ECONOMIC & SOCIAL RESEARCH INSTITUTE

ESRI Working Paper No. 749

April 2023

Are energy performance certificates a strong predictor of actual energy use? Evidence from high-frequency thermostat panel data

Tensay Hadush Meles[a,b*], Niall Farrell[a,b] & John Curtis[a,b,c]

a) Economic and Social Research Institute, Dublin, Ireland
b) School of Economics, Trinity College Dublin, Dublin, Ireland
c) SFI MaREI Centre for Energy, Climate and Marine, Ireland

*Corresponding Author:
Dr Tensay Hadush Meles,
Economic and Social Research Institute,
Whitaker Square, Sir John Rogerson's Quay,
Dublin, Ireland
Email: Tensay.HadushMeles@esri.ie

Acknowledgements:
We are grateful for the funding and support of the Energy Policy Research Centre (EPRC) at the Economic and Social Research Institute. We would also like to thank Hub Controls Ltd. for providing us access to the high-frequency smart thermostat panel data.

Abstract

This paper evaluates the extent with which Energy Performance Certificates (EPCs) reflect observed energy used for heating. We use high-frequency smart thermostat panel data in combination with building characteristics and hourly weather information. We exploit variations in boiler operation in the neighborhood of a steady state indoor temperature to elicit the predictive power of an EPC rating on energy use for heating. We find that the implied energy saving of upgrading from the lowest to highest EPC category is more than 3.5 times greater than that identified through ex-post analysis; boiler time operation is 52% greater among the lowest EPC-rated properties relative to the highest, while the EPC rating itself suggest a 183% difference in energy requirements. The findings cast doubt on the efficacy of public energy efficiency retrofit targets aligned to specific EPC standards.

Keywords: Building energy performance certificates; Smart thermostat; Ex-ante energy use; Ex-post evaluations; Energy efficiency; Climate change.

1 Introduction

This paper evaluates the extent with which building Energy Performance Certificates (EPCs) reflect observed energy used for heating. Improving the energy performance of buildings is commonly cited as an important component of a cost-effective decarbonization strategy (e.g. UNEP, 2021). According to the United Nations Environment Program (UNEP) (UNEP, 2021), investment in building energy efficiency is rising and reached about US$184 billion in 2020. This expenditure is primarily comprised of investment within the European Union.

There is a clear economic rationale for public support. While comfort and cost are private benefits internal to the decision to invest, external public benefits exist in the form of reduced carbon emissions. The privately optimal degree of investment may be misaligned with the social optimum and public support to internalize this externality is justified in many cases.

Policy targets are often defined relative to benchmarks provided by Building Energy Performance Certificates (EPCs). Should EPCs be unreflective of actual energy performance, then policy targets will be inefficient. The Government of Ireland's 2021 Climate Action Plan (CAP, 2021), for instance, states an objective to upgrade 500,000 existing homes to a Building Energy Rating (BER; the Irish EPC standard) of 'B2' by 2030. Energy efficiency retrofit grants are in place to internalize this externality and encourage adoption among households. Similar EPC benchmarks are issued in the European Union, USA, and other countries.[1]

[1]For details of the European Union (EU) Energy Performance of Buildings Directive see https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive. For energy efficiency certifications in the USA see https://www.energystar.gov/about.

With national-level policy formulated relative to the achievement of certain EPC benchmarks, so too are dwelling-level incentives. In Ireland, many supports are available to upgrade a dwelling to at least a 'B2' level on the 'BER' scale, the Irish EPC scale terms Building Energy Rating (SEAI, 2023a). In the UK, landlords of residential privately rented properties are legally required to have their proprieties a minimum of E-rated as part of building decarbonization (Sayce and Hossain, 2020). This implicitly incentivizes households to target a certain EPC rating when investing in the energy performance of their dwelling. The optimality of EPC-related investment incentives is determined by the accuracy of an EPC in capturing dwelling-specific energy performance. Should a discrepancy exist, then household investment will be misaligned with that which is cost-effective. With individual incentives and policy objectives centered around energy consumption as measured by EPCs, a policy question emerges: how well do ex-ante measures of building energy performance, such as those used by EPCs, capture energy and emissions savings?

