation. In Figure C2, we further graph the differences in certification rates by each individual building component.

What may explain this stark difference between the East and the rest of Germany? Since reunification of Germany in 1990, there had been a concerted political and financial effort (via subsidies for renovation and new construction) to improve the quality of the housing stock in East, at least until 2001 (Weiß and Michelsen, 2011). Using identical data on energy performance certificates, Singhal and Stede (2019) document that the thermal retrofit rate was significantly higher for East

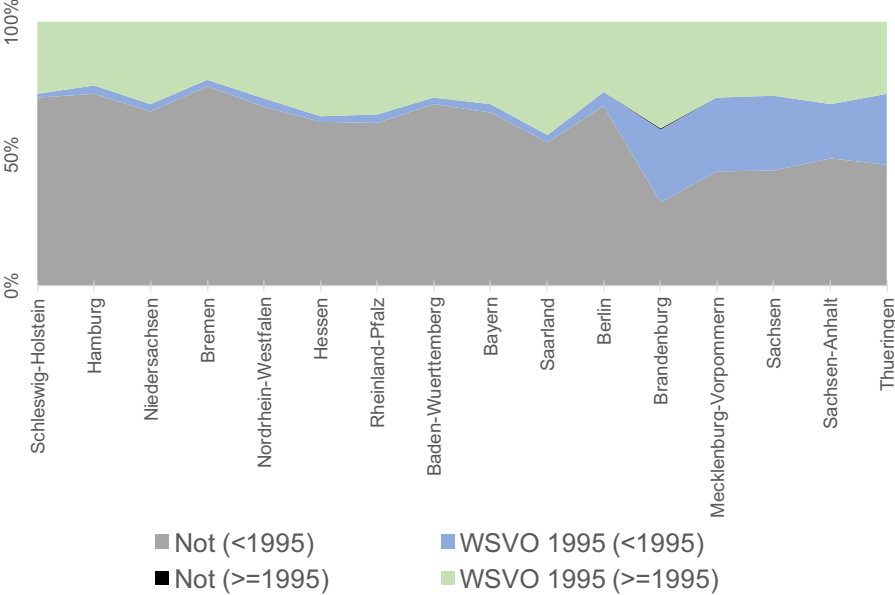
Table 1: Summary Statistics

Bill Start	Energy Use, Price, Building Size		Heating Degree Days		Energy Performance Certificates			
	kWh/m ²	Cents/kWh	# of Apts	Local	Δ from "Typical"	Vintage	WSVO 1995	Heat System
2008	128.5 (116.7)	6.8 (7.6)	9.1 (15.8)	2283.8	-110.3	1966.6 (1949.2)	0.36 (0.52)	1996.5 (1996.7)
2009	128.6 (119.2)	6.4 (7.2)	9.6 (17.1)	2362.4	-31.8	1966.4 (1949.9)	0.37 (0.51)	1996.5 (1996.7)
2010	128.6 (132.4)	5.9 (6.6)	9.8 (17.2)	2652.7	258.5	1966.9 (1950.3)	0.37 (0.52)	1996.6 (1996.8)
2011	118.3 (112.6)	6.5 (7.3)	9.6 (17.1)	2137.6	-256.6	1966.5 (1950.1)	0.37 (0.52)	1996.6 (1996.9)
2012	125.0 (117.2)	7.1 (7.8)	9.7 (17.3)	2405.1	10.9	1966.6 (1950.3)	0.37 (0.52)	1996.7 (1996.8)
2013	122.8 (119.0)	7.2 (7.8)	9.9 (17.4)	2336.9	-57.2	1966.5 (1950.4)	0.38 (0.52)	1996.8 (1996.9)
2014	106.9 (99.8)	7.0 (7.7)	10.1 (17.3)	1947.0	-447.2	1966.2 (1950.4)	0.38 (0.52)	1996.8 (1997.0)
2015	112.5 (105.1)	6.5 (7.2)	10.0 (17.2)	2071.1	-323.0	1966.5 (1950.6)	0.38 (0.52)	1996.9 (1997.0)
2016	118.8 (111.2)	5.9 (6.7)	10.2 (17.7)	2261.0	-133.2	1966.9 (1950.9)	0.38 (0.51)	1996.9 (1997.0)
2017	117.6 (109.9)	5.6 (6.3)	10.4 (18.1)	2159.5	-234.6	1967.6 (1951.5)	0.38 (0.52)	1997.1 (1997.2)
2018	111.4 (103.2)	5.7 (6.4)	11.0 (18.6)	2066.0	-328.2	1969.3 (1953.3)	0.39 (0.53)	1997.2 (1997.2)
All Years	120.7 (113.3)	6.4 (7.1)	9.9 (17.4)	2248.2	-146.0	1966.8 (1950.6)	0.37 (0.52)	1996.8 (1996.9)

Notes: The table reports statistics for the full sample. All values are averages, calculated for each year in which the 12-month billing period starts for the observed building. The values in parentheses are average for the East of Germany located in Berlin, Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, and Thüringen. The second and third columns report on building-level energy use and heating costs for the full sample of buildings described in section 3.1. Energy use is measured as the annual heat energy consumption per square meter of heated living space. Price of heating fuel is given in Euro per kWh and observations above 1 Euro per kWh were removed. Building size is given by the column indicating the number of apartments. The "Local" column reports the average annual sum of heating degree days (with daily mean temperature below 15 degree Celsius) recorded at the nearest weather station from 20080101 to 20181231, while the second column under "Typical" reports the difference in the annual sum of heating degree days recorded at the local station compared to that recorded from 1995 to 2012 at Potsdam. The last three columns report data from energy performance certificates. The WSVO 1995 column reports the share of buildings reporting energy performance certificates from 2014 to 2019 that meet thermal insulation standards effective in 1995.

German buildings during the 90s until 2001. Renovation investments made in the East housing sector in the 1990s are likely driving the improved energy performance of the building stock we observe today.

Figure 1: WSVO 1995 Standards



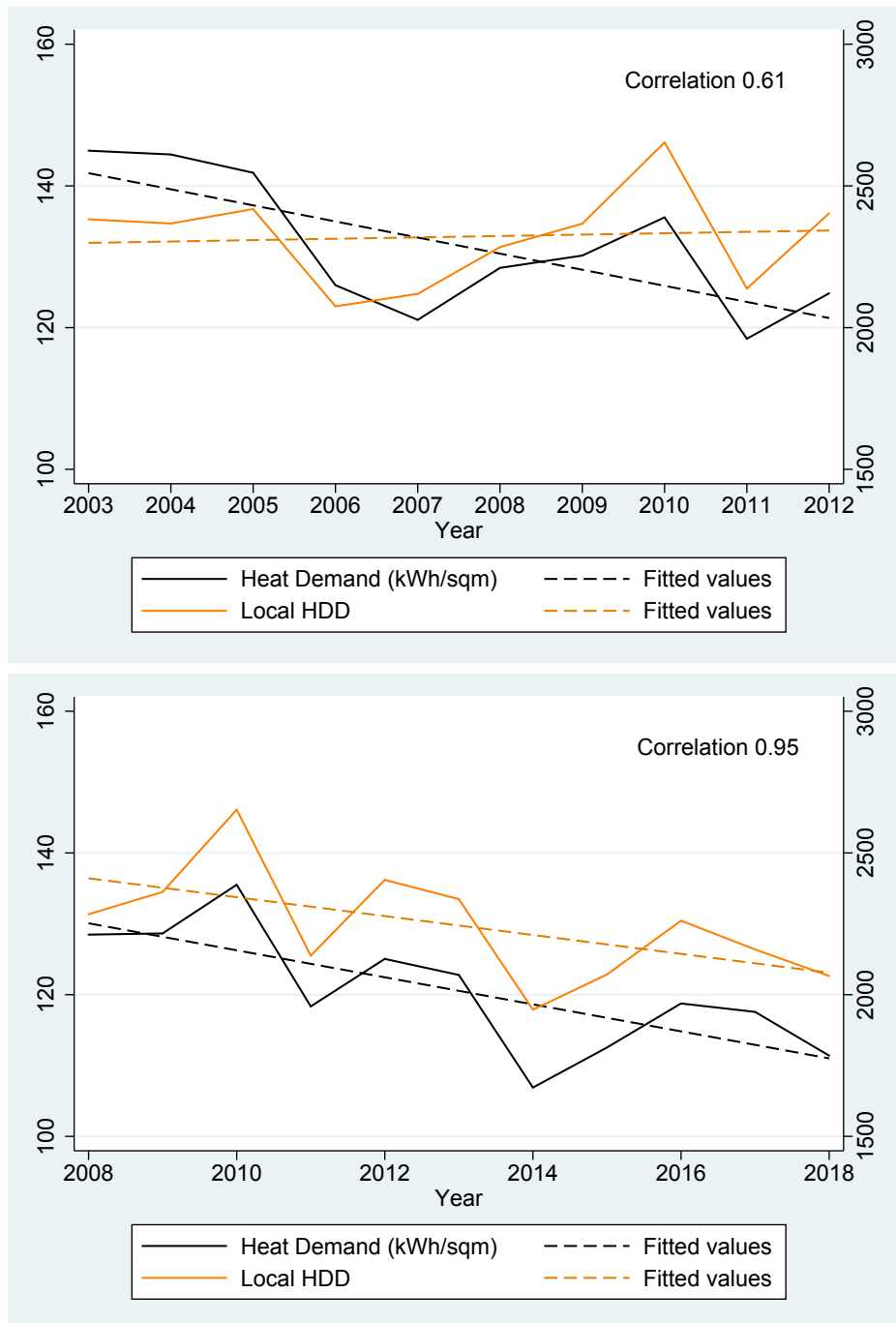
Notes: The graph summarizes the share of buildings that are certified to meet the thermal insulation standards of 1995 (WSVO 1995) by year of construction (built before or after 1995) for each of the sixteen states in Germany (marked on the horizontal axis). The East of Germany consists of Berlin, Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, and Thueringen. A building is defined to meet the 1995 insulation standards, if it is certified that all five building components (the roof, top ceiling, outer wall, windows, and basement) meet the 1995 minimum energy standards. The grey colored area indicates the share of buildings by state that were built before 1995 and do not meet WSVO 1995 insulation standards. Here we clearly show that buildings, built before 1995, in the East German states were much more likely to be renovated to meet the WSVO 1995 standard. We also show that the almost all buildings built after 1995 passed the certification for the WSVO 1995 standard.

3 Trends in Heat Demand

Figure 2 plots yearly averages for the two main variables of interest in this paper: local heating degree days and kWh per sqm of heat energy consumed by buildings. We demonstrate that the correlation between average heat energy demand and annual sum total of heating degree days increased from 2003 - 2012 to 2008 - 2018.

It is further noteworthy that the aggregated time series from 2008 to 2018 exhibit near identical movements over time. Although the downward linear trends (see fitted values) are similar, heat

Figure 2: Decline in Heat Demand (lhs) and Heating Degree Days (rhs)



Notes: The graph plots simple averages of annual heat energy consumption for the full sample against the starting year of each billing or heating period (rolling 12-months). The left y-axis measures the annual kWh per square meter of heated living space per building, while the right y-axis measures the number of heating degree days – heating degrees recorded on days with temperatures below 15 Celsius, recorded at the nearest weather station to each building’s zip code. Source: authors’ calculations using data described in Section 3.

energy demand appears to be decreasing at a slightly faster rate than heating degree days.

Figure 2 suggests that (1) temperature plays an increasingly dominant role in determining heat energy demand, (2) during 2008 - 2018, renovations, demographic changes, and economic factors such as prices played a smaller role (at least on net) in predicting heat energy demand. This is strong suggestive evidence that policies targeting CO₂ emission reductions via technical efficiency of the building sector may not have had the desired effect for the sample population.

Moreover, thermal retrofitting of the existing housing stock is widely considered among German policymakers to be cost-effective in reducing heat energy consumption (Galvin and Sunikka-Blank 2013). But given the declining trend in heating degree days and thus heat demand, the pay-back period for thermal-efficiency investments continues to get longer, and we expect fewer (rather than more) renovation projects to be economically feasible due to climate change. In short, the potential for cost-efficient carbon savings from residential energy efficiency have been declining in the last decade. It remains an empirical question, however, how the currently falling economic incentives for a more energy-efficient housing stock will be altered in response to a carbon-pricing scheme on the residential heating sector.

4 Temperature Response of Heat Demand

The primary aim of this paper is use precise estimates of the effect of temperature fluctuations on heat energy consumption to understand the heat energy requirements of the existing building stock. As discussed already, the estimation sample period covers 8 of the 10 warmest years ever recorded in Germany's history, lending us high statistical power in estimating the response of energy demand to temperature.

4.1 Identification

We seek to isolate the underlying heat-energy requirements (a measure of energy efficiency) at the building level by estimating the direct response of building-level heat demand to temperature variability. The main advantage of the empirical approach is that we control for all time-invariant characteristics of buildings and other time-varying factors that could affect the response of heat de-

mand to temperature. The number of heating degree days are exogenous, conditional on building and time fixed effects, allowing us to identify causal estimates.

We use the following regression to estimate the average response of heat energy requirements to temperature:

$$y_{is} = \alpha + \beta \text{HDD}_{is} + \gamma_i + \phi_s + \lambda_{ft} + \epsilon_{is}$$

where y_{is} denotes annual energy units consumed (kWh) per sqm by building i with billing cycle starting in year s .

- HDD captures the numbers of heating degree days in any given 12-month billing period, calculated as the total sum of differences between the daily mean temperature and the heating threshold of 15°C on days with recorded mean temperatures less than 15°C. Succinctly,

$$\text{HDD} = \sum_{d=1}^{365/366} (T_H - t_d) \times \mathbb{1}[t_d < T_H]$$

We considered different heating thresholds of T_H from 10°C to 20°C and chose the heating threshold of 15°C because it minimizes Root Mean Squared Error (RMSE).

- γ_i captures the building-level fixed characteristics such as vintage, building type, number of apartments, number of floors, quality of building components associated with thermal insulation, aggregate characteristics of residents, ownership status, and fixed-factors related to geographic location.
- ϕ_s are fixed effects for the year (s) in which annual billing cycle starts. This controls for any aggregate time trend in consumer behavior or preferences for heating, influenced by gradually decreasing heating degree days, measured for each building's billing period.
- λ_{ft} are fixed effects for the fuel type (f) by majority billing year (t), which capture year-specific economic shocks that are common to all buildings by each fuel type. Note that $s \neq t$ for bills starting late in the second half of the year (August to December).

- ϵ_{is} the error term, clustered at the zip code level.

The coefficient of interest is β on HDD. β reports the change in annual kWh of heat energy demand per square meter in response to a unit increase in the annual sum total of heating degree days, after controlling for fixed factors associated with buildings, fuel type, residents, and geographic location.

4.2 Results

Table 2 presents results for the full estimation sample. We present four variations of the fixed effects model. We consistently estimate that heat energy consumed by a building per annum fell on average by approximately 3.1 kWh/m² of heated living space in response to a decrease of 100 degree days in the annual sum total of local heating degree days (with mean outside temperature below 15 degree Celsius).

In the first column, conditional on fixed characteristics of buildings and yearly shocks common to all homes in their respective billing periods, we arrive at a coefficient that is insensitive to specification. The second column presents the preferred specification, in which we add fixed effects for the majority billing year interacted with fuel type, which capture yearly fuel price shocks. In the third column, we add time-varying covariates at the building level (price per kWh of energy, building size categories) and zip code level (unemployment rate, purchasing power per household, and density of buildings). Finally, in the fourth column we control for any time-varying factors at the zip-code level that may affect building-level energy efficiency. We also considered other weather parameters and show that the average estimate is robust to the inclusion of precipitation and relative humidity (see Appendix Table B4).

The average annual heat energy requirement of buildings increases by approximately 0.031 kWh per sqm for each heating degree day. This corresponds to an average yearly energy requirement of 79.7 kWh per sqm of heated space. This estimate assumed linearity in temperature response and total heating degree days of 2414, which is the average 12-month sum of heating degree days recorded at 204 weather stations in Germany from 2003 to 2018⁶. Compared to the "typical"

⁶The values are calculated by the author using daily observations from 20030101 to 20181231 at 204 nearest weather stations to 8303 zipcodes in Germany. In the mapping used, average distance between zip code and nearest

climate, the decline in the annual sum of heating degree days measured in the last five years (2014 to 2018) translates this estimate to an average reduction of 8 percent in annual heat energy demand. For a household living in a 100 square meter apartment, expenses in 2018 for space heating fell by approximately 60 euros due to change in climate (HDD relative to "typical").

Table 2: Response of Heat Energy Demand to Temperature

<i>Dependent variable: kWh/m²</i>				
	(1)	(2)	(3)	(4)
Local HDD	0.0312*** (0.0002)	0.0306*** (0.0002)	0.0306*** (0.0003)	0.0303*** (0.0002)
Energy Price (Euros/kWh)			-17.20*** (2.737)	
Unemployment (%)			-0.078 (0.066)	
Purchasing Power per HH			0.147*** (0.022)	
Building Size			-0.134*** (0.048)	
Zip Code Density			-0.018*** (0.003)	
Building FE	Yes	Yes	Yes	Yes
Bill Start Year FE	Yes	Yes	Yes	Yes
Fuel Type FE	Yes			
Year FE	Yes			
Fuel Type x Year FE		Yes	Yes	Yes
Zip Code x Year FE				Yes
<i>N</i>	3,030,063	3,030,063	2,251,079	3,023,892
Adj <i>R</i> ²	0.776	0.776	0.790	0.780

Notes: The full sample covers all buildings using natural gas (high or low calorific), district heating, and oil. The dependent variable is kilowatt hours consumed per square meter of heated living space. Price of heat energy is given in Euros per kWh and observations above 1 Euro per kWh were removed. Standards errors in parentheses are clustered at the zip code level for all specifications. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

weather station is 18.3 km, with standard deviation of 10.4 km, minimum and maximum distance of 0.076 km and 59.86 km respectively. Source: German Weather Service (DWD)

4.3 Heat Demand Adjusted for HDD

The main disadvantage of the method outlined above is that we only consider the linear response of energy demand to changes in annual heating degree days. In the appendix, we considered annual sum of heating degree days in distinct temperature bins below the heating threshold of 15°C for each 12-month billing period. This allowed us to flexibly detect any non-linear response of building-level heat demand to changes in small versus large changes in local weather conditions (see Table B3). The results are largely consistent – for temperatures lower than the heating threshold, every 1 °C decrease in temperature in colder temperature bins require relatively more heat energy. The coefficient on "HDD if (°C < -5)" is significantly lower than what we expect. This is driven by the fact that observations from buildings with high energy-efficiency standards are over-represented in the coldest temperature bin.

Using this more precise specification, which controls for the non-linear and heterogeneous response of heat energy consumption to temperature variation, we show how heat demand from the building sample has evolved since 2008. Figure 3 shows that heat demand was declining until 2014, after which the trend is ambiguous, given the wide confidence intervals.

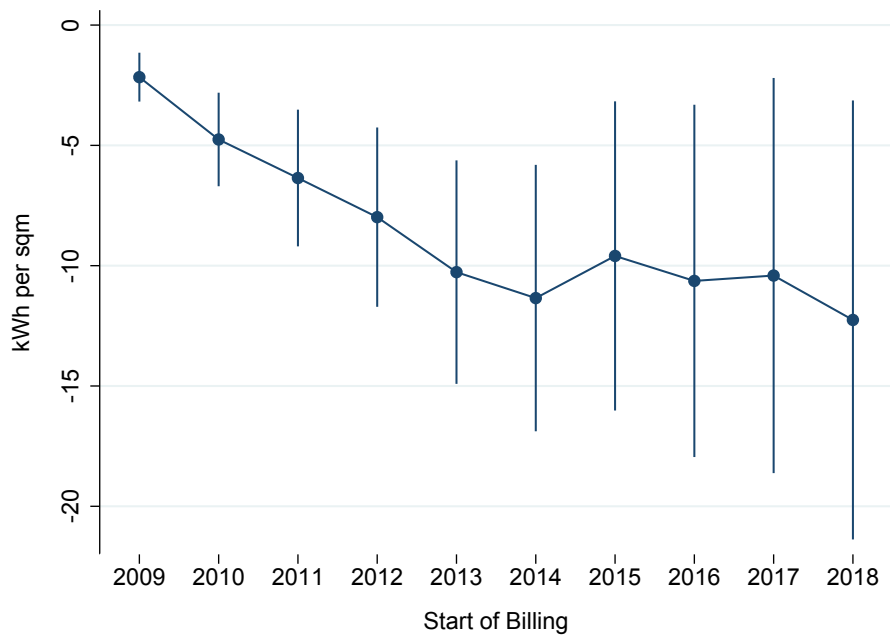
5 Effect Heterogeneity

Now we go beyond estimating the average marginal effects. We expect there to exist considerable heterogeneity in temperature response, not least because newer buildings were subject to stricter building energy codes. In this section, we explore which subgroups of buildings response less or more to temperature, and thus seek to uncover the heterogeneity in heat energy requirements of the existing building stock. We consider those variables that are most likely related to marginal effect heterogeneity: year of construction, size of building, the unemployment rate, and federal states with the West of Germany.

5.1 Building Vintage

To start, we consider energy-efficiency outcomes by building vintage grouped such that they correspond to different energy efficiency codes. Our results suggest that older buildings are less efficient

Figure 3: Decline in Heat Demand Relative to 2008



Notes: The graph illustrates how annual heat demand has changed relative to the base year of 2008. These are coefficient estimates drawn from a regression of heat demand on HDD, disaggregated by 5 temperature bins in the interval $(-\infty, 15)$, with fixed effects for building, fuel type, (majority) year of billing and fuel type by year of billing. 95% confidence intervals provided. Standards errors were clustered at the zip-code level.

than newer buildings in both East and West Germany, but the difference is more pronounced in the West. Energy efficiency regulation in Germany has largely taken the form of building codes, that define the building-aggregate maximum annual energy requirement per square meter of living space for newly constructed homes. Table 3 summarizes the time development of building codes in Germany. The Heat Insulation Ordinance was first introduced in 1978, amended and made progressively more stringent in 1984 and 1995. It was replaced by the Energy Saving Ordinance in 2002 and subsequently amended in 2009. Note that there were no minimum energy standards for buildings built before 1978.

Table 3: Standards for New Construction

Year	Regulation	Max. per annum
Pre-1978	No regulation	
1978	Heat insulation (WSchV)	250 kWh/m ²
1984	Amendment of WSchV	220 kWh/m ²
1995	Amendment of WSchV	150 kWh/m ²
2002	Energy saving (EnEV)	100 kWh/m ²
2009	Amendment of EnEV	60 kWh/m ²
2016	Amendment of EnEV	45 kWh/m ²

Notes: The first column indicates the year in which the regulation became effective.

Sources: Galvin and Sunikka-Blank (2013), El-Shagi et al. (2017)

In Table 4, we estimate the response of energy requirements to temperature separately for buildings in the West and East of Germany⁷. For the West, the average heating requirement per heating degree day is monotonically decreasing with each successive revision (tightening) of building codes, until 2009. For instance, the difference in the average temperature response between buildings built in 1984–1994 and those built in 1995–2001 is 32%. Compared to buildings built from 2009 onwards, buildings built before 1978 are on average less energy-efficient by a factor of 2.

This decreasing pattern is less clear, however, when we consider only the buildings located in the East of Germany. These results are a strong indication that other factors such as renovations targeting energy-efficiency of older buildings, other building attributes, or state-level policy efforts

⁷East is located in the following federal states: Berlin, Brandenburg, Mecklenburg-Vorpommern, Freistaat Sachsen, Sachsen-Anhalt, and Freistaat Thüringen.

could be important in explaining the differences in energy requirements of buildings in the East.

Comparing across the regions, the coefficients in the first column shows that buildings built before 1984 in the East perform comparatively well, while the energy-requirements are lower for buildings built from 1995 in the West. Interestingly, pre-1978 (before building regulation) buildings in the East are significantly older, but on average demand less heat energy per heating degree Celsius.