Recent ex-post evaluations of energy efficiency investment cast doubt on the extent with which EPCs capture energy performance (e.g., Levinson, 2016; Fowlie et al., 2018; Davis et al., 2020; Coyne and Denny, 2021), with much of the literature focusing on the impact that behavior may have on any discrepancy. There are many ways in which behavior may drive this difference. For instance, the projected energy savings will be void if households in more energy-efficient homes respond to the lowered cost of energy by using more, which is known as a rebound effect (Gillingham et al., 2020; Aydin et al., 2017; Sorrell et al., 2009). There are further behavioral drivers. It may be the case that households with high income and large families, who are expected to have a greater energy demand, may self-select to dwellings with high energy efficiency.

Households with strong preferences for warm temperatures, or people who stay and work from home, may sort into more energy-efficient homes. These potential selection issues may lead to a spurious lack of observed differences in energy savings between more and less energy-efficient homes and result in understating the actual benefits of home energy efficiency programs and undermining their policy relevance.

In an attempt to address these identification problems, previous studies have exploited the introduction of new building energy codes (Jacobsen and Kotchen, 2013; Levinson, 2016; Kotchen, 2017; Novan et al., 2022); some have used eligibility to a home upgrade program as an instrumental variable (Fowlie et al., 2018; Davis et al., 2020; Hancevic and Sandoval, 2022); while others use the exogenous variation associated with external temperature response to account for unobserved heterogeneity (Chong, 2012; Liang et al., 2018). Still, evidence on the actual returns to energy efficiency investments remains inconclusive. This is mainly due to occupants' behavior. Davis et al. (2020) find no detectable impacts of home upgrades on electricity use or thermal comfort from a field trial in Mexico, suggesting that most households open their windows on hot days.

While the influence of behavior is the focus of much work, this is not the only factor that may drive a difference between estimated and observed rates of energy performance. Many of the calculations that comprise an EPC are based on standard values, rather than dwelling-specific values. Such standardization may arise during evaluation, where the inspector has incomplete information or an assumed value is used (DEAP, 2022). For instance, the degree of home energy efficiency and energy demand may vary with factors such as materials or degree of dwelling occupancy, factors that are unrecorded in standardized EPCs. This may introduce noise or bias into the estimated energy performance. A study by Christensen et al. (2021) examine the factors that drive a performance gap between expected and realized energy savings, focusing on dwelling characteristics. They find a bias in model projections and workmanship as primary determining factors, contributing up to 41% and 43% of the gap, respectively. Rebound effects are estimated to contribute only 6%.

Our study isolates the contribution of building fabric from occupants' behavior to evaluate the difference between assumed and observed energy consumption for Irish homes. Existing empirical studies to some extent concentrate on the USA (e.g. Christensen et al., 2021; Zivin and Novan, 2016). Evidence across different locations is limited although residential energy demand is climate sensitive. This paper attempts to fill in this gap by analyzing the impacts of building energy performance ratings on energy use for heating among existing Irish residential buildings, where the weather is relatively mild during winter and does not experience extremely cold temperatures like many other countries at similar latitude. Closely related to our study is Coyne and Denny (2021), who consider deviations in energy performance inclusive of behavioral effects. Coyne and Denny (2021) find a small difference in actual energy use, in contrast to ex-ante predicted energy use. While Coyne and Denny (2021) use bi-monthly metered energy data that necessarily includes unobserved behavioral impacts, the present analysis exploits high-frequency data. This allows us to focus on time periods where the effect of building fabric on energy use can be uniquely identified, isolating building fabric effects.

In providing this contribution, we adopt a novel methodological approach to isolate measurements of energy performance from unobserved occupant behaviors. As such, we focus on discrepancies associated with building fabric. We exploit variations in boiler operation while the indoor temperature is around a certain threshold of the thermostat set point during the main winter heating months. We consider the extent with which boiler operation varies across EPC ratings to maintain indoor temperature at thermostat set point levels. We use this as a proxy measure of the variations in energy use across building energy ratings attributable to building fabric alone, isolating this effect from occupant behavior.

The availability of high-frequency panel data (approximately every three minutes) from a smart thermostat company in combination with data on building characteristics and local weather allows us to clearly identify for how long a home heating unit was in operation. We model the relationship between the duration a boiler operates and the building energy performance ratings, conditioning on hourly outdoor temperature, relative humidity, wind speed, levels of thermostat set point, and several building characteristics.

We find a considerable deviation between ex-post estimates and projected energy requirements along the EPCs. For our sample, improving the EPC from E–G to B-rated decreases boiler operation by 52%, which is considerably less than the 183% difference in energy requirement predicted by the EPC ratings associated with the same properties. This is in line with the findings by Zivin and Novan (2016) and Meles et al. (2022). Zivin and Novan (2016) isolate the effects of energy efficiency upgrades and behavioral treatments and find energy savings from efficiency upgrades substantially smaller than ex-ante projections. Meles et al. (2022) evaluate variations in heat loss across EPC ratings during early morning hours when a heating system is turned off and occupants' behavioral effect is expected to be minimal. The study finds the performance gradient along the EPC ratings less than predicted ex-ante.

Our results also show that energy use for heating is outweighed by other factors than variations in EPCs. Thermostat set point temperature has a greater impact on boiler run time. Reducing the thermostat set point could reduce energy use for heating by a greater magnitude than improving a dwelling's energy performance ratings. From a policy perspective, a short-term change in occupant behavior relative to temperature preferences may be more effective in reducing energy use and CO_2 emissions, compared to costly energy-efficiency building upgrades.

The remainder of the paper is structured as follows. Section 2 describes the data and provides descriptive statistics. Section 3 outlines the empirical strategy. Section 4 presents and discusses the results and Section 5 concludes.

2 Data

The data used for this analysis are from a 'Hub Controller', an automatic energy manager device with a smart thermostat functionality, which we refer to as a 'smart thermostat'. This is installed in the main living area of the dwelling by Hub Controls Ltd.[2] The smart thermostat data covers a 2-year sample period from the 1st October 2019 to the 30th September 2021 for more than 10,000 dwellings with gas or oil boiler and built before 2006. This is a panel dataset that provides high-frequency observations on the following variables: household thermostat set point, indoor temperature, relative humidity, heating unit operating status, and heating mode status (on, off, 'boost'). These data are observed, on average, at 3-minute intervals. For our analysis, we construct relevant variables at an hourly level from the raw smart thermostat data.

[2]See https://thehubcontroller.com/ for more information.

These data are merged with other relevant data sources. Data on building energy performance and associated building characteristics are obtained from the Building Energy Rating (BER) database, which is available from the Sustainable Energy Authority of Ireland (SEAI) in an anonymized form.[3] The BER is the name for the Irish Energy Performance Certificate.

The BER assessment is based on the Dwelling Energy Assessment Procedure (DEAP), the official procedure for calculating and rating the energy performance of dwellings (DEAP, 2022). The procedure calculates the energy required for space heating, water heating, ventilation, and lighting, less savings from energy generation technologies. It generates a dwelling's annual delivered and primary energy consumption and associated CO_2 emissions using standard assumptions regarding occupancy, levels and duration of space heating and cooling, hot water demand and electricity usage for ventilation, pumps and lighting. For example, DEAP calculates the energy demand for space heating assuming the heating system operates 8 hours per day (07:00–09:00 and 17:00–23:00) during heating seasons running from October to May inclusive (8 months a year) to maintain the indoor temperature of the living room area to 21°C and the rest of the dwelling to 18°C.

Among others, the BER database contains a dwelling's BER in kWh/m^2/year and in letter grade that ranges from A1-rating (lowest primary energy usage) to G-rating (largest primary energy use). See Figure A.1 in the appendix for an example of a BER certificate that also displays all the BER scales and corresponding primary energy use in kWh/m^2/year. In addition to the energy required for space and water heating, the BER constitutes electricity use for ventilation and lighting, and takes into account

[3]The Building Energy Rating (BER) data is publicly available at https://ndber.seai.ie/BERResearchTool/ber/search.aspx

multiple other factors such as dwelling dimension, dwelling fabric, heating controls, fuel type, and renewable energy technologies (DEAP, 2022); thus it is an aggregate indicator of building energy efficiency.[4] The BER database contains information on location (at a county level), dwelling type and size, year of construction, number of storeys, the main fuel for space heating, space heating boiler efficiency, year of BER assessment, and purpose of assessment.

We subsequently match these data with hourly weather variables sourced from Ireland's National Meteorological Service, Met Éireann.[5]. These variables include outdoor temperature, relative humidity, and wind speed. Weather variables are taken from the local station at Dublin Airport. Upon merging, our final dataset comprises 798 residential buildings in Dublin and adjacent counties.