Table 4: Temperature Response By Year of Construction

West Germany						
Response of kWh/m ²						
	(1)	(2)	(3)	(4)	(5)	(6)
	Pre-1978	1978-1983	1984-1994	1995-2001	2002-2008	2009+
Local HDD	0.0367*** (0.0004)	0.0309*** (0.0006)	0.0248*** (0.0004)	0.0169*** (0.0004)	0.0109*** (0.0005)	0.0176*** (0.0018)
$N = 1,045,616$						
Adj R ² = 0.7706						
Share	0.50	0.10	0.22	0.14	0.04	0.001
Avg. Vintage	1952	1981	1990	1997	2004	2010
East Germany						
Response of kWh/m ²						
	(1)	(2)	(3)	(4)	(5)	(6)
	Pre-1978	1978-1983	1984-1994	1995-2001	2002-2008	2009+
Local HDD	0.0300*** (0.0001)	0.0237*** (0.0013)	0.0238*** (0.0011)	0.0207*** (0.0001)	0.0171*** (0.0016)	0.0303*** (0.0077)
$N = 317,401$						
Adj R ² = 0.8301						
Share	0.66	0.04	0.10	0.18	0.01	0.001
Avg. Vintage	1929	1980	1990	1997	2004	2010

Notes: Table presents coefficients on the interactions between Local HDD and building codes. The specification includes fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. The sample covers all buildings using natural gas (high or low calorific), district heating, and oil. The dependent variable is kilowatt hours consumed per square meter of heated living space. Standards errors in parentheses are clustered at the zip code level for all specifications. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5.2 Building Size

Now we explore how the energy requirements of buildings differ by the size of buildings, again analyzing buildings in the West and East of Germany separately. The results in Table 5 demonstrate that for both regions, the response of annual kWh/sqm heat demand to an additional heating degree day is falling with building size. Specifically, the largest subgroup of buildings in the West (with more than 20 apartments or households) demand on average 26 percent less annual kWh per sqm in response to each heating degree day compared to two-family homes. Interestingly, relative to the West, buildings located in the East have on average better energy efficiency standards within each building size category. Taken all together, these results strongly suggest that larger buildings are significantly more energy-efficient.

5.3 By Unemployment Rates

Is the distribution of energy-efficiency outcomes inequitable? An important limitation of our data is that we do not observe the socio-economic characteristics of households or the ownership status of the buildings. For instance, the income and wealth of the owner of the building or its residents may play an economically significant role in the incentives to improve the energy-efficiency of houses.

Instead we use data available at the zip code level and make the plausible assumption that the location of buildings is strongly correlated with the socio-economic background of residents, i.e. poorer zip codes tend to accommodate poorer households. Here we consider effect heterogeneity by unemployment rates, which is an important dimension of socio-economic disparities across regions.

We pay particular attention to the spatial distribution of unemployment rates in this paper. because (1) we study primarily multi-apartment buildings and (2) according to official statistics, the large majority (over 80 percent) of the unemployed population resides in multi-apartments buildings (Destatis 2018). Specifically, we seek to answer whether less energy-efficient houses are located in localities or communities marked with higher unemployment rates. In other words, do the least economically-secure households live in buildings that do not perform as well in terms of energy-

Table 5: Temperature Response By Size of Building (# of Apartments)

West Germany					
Response of kWh/m ²					
	(1)	(2)	(3)	(4)	(5)
	2 HH	3 - 6 HH	7 - 12 HH	13 - 20 HH	21+ HH
Local HDD	0.0354*** (0.0003)	0.0314*** (0.0003)	0.0285*** (0.0003)	0.0271*** (0.0004)	0.0261*** (0.0004)
<i>N</i> = 2,460,694					
Adj R ² = 0.9667					
Share	0.15	0.51	0.21	0.07	0.07
Avg. Building Size	2 Units	4.22	8.99	15.98	42.97
East Germany					
Response of kWh/m ²					
	(1)	(2)	(3)	(4)	(5)
	2 HH	3 - 6 HH	7 - 12 HH	13 - 20 HH	21+ HH
Local HDD	0.0325*** (0.0011)	0.0287*** (0.0007)	0.0267*** (0.0007)	0.0256*** (0.0007)	0.0243*** (0.0007)
<i>N</i> = 569,369					
Adj R ² = 0.8184					
Share	0.04	0.31	0.27	0.13	0.25
Avg. Building Size	2 Units	4.64	9.23	16.23	45.57

Notes: Table presents coefficients on the interactions between Local HDD and building size categories. The specification includes fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. The sample covers all buildings using natural gas (high or low calorific), district heating, and oil. The dependent variable is kilowatt hours consumed per square meter of heated living space. Standards errors in parentheses are clustered at the zip code level for all specifications. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

efficiency?

Results in Table 6 paint an interesting picture. The majority of the buildings in the East of Germany fall into the top tercile of the unemployment rate distribution, with average unemployment rate of 9.67%. However, we estimate the lowest average heat demand response per heating degree day for this subgroup of buildings. In contrast, households in the West of Germany living in buildings that are more energy-efficient are located in the zip codes that fall in the lowest tercile of the unemployment rate distribution. The conditional average marginal effects estimated for the West of Germany also highlight potentially large economic inequity in the distribution of energy-efficiency outcomes, compared to the localities in the East.

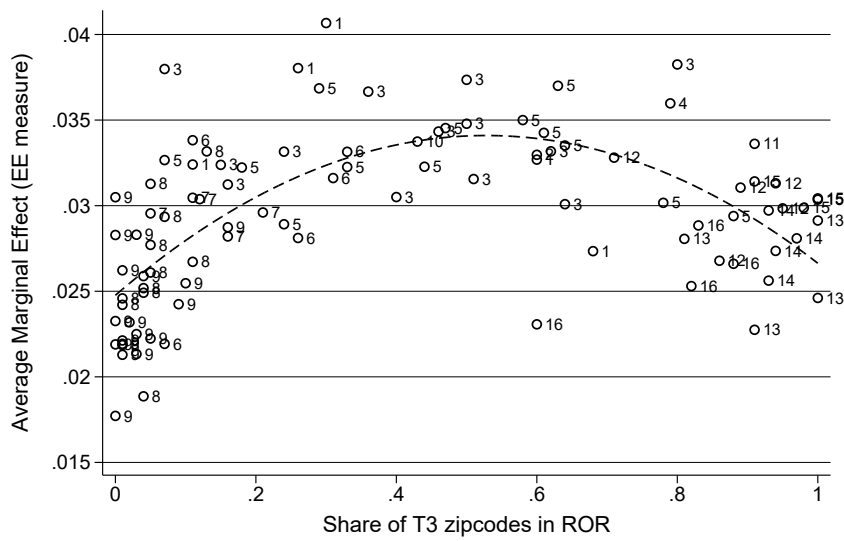
We explore these regional disparities in more detail in Figure 4. The labels indicate the corresponding federal state in Germany. Overall, it shows that planning regions with the lowest shares of zip codes with high unemployment rates have some of the most energy-efficient buildings. On the flip side, buildings in planning regions with a large majority of zip codes experiencing high unemployment rates are reasonably energy-efficient. A cursory look at the labels reveal that regions in the South states perform the best, followed by regions in the East, and in the middle of the graph lie regions that fall in the North-West. We explore differences in energy-efficiency outcomes at the state-level in the next section.

Table 6: Temperature Response By Terciles of the Unemployment Rate

West Germany			
Response of kWh/m ²			
	(1)	(2)	(3)
	T1	T2	T3
Local HDD	0.0279*** (0.0004)	0.0306*** (0.0003)	0.0332*** (0.0003)
<i>N</i> = 2,107,369			
Adj R ² = 0.7737			
Share	0.25	0.39	0.36
Avg. Unemployment %	2.26	4.22	8.65
East Germany			
Response of kWh/m ²			
	T1	T2	T3
Local HDD	0.0309*** (0.0020)	0.0275*** (0.0001)	0.0270*** (0.0006)
<i>N</i> = 569,369			
Adj R ² = 0.8184			
Share	0.01	0.09	0.90
Avg. Unemployment %	2.14	4.46	9.67

Notes: Table presents coefficients on the interactions between Local HDD and indicators for terciles of the unemployment rate distribution. The specification includes fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. The sample was limited to years 2009 to 2017, for which the unemployment rates were available. The last row reports the average unemployment rate for each tercile. The sample covers all buildings using natural gas (high or low calorific), district heating, and oil. The dependent variable is kilowatt hours consumed per square meter of heated living space. Standard errors in parentheses are clustered at the zip code level for all specifications. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 4: Regional Distribution of Energy Efficiency



Notes: The graph shows how energy efficiency varies with socio-economic status of the regions (measured by the unemployment rate). The vertical axis plots the average marginal effect for each of the 96 spatial planning region (ROR or Raumordnungsregion) in Germany. The labels indicate the corresponding federal state. These are coefficients on the interaction term between spatial planning regions and heating degree days. The regression included fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. The sample covers all buildings using natural gas (high or low calorific), district heating, and oil. Standards errors were clustered at the zip code level and all coefficients were statistically significant at the 1% level. The horizontal axis measures the share of postal codes in each region that fall in the top tercile (T3) of the national unemployment distribution.

5.4 Regional Energy Efficiency

In addition to the East-West divide in Germany, we further expect that differences between states in the West could arise for a number of reasons. These may include differences in the availability and take-up of financing initiatives for retrofits, state level mandates such as the 2015 Renewable Heat Act (EWärmeG) in Baden Württemberg that was intended to increase the share of renewable energy in heat supply, or even greener preferences at the household level.

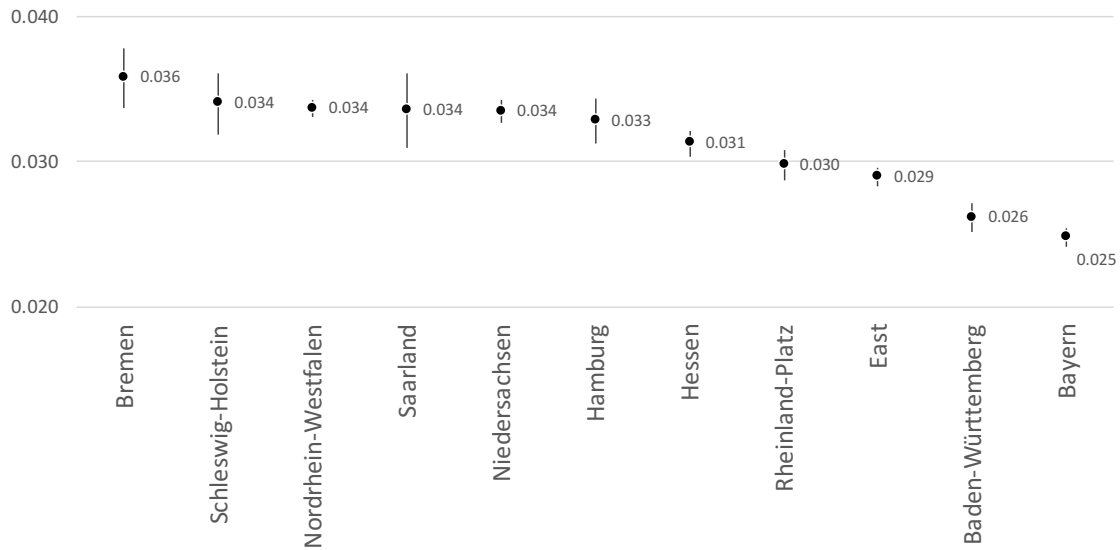
In this section, we show that states in the East are indeed endowed with more energy-efficient buildings compared to almost all buildings located in the West of Germany. However, there are

two large and thus important exceptions: the states of Bavaria and Baden Württemberg.

Figure 5 plots the average marginal effects for buildings located in each of the states in West Germany and East defined as one bloc. The differences are ordered in magnitude and associated 95% confidence intervals are shown. We highlight that buildings observed in Baden Württemberg and Bavaria are more energy-efficient compared to the East – demanding less heat energy per square meter on average for each heating degree day.

This is in contrast to the conclusion one would draw from energy performance scores from consumption based EPCs. Scores on energy-performance certificates are unable to purge the effect of factors such as consumer behaviour, energy prices, and socio-economic factor on the estimated energy performance of buildings. Indeed, households residing in East of Germany are significantly poorer than those in the South (Bavaria and Baden Württemberg). Low energy scores on performance certificates, thus, indicate both that buildings in the East are relatively energy-efficient buildings and the fact that energy consumption is dampened due to low incomes. Indeed, energy performance scores would lead us to conclude that buildings in the East are more energy-efficient than those in the rich states of the South. Using our methodological improvement over the use of energy performance certificates, we are closer to estimating the true potential of buildings – buildings in the South are in fact performing better than those in the East once we account for fixed differences.