We are interested in estimating the variations in energy use across dwellings with different energy performance ratings, using boiler operation as a proxy for energy demand. Our outcome of interest is the duration a boiler operates (in minutes) in a given hour h at dwelling i, conditional on the measured indoor temperature being within a certain threshold, ϵ, of the thermostat set point temperature and the heating mode is turned on throughout the hour. We consider a threshold of ϵ=0.50°C for the main regression specification and check the sensitivity of the results by considering smaller and larger values: ϵ=0.35, 0.40, 0.45,...,0.60°C. By focusing on boiler operation when indoor temperature is close to the thermostat set point, we are capturing the length of boiler operation per hour to maintain temperature levels, akin to operation in a steady state. This approach excludes time periods when the

[4]Note that BER does not include electricity used for running home appliances like cookers, fridges, and washing machines.

[5]https://www.met.ie

boiler becomes operational and is raising temperatures to the set point level, i.e., a non-steady state.

We focus on sample dwellings with a gas or oil boiler as their main space heating fuel, excluding a small number of properties with heat pumps. We exclude irrelevant thermostat set point temperatures such as those less than the outdoor temperature and the base temperature for heating degree days (HDD) in Ireland (i.e., 15.5°C). We also restricted our data for analysis to the main winter heating months (December, January, and February) in Ireland to further ensure that we exclude indoor temperature readings around the set point threshold, particularly when a boiler is operating for water heating, not space heating. We can not distinguish whether a boiler is operating for space or water heating, though both are incorporated with the BER assessment. After dropping irrelevant observations for our analysis, we have 49,149 hourly observations of 'steady-state' operation across 320 dwellings. Since the share of each of the 15-point BER scale in the final sample is relatively small, we regroup them in the following BER categories: B, C, D, and E–G.[6]

Table 1 presents descriptive statistics of the 320 dwellings in the final sample across the BER scales. All the dwellings are located in the greater Dublin area, with an average ex-ante primary energy usage of about 242 kWh/m^2/year. Most of the dwellings are either semi-detached (47%) or terrace houses (43%), with most having two storeys (79%). A typical dwelling has a total floor area of 99 m^2 and is about 45 years since its construction. Gas boilers account for 89% of the main space heating, with an average efficiency of 82%. The average number of years since their BER

[6]Of the 320 sample dwellings, 4 are B2-rated, 28 are B3-rated, 30 are C1-rated, 51 are C2-rated, 57 are C3-rated, 46 are D1-rated, 46 are D2-rated, 21 are E1-rated, 11 are E2-rated, 16 are F-rated, and 10 are G-rated.

assessment is 6 years.

Table 1: Descriptive statistics of sample dwellings across BER scales

Variables	BER scales:				
	All	B	C	D	E–G
BER (kWh/m^2/year)	242.13	136.35	192.68	262.14	386.40
Dwelling type (%):					
Detached house	6.88	6.25	7.25	4.35	10.34
Semi-detached house	47.19	53.13	51.45	42.39	41.38
End of terrace house	16.56	6.25	11.59	25.00	20.69
Mid-terrace house	26.56	34.38	27.54	25.00	22.41
Ground-floor apartment	1.88	0.00	0.72	2.17	5.17
Mid-floor apartment	0.31	0.00	0.72	0.00	0.00
Top-floor apartment	0.63	0.00	0.72	1.09	0.00
Number of storeys (%):					
One storey	6.88	0.00	3.62	9.78	13.79
Two storeys	79.38	71.88	79.71	79.35	82.76
Three storeys	13.75	28.13	16.67	10.87	3.45
Building age in 2022 (in years)	44.86	41.31	37.49	47.46	60.22
Total floor area of a dwelling (m^2)	99.08	114.08	102.31	97.04	86.34
Area of a living room (m^2)	18.53	19.94	18.23	18.47	18.56
Main space heating fuel (%):					
Gas boiler	89.06	87.50	89.86	88.04	89.66
Oil boiler	10.94	12.50	10.14	11.96	10.34
Efficiency of main space boiler (%)	82.41	90.70	85.17	79.39	76.04
Years since BER assessment (in 2022)	6.07	3.34	5.86	6.60	7.24
Number of dwellings	320	32	138	92	58

Table 2 reports summary statistics of the hourly smart thermostat readings and weather variables for the dwellings across the BER scales. The reported values are only for hours that satisfy the condition: indoor temperature is within 0.50°C of the thermostat set point temperature and the heating mode is turned on throughout the hour. In the regression estimates, we check the sensitivity of the results to different temperature thresholds (0.35°C–0.60°C) of the thermostat set point temperature. The duration a boiler was operating while the indoor temperature was within 0.50°C of the mean thermostat set point temperature of 20°C (which varies from 15.5°C–30°C) ranges from zero to 60 minutes, with an average duration of approximately 17 minutes. During this period, the hourly outdoor temperature was between -5.6°C and +14.1°C, with an average of 5.5°C. The mean duration a boiler was operating increases along the BER scales, with a difference of 6 minutes between the best (B-rated) and worst (E–G rated) energy efficiency rated dwellings.