Figure 5: Energy Efficiency by States



Notes: The graph presents the differences in the average marginal effect for each state in West Germany and states in the East pooled in one group. Corresponding 95 % confidence intervals are shown. The estimates are coefficients on the interaction between HDD and regions in a regression with fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. The sample covers all buildings using natural gas (high or low calorific), district heating, and oil. Standards errors were clustered at the zip code level.

6 Machine Learning - Causal Forests

The causal forests approach proposed by [Wager and Athey \(2018\)](#) makes it possible to estimate building-specific marginal effects, providing a much more detailed picture of the full extent of heterogeneity in energy-efficiency outcomes.

Causal forests are an extension of the widely used random forest machine learning algorithm described by [Breiman \(2001\)](#) that provide consistent treatment effect estimates conditional on covariates at the individual level. In simplified terms, the method works by splitting the dataset into parts, one of which is used to determine a set of subgroups for estimating heterogeneous effects and the other of which is used to estimate the effects.

Causal forests can fit very flexible nonlinear functions and uncover nonlinear interactions between

variables, generating individual marginal effect estimates for each observation. As discussed above, in this case, the marginal effect measures the impact of heating degree days on energy consumed per square meter.

We apply causal forests to examine heterogeneous effects over the following variables: building size, heating fuel type, building age, heating system age, zip-code level unemployment rates and household purchasing power, state, and regional planning regions (ROR).

In addition, before applying the causal forest, we apply building and year fixed effects by demeaning all continuous variables (including the dependent variable) at the building and year level.

Intuitively, our approach allows the treatment effect to vary in nonlinear ways over each of the included variables, as well as over interactions of the variables. We therefore obtain building-specific estimates of the relationship between heating degree days and heat energy consumption per square meter as a function of building attributes. The results we share in this section explore how those estimates vary over a single variable holding others constant at the sample mean.

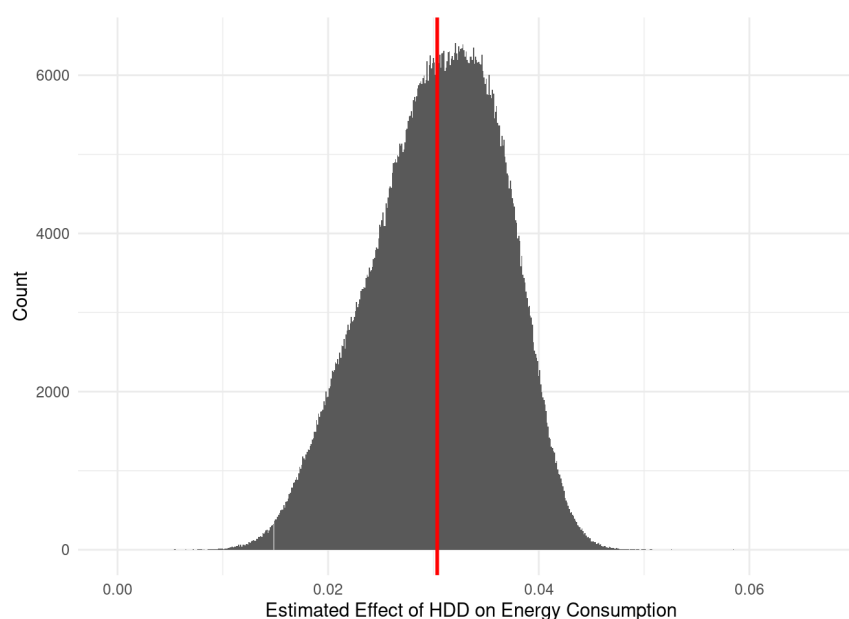
Figure 6 is a histogram showing the full distribution of estimated marginal effects. The average marginal effect is about 0.03, which is similar to the results from the fixed-effects regressions reported above. The distribution is slightly negatively skewed; there are more outliers on the efficient (non-responsive to temperature) end of the spectrum than at the other extreme, perhaps because building standards and retrofits have been effective in ensuring a minimum level of energy efficiency. Figure C3 shows the geographical distribution of estimated marginal effects.

We contrast our findings with those coming from the "traditional" approach – energy ratings on energy performance certificates. We show how building-specific marginal effects correspond to the reported energy scores in the data on energy performance certificates in Figure C4 – which suggests that buildings with lower energy scores (higher energy-efficiency standards as measures on the EPC) are more likely to underperform on energy performance predictions⁸. This is con-

⁸Although there is a positive correlation in the data, there are many cases with high marginal effects and low energy scores. This suggests that our methodology may be catching large departures for the best case energy performance settings/buildings. A definite interpretation of this graph is however difficult because we are unable to distinguish the type of EPC reported in the data, i.e. whether the energy score is based on engineering estimates (Demand Certificate) or past energy consumption on inhabitants (Usage/Consumption Certificate).

sistent with rigorous studies from the field that use actual metered consumption data to show that realized energy efficiency of homes are not as high as those predicted ex-ante by engineering models. Based on the results in this paper we argue that despite the best of intentions to provide transparent and reliable information on the energy quality of residential buildings, scores on energy performance certificates are not without their shortcomings.

Figure 6: Distribution of Marginal Effects

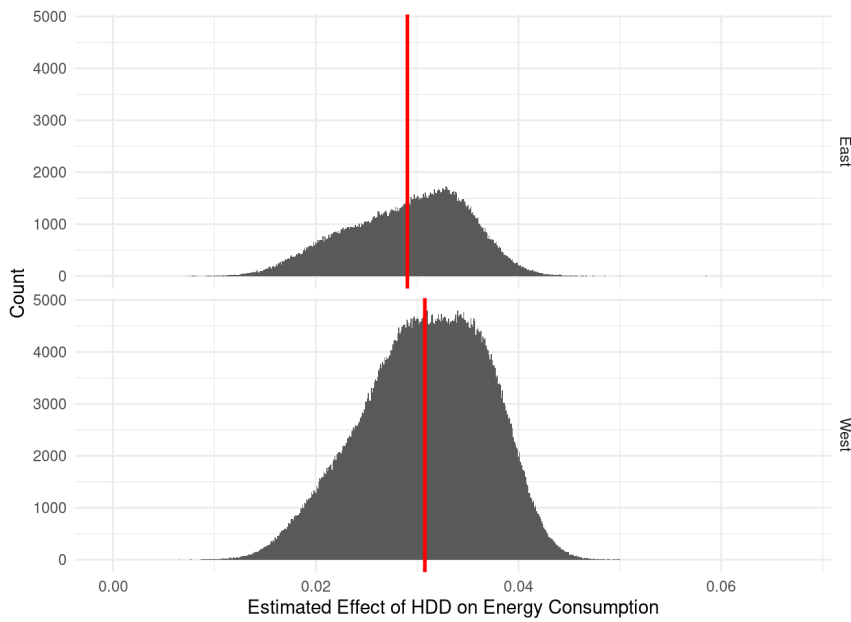


Notes: The graph is a histogram showing the full distribution of building-specific marginal effects, generated using causal forests. In this case, the marginal effect measures the impact of heating degree days on energy consumed per square meter.

Examining East and West Germany separately yields an even more interesting picture, as shown in Figure 7. The negative skewness is particularly pronounced in the East and much less noticeable in the West, perhaps because of intensive retrofit programs in Eastern states. By further breaking down the distributions by vintage in Figure 8, we can see that an important driver of the greater average efficiency in East Germany is greater efficiency of the old buildings that were built before 1978 – retrofit programs likely explain the significant mass on the left side of the distribution. This East-West divide is likely magnified because a much higher (lower) share of the the building stock in the East is heated via district heating (heating oil)⁹.

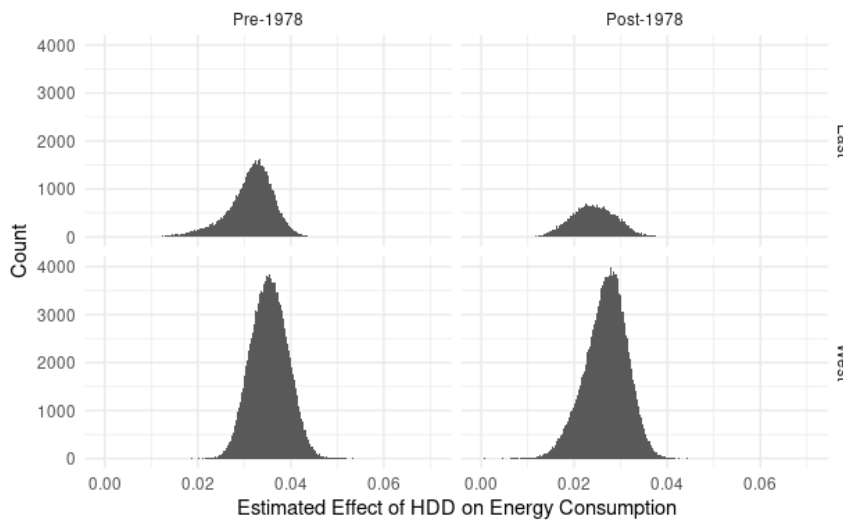
⁹Figure C5 shows the distribution of marginal effects by fuel type, in particular we show that buildings using

Figure 7: Distribution by East and West Germany



Notes: The graph shows the full distribution of building-specific marginal effects for East and West Germany generated using causal forests.

Figure 8: New Versus Old Buildings: East & West



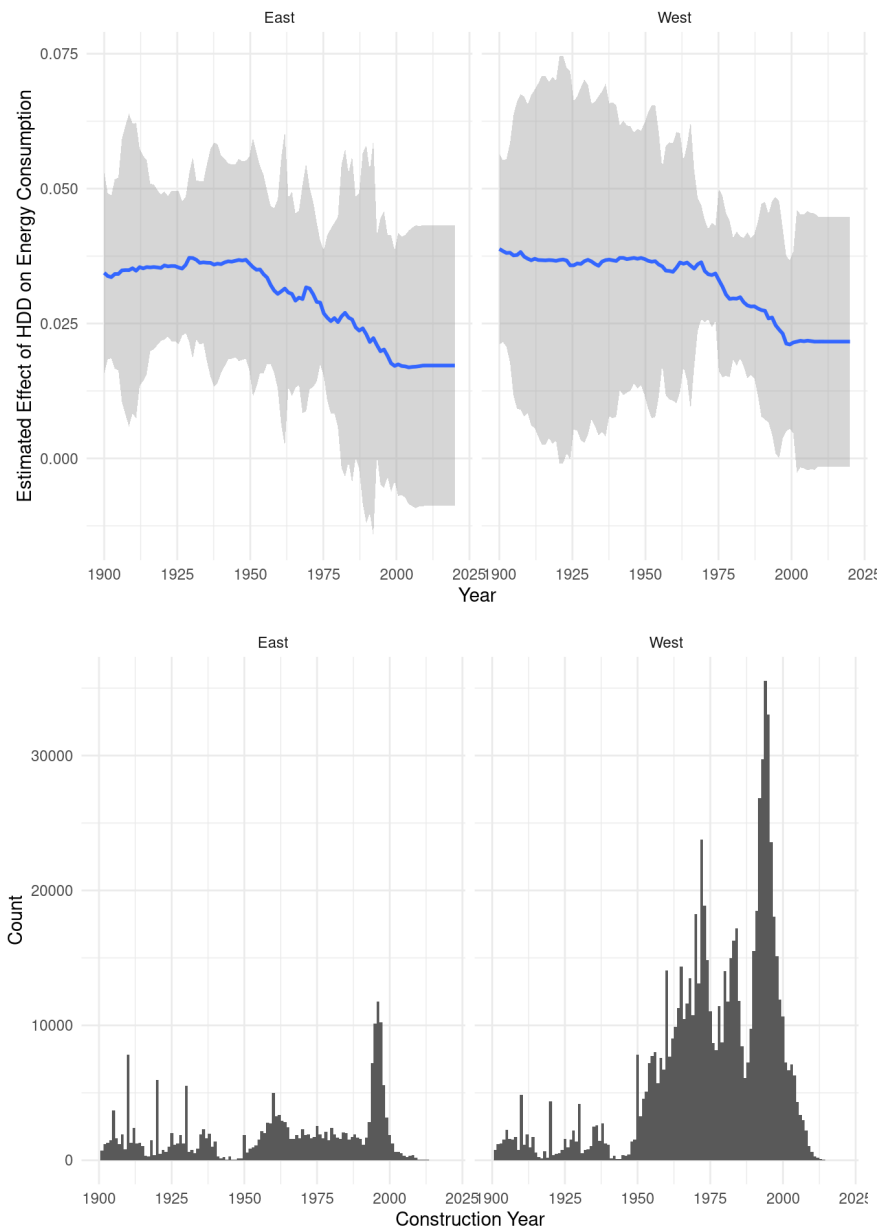
Notes: The graph shows how the distribution of building-specific marginal effects varies for East and West, separately for buildings before and after 1978, when energy standards became effective.

We can also examine how building efficiency varies based on the exact year when they were built, and the results provide further support for the hypothesis that retrofit programs in East Germany heating oil are relatively less energy-efficient.

led to a more egalitarian distribution of energy efficiency. Given that the retrofit rate of existing buildings has been abysmally low over the last two decades ([Galvin and Sunikka-Blank, 2013](#)), the main candidate to have improved the energy-performance of buildings are building codes, that have been rising in stringency since 1978, when they were first introduced. We do not quantify the causal link in this paper, but using precise measures we provide evidence on the strong negative association between realized heat energy-requirements of buildings and the stringency of federally-enforced building codes in Germany.

As shown in [Figure 9](#), newer buildings are more energy efficient. In particular, efficiency seems to start increasing around 1975 in the West and then increases dramatically again in the 1990s. But more interestingly, significant efficiency improvements are also evident for buildings built between 1950 to 1975 in the East, before national building codes came into effect. This can help us further understand the East/West distributions we studied in [Figure 7](#). We also show the age distribution of the building stock by region.

Figure 9: Marginal Effects by Construction Year

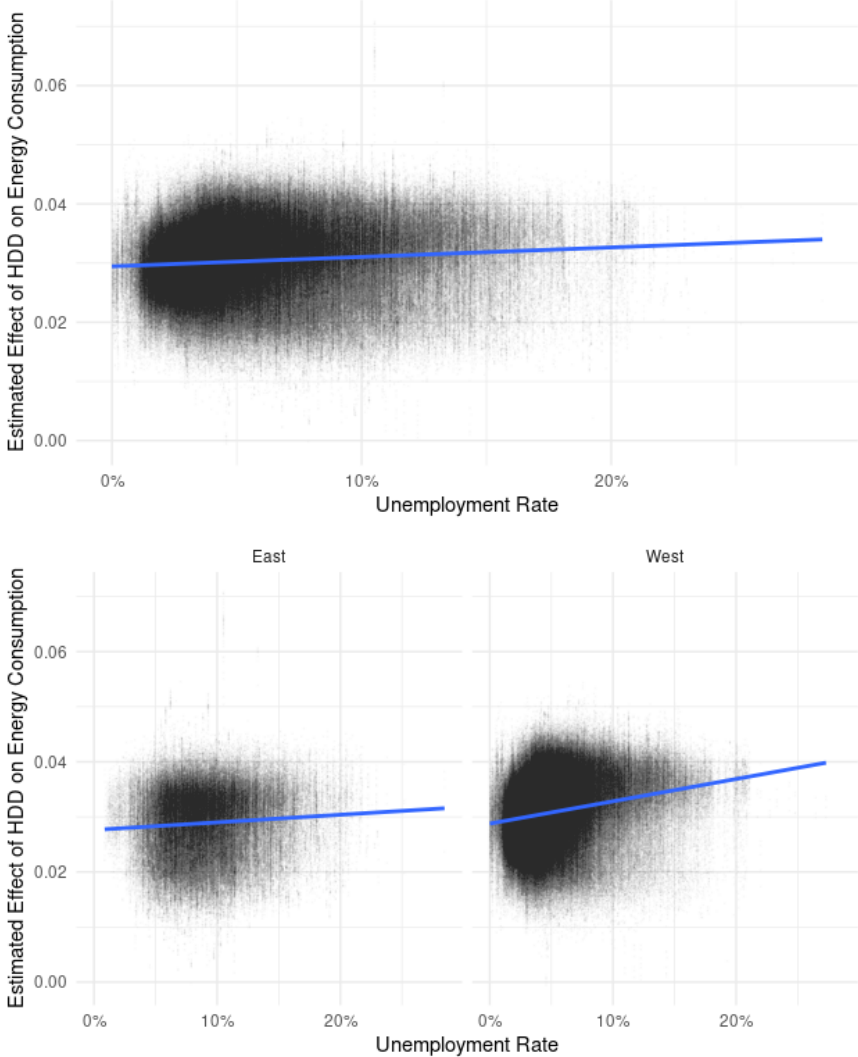


Notes: The first graph shows how the marginal effects generated using causal forests vary over construction year, holding all other variables constant at the sample mean, separately for East and West Germany. Grey bands indicate the 95% confidence intervals. The second graph in the panel plots the frequency of buildings observed by construction year.

Finally, we can take a more direct look at how estimated effects vary with regional economic conditions by plotting estimates against unemployment rates in each building’s postal code. The first panel in Figure 10 shows that zip codes with higher unemployment rates appear to have less efficient buildings on average. As shown in the second panel, this effect is much more pronounced

in West Germany than in East Germany, again likely due to East German retrofit efforts in the 1990s. This reinforces the point made above: East German retrofit efforts appear to have yielded substantial improvements in energy efficiency for those living in less prosperous zip codes.

Figure 10: Marginal Effects By Unemployment Rates

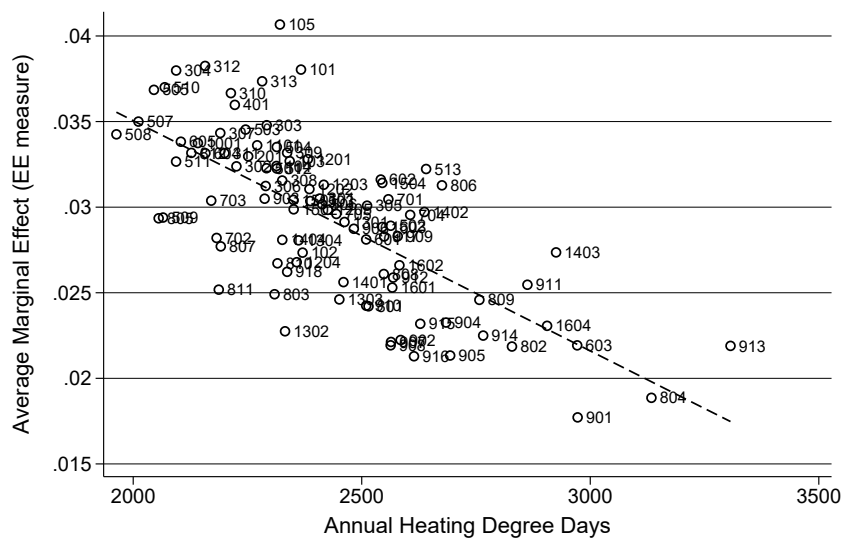


Notes: The graphs shows how ML-based building-specific marginal effects vary with the corresponding zipcode’s average unemployment rate. The second panel shows the correlations separately for East and West Germany. Regression lines in blue shows that areas with higher unemployment have somewhat lower estimated efficiency on average and this relationship is stronger in the West than in the East.

7 Explaining the Distribution of Energy Efficiency

What may explain the seemingly higher number of thermal retrofit measures in the East while the regions in the North-West lagging behind? We explore further the distribution of heating degree days across these regions and find that the spatial distribution of energy-efficiency is driven by heating needs. Figure 11 shows a strong negative relationship between the energy performance of the average building in each region and recorded heating degree days from 2008 to 2018. These results suggest that energy efficiency of buildings is likely not a low-hanging fruit and that lower expected returns to investments explain why low-income zip codes in the West do not catch up. In other words, the regional inequality in energy-efficiency of buildings are likely optimal outcomes. Thus the promotion of energy-efficiency in regions lagging behind may be a cost-inefficient policy, especially given warmer heating seasons.

Figure 11: Heating Needs and Energy Efficiency Outcomes



Notes: The graph shows how energy efficiency varies with regional heating needs, measured using recorded heating degree days from 2008 to 2018. The vertical axis plots the average marginal effect for each of the 96 spatial planning region (ROR or Raumordnungsregion) in Germany. The labels indicate the spatial planning region number. These are coefficients on the interaction term between spatial planning regions and heating degree days. The regression included fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. The sample covers all buildings using natural gas (high or low calorific), district heating, and oil. Standards errors were clustered at the zip code level and all coefficients were statistically significant at the 1% level.

8 Conclusion

By combining over 3 million yearly heating bills with daily temperature data at local weather stations, we examine the long-term realization of energy-efficiency outcomes in the housing sector. By examining the distribution of outcomes we find large regional differences. In particular, results suggest that early investments in retrofitting has had a lasting impact in the East of Germany – which is largely missing for the poorer zip codes in the West of Germany. A closer look reveals that although the distribution of energy-efficiency outcomes is not equitable, it is not necessarily cost inefficient. The prevailing distribution of energy efficiency conforms to the distribution of heating degree days. A targeted EE policy may aim to close the inequality in energy performance of buildings between the regions. However, this may come at a real cost in terms of efficiency. A gradual fall in heating degree days due to global warming may be undermining economic incentives for energy-efficiency in the housing sector. We find strong evidence that heat demand declined up until 2014, after which gains in energy-efficiency were not statistically significant.

We also considered the non-linear response to temperature shocks in the last decade, in addition to controlling for changes in precipitation and wind conditions. Thus far, we have not examined how effect heterogeneity (by construction year or size of building, for example) varies with heating degree days in different temperature bins. It remains a very interesting extension to the paper, given that the warming climate brings milder winters, leading to the shift of heating degree days from colder to relatively warmer temperature intervals.

The main shortcoming of the paper stems from data limitations. We did not have access to high-frequency data on energy consumption and crucial information on the socio-economic characteristics of households that reside in the buildings observed. Barring these caveats, we hope to show in this paper that annual billing data is nevertheless useful to get an accurate picture of the underlying energy-efficiency performance of buildings. The main strength of the empirical approach in this paper vis-a-vis ratings on energy performance certificates is that by accounting for fixed differences (heating preferences and incomes of households, for example) between buildings we are closer to approximating the true energy performance potential of homes. These approximations may in turn be more reliable than energy performance scores on energy performance certificates,

allowing engineers to identify and target buildings that would benefit the most from renovations.

References

- AGEB (2018). Anwendungsbilanzen für die Endenergiesektoren in Deutschland in den Jahren 2013 bis 2017. Report, AG Energiebilanzen e.V. (available online, in German).
- Bento, A. M. (2013). Equity impacts of environmental policy. *Annual Review Resource Economics*, 5(1):181–196.
- BMWi (2019). Zahlen und Fakten, Energiedaten. Report, Federal Ministry for Economic Affairs and Energy. (available online, in German).
- Borenstein, S. and Davis, L. W. (2016). The Distributional Effects of U.S. Clean Energy Tax Credits. *NBER Tax Policy and the Economy, University of Chicago Press*, 30(1):191–234.
- Breidenbach, P. and Eilers, L. (2018). RWI-GEO-GRID: Socio-economic data on grid level. *Journal of Economics and Statistics*, 238(6):609–616.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.
- Bruegge, C., Deryugina, T., and Myers, E. (2019). The Distributional Effects of Building Energy Codes. *Journal of the Association of Environmental and Resource Economists*, 6(2):95–127.
- Christensen, P., Francisco, P., Myers, E., and Souza, M. (2020). Decomposing the Wedge Between Projected and Realized Returns in Energy Efficiency Programs. *E2e Working Paper*, 046.
- Destatis (2012). Bauen und Wohnen. Mikrozensus - Zusatzerhebung 2010. Report, Statistisches Bundesamt, Wiesbaden. (available online, in German).
- Destatis (2016). Bauen und Wohnen. Mikrozensus - Zusatzerhebung 2014. Report, Statistisches Bundesamt, Wiesbaden. (available online, in German).
- Destatis (2018). Wirtschaftsrechnungen: Einkommens- und Verbrauchsstichprobe Wohnverhältnisse privater Haushalte. Fachserie 15 Sonderheft 1, Statistisches Bundesamt. (available online, in German).
- Destatis (2019). Wohnen in Deutschland - Zusatzprogramm des Mikrozensus 2018. Report, Statistisches Bundesamt, Wiesbaden. (available online, in German).

- El-Shagi, M., Michelsen, C., and Rosenschon, S. (2017). Empirics on the Long-Run Effects of Building Energy Codes in the Housing Market. *Land Economics*, 93(4):585–607.
- Fowlie, M., Greenstone, M., and Wolfram, C. (2018). Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program. *The Quarterly Journal of Economics*, 133(3):1597–1644.
- Galvin, R. and Sunikka-Blank, M. (2013). *A Critical Appraisal of Germany's Thermal Retrofit Policy: Turning Down the Heat*. London: Springer.
- Giraudet, L.-G., Houde, S., and Maher, J. (2018). Moral hazard and the Energy Efficiency Gap: Theory and Evidence. *Journal of the Association of Environmental and Resource Economists*, 5(4):755–790.
- Jacobsen, G. D. (2019). An examination of how energy efficiency incentives are distributed across income groups. *The Energy Journal*, 40(6):171–198.
- Kotchen, M. J. (2017). Longer-run evidence on whether building energy codes reduce residential energy consumption. *Journal of the Association of Environmental and Resource Economists*, 4(1):135–153.
- McCoy, D. and Kotsch, R. (2020). Quantifying the distributional impact of energy efficiency measures. *GRI Working Paper*, 306.
- RWI and microm (2020). RWI-GEO-GRID: Socio-economic data on grid level (wave 9). Version: 1. RWI – Leibniz Institute for Economic Research. Dataset. <http://doi.org/10.7807/microm:V8>.
- Singhal, P. and Stede, J. (2019). Heat monitor 2018: Rising heating energy demand, thermal retrofit rate must increase. *DIW Weekly Report*, Volume 9, Issue 35/36, 303–312.
- Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.
- Wei, D. and Michelsen, C. (2011). The improvement of housing conditions in post communist germany - market mechanisms and subsidy impacts. Conference paper. 51st Congress of the ERSA: 30 August - 3 September 2011, Barcelona, Spain.

A Appendix

Institutional Setting

Here we provide some context for the analysis – we describe the residential building sector in Germany and how energy ratings on energy performance certificates are calculated.

A.1 National Context

Space heating accounted for approximately 70% of the final energy consumed by the private housing sector in 2017, and residential heating is almost one-fifth of Germany’s total final energy consumption (AGEB 2018 and BMWi 2019).

The German Federal Statistical office conducts a special microcensus¹⁰ survey every four years to report on the current housing situation in Germany. Based on the 2014 microcensus results (Destatis 2016), 50% of the rented German housing units were using natural gas for home heating, 22% were using district heating, another 19% using heating oil, and only about 2% using electricity. These figures were approximately the same in the 2010 and 2018 microcensus reports (Destatis 2012 and Destatis 2019). In this paper, we study buildings that were using natural gas (high calorific or low calorific), district heating, and mineral oil to heat their homes from 2008 to 2018. This amounts to excluding those buildings using LPG, pellets, electricity, wood, coal, brown coal, steam, and coke as main heat fuel energy.

Table A1 shows the distribution of buildings represented in the sample by fuel type and number of housing units. Compared to the 2014 microcensus survey results, district heating as the main fuel type is under-represented in the sample, while natural gas and heating oil are both over-represented. In the billing data sample, we observe all residential building types except those with only one housing unit, i.e. single-family homes. The sample of buildings is indeed not representative of the full population of the building stock in Germany – large housing blocks (with 13 or more apartments) are significantly overrepresented. The sample has a higher share of buildings with 3 to 6 and 7 to 12 apartment units. Nevertheless, analysis based on the building

¹⁰The microcensus surveys a representative one percent sample of the population.

sample in this paper is informative because more than two-thirds of German households live in buildings with 2 or more homes or apartments (Destatis 2019).

Table A2 shows the regional distribution of the full sample. The sample’s regional coverage mirrors the distribution of the population of buildings (Destatis 2016). In total, we observe buildings from 7769 zip codes, in all sixteen federal states in Germany.

Figure A1 shows the distribution of main fuel types used by buildings in the sample by terciles of the unemployment rate distribution. Tables B1 and B2 compare the building sample with statistics at the planning regional level (Raumordnungsregion or ROR) from the 2010 Microcensus for owner-occupied and rented buildings with 2 or more housing units. We show some evidence that distribution by fuel type is likely representative for the population of multi-apartment buildings ¹¹.

Table A1: Buildings by Fuel Type and Number of Apartments

Fuel Type	Building Size by Number of Apts					
	All	2	3 - 6	7 - 12	13 - 20	21 +
Natural Gas H	60%	6.4	28.6	15.1	5.0	4.6
Natural Gas L	1%	0.1	0.4	0.4	0.2	0.2
District Heating	10%	0.2	1.8	2.4	1.6	4.1
Oil	29%	6.4	16.3	4.1	1.1	1.1
Total	100%	13%	47%	22%	8%	10%

Notes: This table show the distribution of the 3,030,063 sample of heating bills by fuel type and building size. The sample covers buildings that were billed starting from January 2008 to June 2018.

A.2 Limitations of Energy Performance Certificates

In this section, we explain why energy performance scores that use past consumption to measure the energy-efficiency rating of buildings are imperfect measures of energy-efficiency. We start with a brief description of how energy performance scores based on consumption (as opposed to

¹¹Since the microcensus is conducted at the household level, the data describes the distribution of apartments and not the building stock. Thus, we can only compare the data sample (at the building level) to the Microcensus statistics (at the apartment level) in a very general and rough manner.

Table A2: Regional Shares

State	Observations	Percent	# of Zip Codes	Buildings
Schleswig-Holstein	56,792	1.9	361	7,245
Hamburg	56,900	1.9	97	8,101
Niedersachsen	291,197	9.6	782	36,623
Bremen	27,147	0.9	34	3,314
Nordrhein-Westfalen	769,198	25.4	864	100,811
Hessen	303,301	10.0	535	38,287
Rheinland-Pfalz	127,026	4.2	627	15,580
Baden-Württemberg	379,418	12.5	1146	46,072
Bayern	427,125	14.1	1864	52,366
Saarland	22,590	0.8	68	2,799
Berlin	117,986	3.9	189	15,538
Brandenburg	63,946	2.1	210	7,913
Mecklenburg-Vorpommern	61,771	2.0	186	7,932
Sachsen	194,499	6.4	388	24,230
Sachsen-Anhalt	80,564	2.7	207	10,712
Thüringen	50,603	1.7	211	6,404
Total	3,030,063	100	7,769	384,223

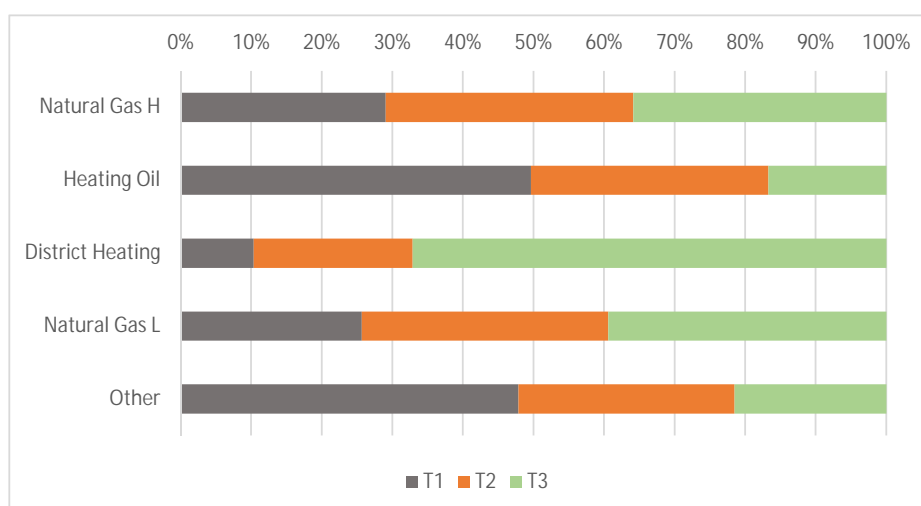
Notes: This table shows the number of heating bills observed in the full sample for each of the 16 states.

theoretical energy-requirements using engineering calculations) are calculated:

In accordance with the official guidelines¹², energy performance ratings on consumption-based ("Verbrauchsausweis") energy performance certificates (EPC) are calculated using building-level consumption in the past three years (in kWh per square meter of usable living space) and a climate factor to adjust for weather and climate differences. The climate factor anchors "typical climate" to the reference location of Potsdam, taking the average of annual heating degree days recorded from 1995 to 2012. Thus, in any given year if the local weather is warmer (colder) than "typical climate" in Potsdam, then consumption is adjusted upwards (downwards) accordingly by multiplying the annual consumption in kWh per square meter by the ratio of the annual sum total of heating degree

¹²Bekanntmachung der Regeln für Energieverbrauchswerte im Wohngebäudebestand Vom 7. April 2015

Figure A1: Fuel Type by Terciles of the Unemployment Rate



Notes: This figure shows the distribution of main fuel types used by buildings in the sample by terciles of the unemployment rate distribution, defined using data from RWI-GEO-GRID.

days recorded in Potsdam and at the local weather station. This allows buildings nation-wide to be compared to each other in terms of energy performance.

Using this method, however, does not give a transparent rating of the heat energy requirements of a building because the performance scores are not independent of the behaviour of tenants vs. residents, energy prices by fuel type, and other demographic and socio-economic factors. More importantly, correcting consumption for temperature and climate change to determine the "true" trend in heat demand may require much more than multiplicative climate factors. It requires us to consider heterogeneous response to changes in temperature that vary by regions, income, tastes, building attributes, and other energy-efficiency parameters. In comparison, the energy demand certificate ("Bedarfsausweis") is a more objective measure because it is the theoretical energy-requirement of a building.

In absence of sufficient data on energy-demand certificates for buildings, it is useful to consider how consumption responds to temperature as an improvement over consumption-based energy performance scores on EPCs. As we will show in the results sections, the strong co-movement between heat energy demand and heating degree days allows us to precisely estimate the energy performance of the existing building stock in a fixed-effects framework and detect even very small differences in energy efficiency between buildings and regions using machine learning.

B Appendix

Additional Tables

Table B1: Shares of Fuel Type by Planning Regions (ROR) - Tercile 1

T1 of the Unemployment Rate

State	ROR No.	Microcensus 2010			Building Sample		
		District Heat	Oil	Natural Gas	District Heat	Oil	Natural Gas
Ingolstadt	907	0.10	0.34	0.55	0.05	0.54	0.41
Landshut	908	0.00	0.54	0.46	0.02	0.49	0.48
Allgäu	901	0.06	0.57	0.38	0.03	0.53	0.44
Donau-Iller (BY)	904	0.00	0.57	0.43	0.04	0.52	0.44
Donau-Iller (BW)	802	0.18	0.37	0.45	0.09	0.41	0.51
Südostoberbayern	916	0.08	0.61	0.31	0.07	0.60	0.34
Augsburg	902	0.08	0.32	0.61	0.08	0.35	0.58
Westmittelfranken	917	0.00	0.55	0.45	0.01	0.59	0.40
Regensburg	915	0.00	0.57	0.43	0.02	0.52	0.46
München	910	0.25	0.30	0.45	0.12	0.37	0.52
Würzburg	918	0.00	0.35	0.65	0.06	0.35	0.59
Average		0.07	0.46	0.47	0.05	0.48	0.47

Notes: The table compares the shares of main fuel types by planning regions from the 2010 Microcensus statistics on housing and the billing panel used in the paper. A planning region falls into the tercile (T1) with the lowest unemployment rates, if more than 75 % of the zipcodes fall into T1. The data for the unemployment rate come from [RWI and microm \(2020\)](#) and Microcensus data from [Destatis \(2012\)](#). The Microcensus reports the number of apartment units heated with different fuel types, while the data used in the paper is at the building level. Thus, these comparison are an approximation at best.

Table B2: Shares of Fuel Type by Planning Regions (ROR) - Tercile 3
T3 of the Unemployment Rate

State	ROR No.	Microcensus 2010			Building Sample		
		District Heat	Oil	Natural Gas	District Heat	Oil	Natural Gas
Mecklenburgische Seenplatte	1301	0.65	0.12	0.24	0.15	0.32	0.53
Vorpommern	1303	0.62	0.10	0.28	0.34	0.08	0.58
Altmark	1501	0.43	0.14	0.43	0.30	0.04	0.65
Halle/S.	1503	0.45	0.15	0.40	0.27	0.13	0.60
Anhalt-Bitterfeld-Wittenberg	1502	0.42	0.16	0.41	0.23	0.11	0.67
West Sachsen	1404	0.42	0.08	0.50	0.20	0.07	0.73
Uckermark-Barnim	1205	0.58	0.08	0.34	0.33	0.16	0.51
Südsachsen	1403	0.33	0.11	0.56	0.16	0.09	0.76
Oderland-Spree	1203	0.47	0.08	0.45	0.42	0.03	0.55
Oberlausitz-Niederschlesien	1402	0.38	0.15	0.47	0.25	0.10	0.65
Oberes Elbtal/Osterzgebirge	1401	0.47	0.06	0.47	0.40	0.05	0.55
Mittleres Mecklenburg/Rostock	1302	0.62	0.07	0.31	0.43	0.03	0.54
Magdeburg	1504	0.38	0.09	0.53	0.17	0.07	0.76
Berlin	1101	0.44	0.21	0.35	0.28	0.24	0.48
Lausitz-Spreewald	1202	0.54	0.10	0.35	0.23	0.12	0.66
Nordthüringen	1602	0.38	0.14	0.48	0.13	0.06	0.81
Emscher-Lippe	509	0.23	0.13	0.64	0.21	0.13	0.65
Prignitz-Oberhavel	1204	0.43	0.13	0.44	0.35	0.10	0.56
Ostthüringen	1603	0.43	0.12	0.45	0.18	0.09	0.73
Mittelthüringen	1601	0.41	0.08	0.51	0.20	0.05	0.75
Westmecklenburg	1304	0.49	0.07	0.44	0.31	0.06	0.63
Ost-Friesland	312	0.04	0.07	0.89	0.04	0.05	0.91
Bremen	401	0.27	0.18	0.55	0.04	0.27	0.68
Dortmund	506	0.10	0.14	0.76	0.10	0.18	0.72
Average		0.42	0.11	0.47	0.24	0.11	0.65

Notes: The table compares the shares of main fuel types by planning regions from the 2010 Microcensus statistics on housing and the billing panel used in the paper. A planning region falls into the tercile (T3) with the highest unemployment rates, if more than 75 % of the zipcodes fall into T3. The data for the unemployment rate come from [RWI and microm \(2020\)](#) and Census data from [Destatis \(2012\)](#). The Microcensus reports the number of apartment units heated with different fuel types, while the data used in the paper is at the building level. Thus, these comparison are an approximation at best.

Table B3: Non-Linearities in Temperature Response

<i>Dependent variable: kWh/m²</i>			
	(1)	(2)	(3)
	Full	"Altbau"	Not WSVO1995
HDD if ($^{\circ}\text{C} < -5$)	0.0255*** (0.0003)	0.0280*** (0.0005)	0.0279*** (0.0006)
HDD if ($-5 \leq ^{\circ}\text{C} < 0$)	0.0307*** (0.0003)	0.0330*** (0.0004)	0.0322*** (0.0005)
HDD if ($0 \leq ^{\circ}\text{C} < 5$)	0.0266*** (0.0004)	0.0284*** (0.0005)	0.0282*** (0.0008)
HDD if ($5 \leq ^{\circ}\text{C} < 10$)	0.0232*** (0.0005)	0.0253*** (0.0007)	0.0255*** (0.0011)
HDD if ($10 \leq ^{\circ}\text{C} < 15$)	0.0210*** (0.0010)	0.0259*** (0.0014)	0.0275*** (0.0021)
Interaction Terms			
		X "Neubau"	X WSVO1995
HDD if ($^{\circ}\text{C} < -5$)		-0.0125*** (0.0005)	-0.00681*** (0.0009)
HDD if ($-5 \leq ^{\circ}\text{C} < 0$)		-0.0149*** (0.0004)	-0.0105*** (0.0005)
HDD if ($0 \leq ^{\circ}\text{C} < 5$)		-0.0123*** (0.0005)	-0.00887*** (0.0008)
HDD if ($5 \leq ^{\circ}\text{C} < 10$)		-0.0139*** (0.0007)	-0.0102*** (0.0010)
HDD if ($10 \leq ^{\circ}\text{C} < 15$)		-0.0182*** (0.0014)	-0.0123*** (0.0024)
<i>N</i>	3,030,063	1,363,017	439,126
Adj <i>R</i> ²	0.777	0.785	0.800
Zip Codes	7,769	6,860	5,340

Notes: The full sample covers all buildings using natural gas (high or low calorific), district heating, and oil. The dependent variable is kilowatt hours consumed per square meter of heated living space. "Altbau" equals 1 if the building was constructed before 1995. "Neubau" are buildings built starting 1995. WSVO 1995 is a composite measure indicating whether the roof, loft ceiling, windows, outer wall, and basement ceiling met thermal insulation standards under the 1995 building codes. This information was only available from energy performance certificates issued from 2014 to 2019. Standards errors in parentheses are clustered at the zip code level for all specifications. All regressions include fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B4: Accounting for Other Weather Variation

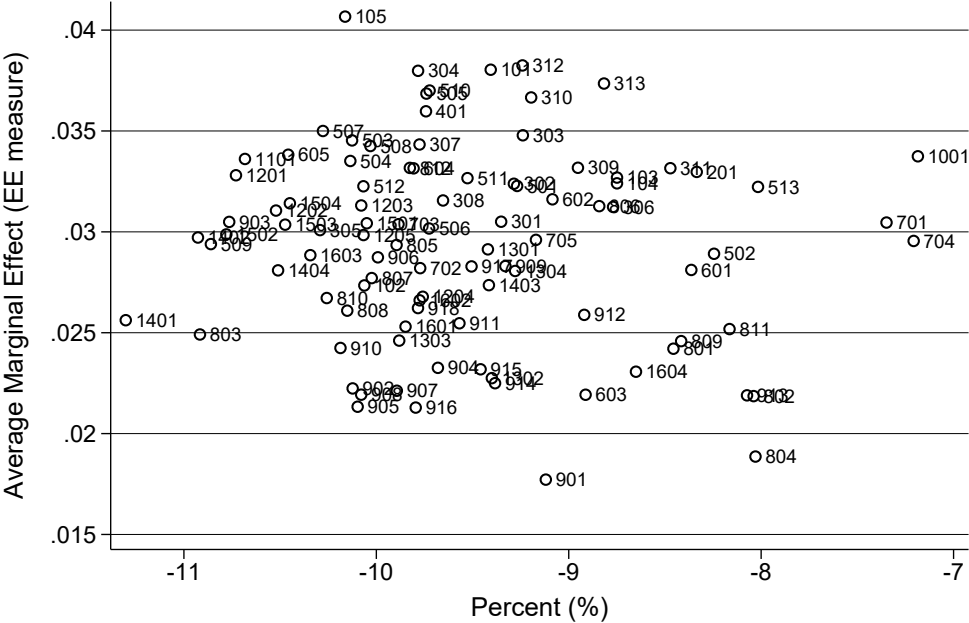
<i>Dependent variable: kWh/m²</i>				
	(1)	(2)	(3)	(4)
HDD	0.0306*** (0.0002)		0.0306*** (0.0002)	
HDD if ($^{\circ}\text{C} < -5$)		0.0255*** (0.0003)		0.0253*** (0.0003)
HDD if ($-5 \leq ^{\circ}\text{C} < 0$)		0.0307*** (0.0003)		0.0305*** (0.0003)
HDD if ($0 \leq ^{\circ}\text{C} < 5$)		0.0266*** (0.0004)		0.0262*** (0.0004)
HDD if ($5 \leq ^{\circ}\text{C} < 10$)		0.0232*** (0.0005)		0.0227*** (0.0005)
HDD if ($10 \leq ^{\circ}\text{C} < 15$)		0.0210*** (0.0010)		0.0197*** (0.0010)
Precipitation (mm)			-0.0565** (0.0212)	-0.0547** (0.0200)
Humidity (%)			0.691*** (0.112)	0.907*** (0.105)
<i>N</i>	3,030,063	3,030,063	3,030,063	3,030,063
Adj <i>R</i> ²	0.776	0.777	0.776	0.777

Notes: The full sample covers all buildings using natural gas (high or low calorific), district heating, and oil. The coefficient on the HDD should be interpreted as the increase in kWh per square meter living space for each 1°C increase in the annual sum of heating degree days, with mean temperatures below 15°C. Similarly in the sixth row, the coefficient should be interpreted as the increase in kWh per square meter for each 1°C increase in the annual sum of heating degree days, with mean temperatures inside the interval ($10 \geq ^{\circ}\text{C} < 15$). Precipitation is the average of total precipitation height (mm) on days with mean temperatures below 15°C. Relative humidity is the average of the mean relative humidity (%) on days with mean temperatures below 15°C. Standards errors in parentheses are clustered at the zip code level for all specifications. All regressions include fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

C Appendix

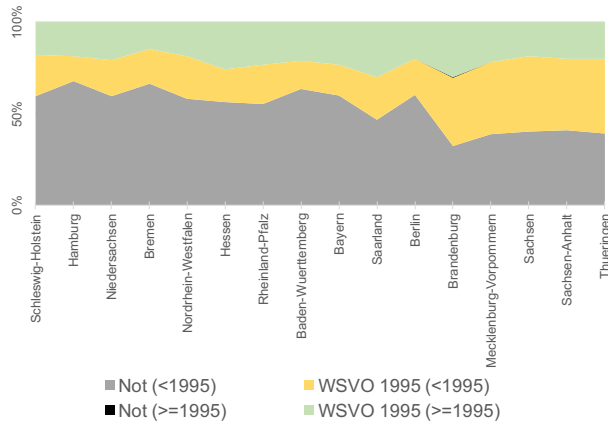
Additional Figures

Figure C1: Fall in Heating Degree Days

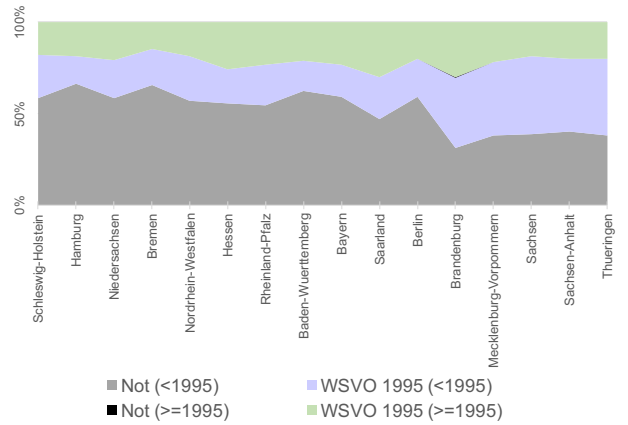


Notes: The horizontal axis shows the percent change in heating degree days (HDD), recorded from 2013 to 2018 relative to that from 2003 to 2012). The labels indicate the spatial planning region number. The vertical axis measures the average building-level marginal effect of HDD on heat demand for each of the 96 spatial planning region (ROR or Raumordnungsregion) in Germany. These are coefficients on the interaction term between spatial planning regions and heating degree days. The regression included fixed effects for building, fuel type, (majority) year of billing, the starting year of each billing period, and fuel type by year. The sample covers all buildings using natural gas (high or low calorific), district heating, and oil. Standards errors were clustered at the zip code level and all coefficients were statistically significant at the 1% level.

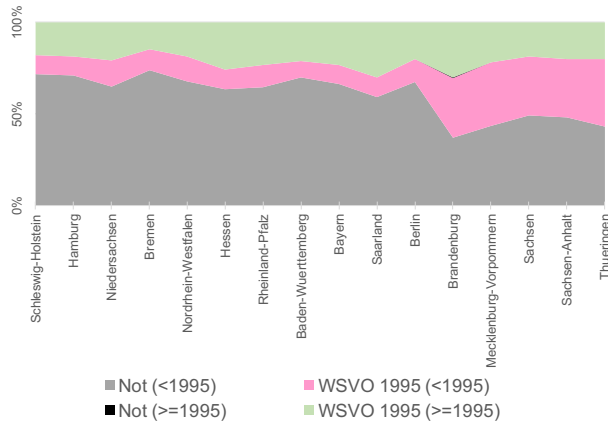
Figure C2: WSVO 1995 By Component



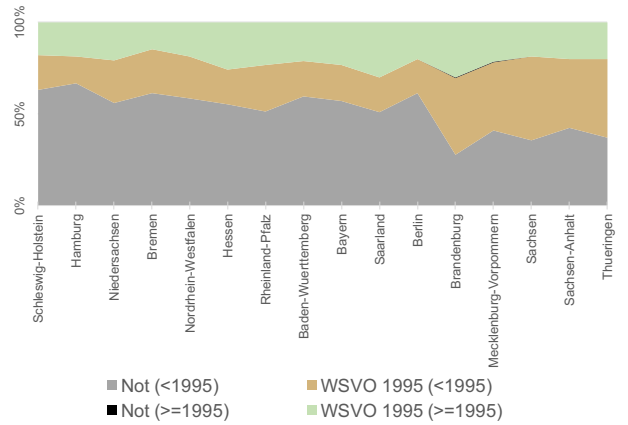
(a) Roof



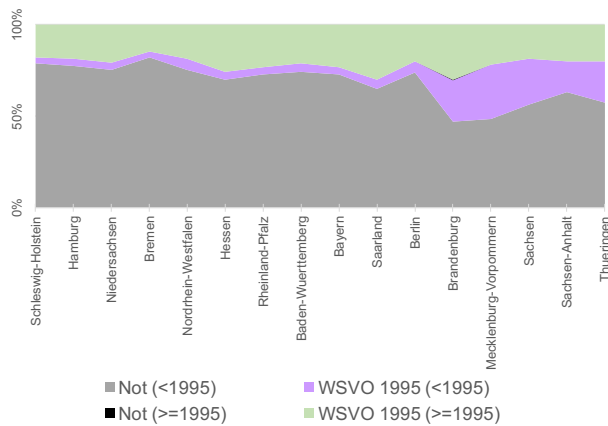
(b) Loft Ceiling



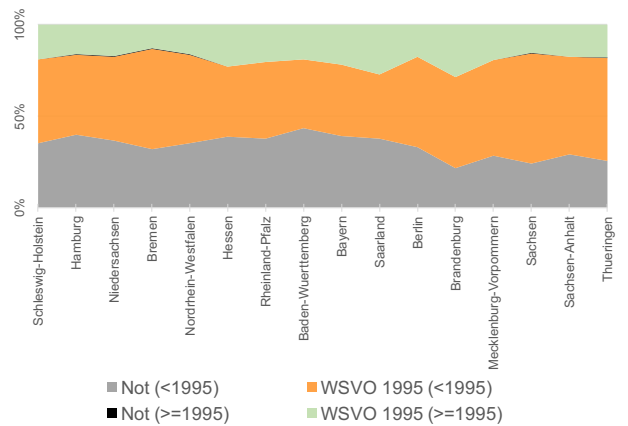
(c) Outer Wall



(d) Windows



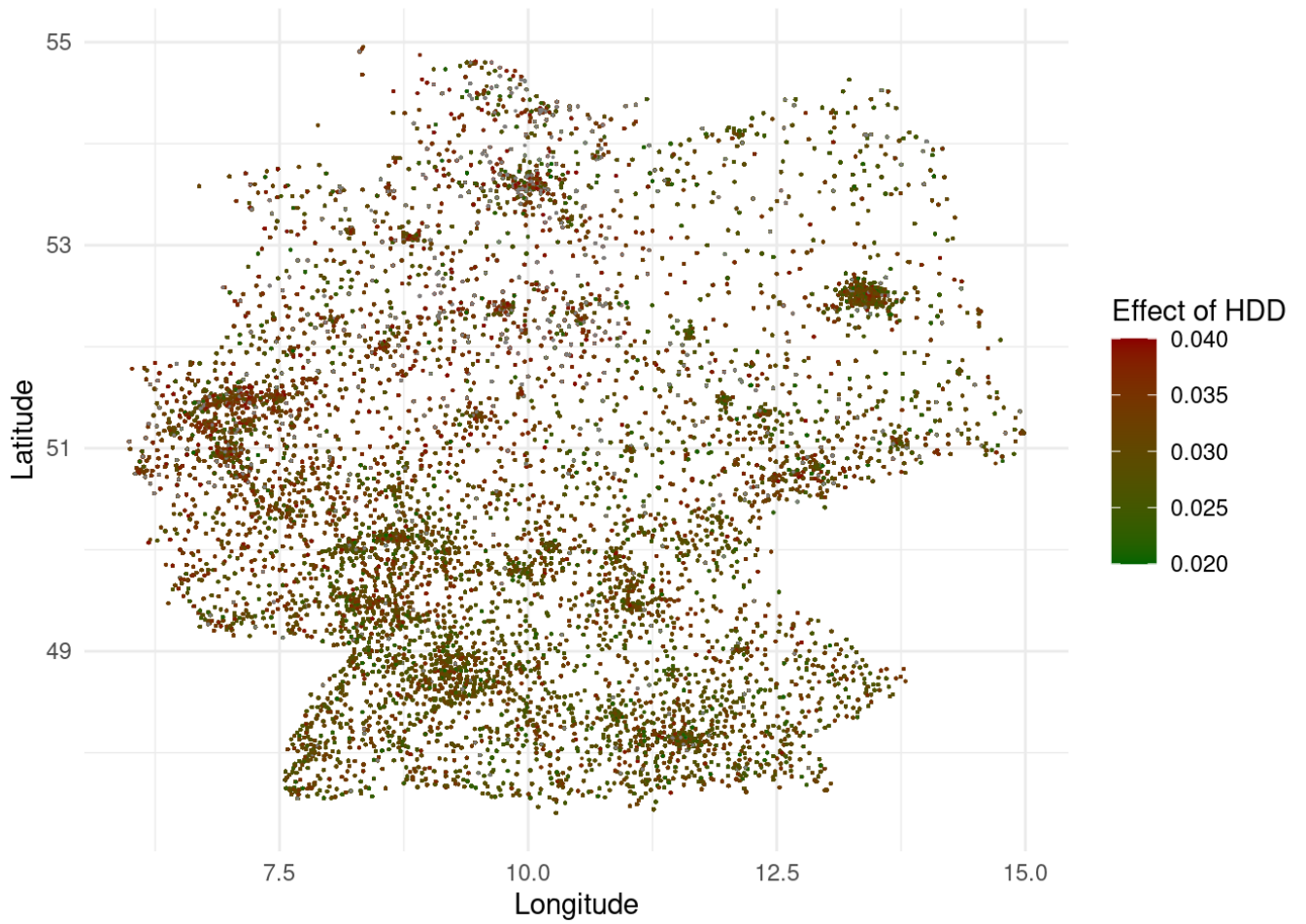
(e) Basement



(f) Heating System

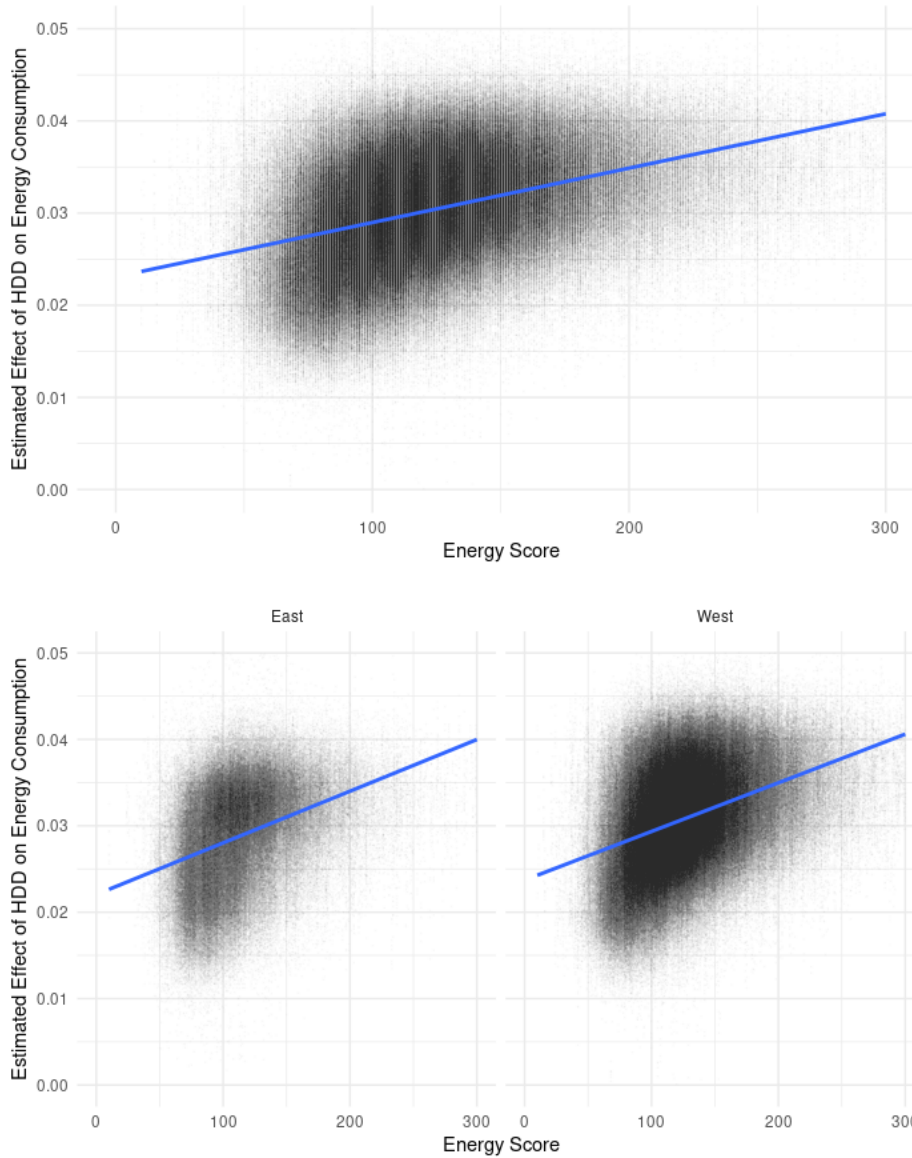
Notes: The graphs summarize the share of buildings that are certified to meet the minimum thermal insulation standards of 1995 for all six building components (the roof, loft or top ceiling, outer wall, windows, basement ceiling, and the heating system).

Figure C3:
Geographical Distribution of Buildings and
Marginal Effects from Causal Forests



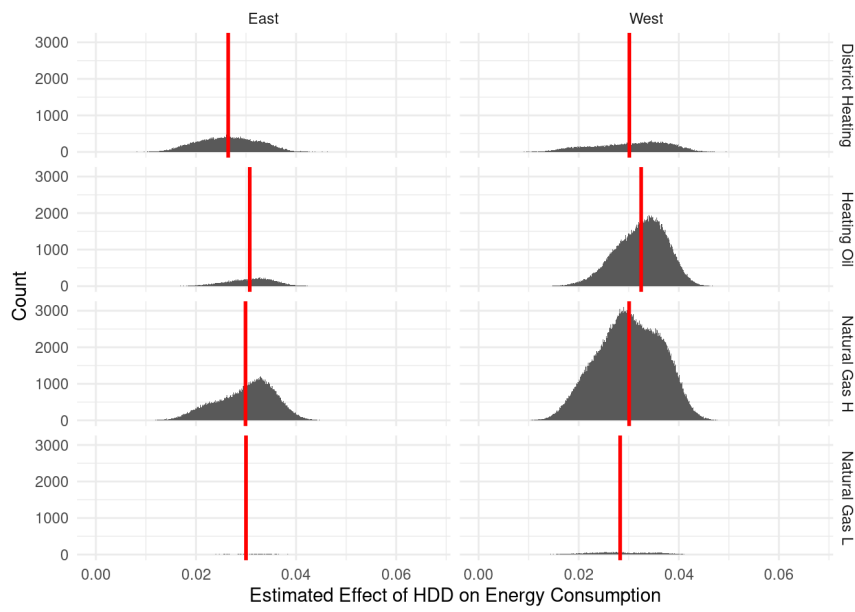
Notes: The map shows those buildings which were used in the main analysis. Each dot represents a building in the sample and depicts building-specific marginal effects generated using causal forests.

Figure C4:
Energy Performance Scores Versus Marginal Effects from Causal Forests



Notes: The graphs plot the energy performance scores reported on Energy Performance Certificates (issued from 2008 to 2019) against the building-specific marginal effects generated using causal forests. The energy scores are measured in kWh/square meter of living space required for heat energy annually. The type of EPC (demand or consumption-based) was not distinguished in the data sample.

Figure C5:
Distribution by Fuel Type for East and West Germany



Notes: The graph shows the distribution of ML-based building-specific marginal effects by main fuel type in East and West Germany.