The hourly local weather variables (outdoor temperature, relative humidity and wind speed) are from the weather station at Dublin airport. The small differences in the reported hourly weather variables across the BER scales in Table 2 are due to variations in the hours at which the restriction of the temperature being within 0.50°C of the set point is satisfied. To control for this in regression modeling, we introduce dummies for hour of day, day of year, and year.

Table 2: Average values of the hourly smart thermostat readings and weather variables of the sample dwellings across BER scales

Variables	BER scales:				
	All	B	C	D	E–G
Duration a boiler was operating (in minutes)	16.66	14.42	16.11	17.78	20.31
Thermostat set point temperature (^{0}C)	19.78	18.84	19.93	20.04	20.40
Average indoor temperature (^{0}C)	19.82	18.87	19.97	20.06	20.40
Average indoor relative humidity (%)	46.44	46.48	46.04	46.08	48.78
Average outdoor temperature (^{0}C)	5.50	5.40	5.44	5.54	5.76
Average outdoor relative humidity (%)	85.12	85.12	85.23	85.03	84.94
Average wind speed (knot)	10.30	10.14	10.32	10.41	10.22
Observations	49,149	10,070	20,800	12,965	5,314
Number of dwellings	320	32	138	92	58

3 Empirical strategy

In order to examine the effects of building energy performance certificates on heating demand, we specify and estimate the following panel data model.

$$Y_{ihdmt} = \beta BER_i + \gamma_k \sum_k Temp_{k,hdmt} + \delta_j \sum_j Setpoint_{j,ihdmt} + \theta' X + \lambda_{h,d,t} + U_{ihdmt} \quad (1)$$

Where the outcome variable Y_{idmt} is the duration a boiler was in operation (in minutes) while the indoor temperature is within a certain threshold, ϵ, of the thermostat set point temperature and the heating mode was turned on throughout hour h at dwelling i on day d, month m, and year t. Boiler time is a proxy measure of energy use, indicating for how long the boiler was turned on and fired to keep the indoor temperature within ϵ°C of the set point. We would expect the duration a boiler operates to be shorter for dwellings with better building energy ratings.

BER_i is our measure of building energy efficiency of dwelling i, which is specified in grade letters from B to E–G. While B-rated dwellings are the most energy efficient (with a primary energy demand of 75 kWh/m^2/year – 150 kWh/m^2/year), E–G rated dwellings are the least energy efficient dwellings (with a primary energy demand of above 300 kWh/m^2/year). The coefficient of interest, β, is interpreted as the impact of BER scales on boiler operation (a proxy for energy use), with B-rated dwellings as a reference category.

$Temp_{k,hdmt}$ is outdoor temperature, which is a main determinant of space heating demand. To capture the non-linear effects of outdoor temperature, like in Deschênes and Greenstone (2011) and Ge and Ho (2019), we construct seven temperature bins with 2°C intervals for the average hourly outdoor temperature that ranges from -5.6°C to +14.1°C over sample period.[7] $Temp_{k,hdmt}$ equals one if the hourly outdoor temperature is in the kth of the seven temperature bins. The coefficient γ_k is the impact of outdoor temperature in the kth bin on the duration a boiler was operating, with 10°C or above as a baseline temperature bin.

We also control the thermostat set point $Setpoint_{j,ihdmt}$ at dwelling i as the duration a boiler turns on and runs to keep the indoor temperature within ϵ°C depends on the level of the set point temperatures. For instance, at the average hourly outside temperature of 5.5°C, a boiler need to operate for a longer duration to keep the indoor temperature within 0.50°C of a set point of 20°C compared to 18°C. Similar to the hourly outdoor temperature, we construct seven thermostat set point temperature bins with 1°C intervals for the hourly thermostat set point temperatures that range from 15.5°C to 30°C across the sample dwellings over the sample period. The coefficient δ_j captures the effects of a thermostat set point temperature in the jth bin on the duration a boiler was operating, with 17°C or less as a baseline thermostat set point temperature bin.

In addition, we control for a vector of other variables, X, that potentially affect space heating demand. This includes the time-invariant building characteristics such as dwelling area, dwelling type, building age, number of storeys, the main fuel for space heating (gas or oil boiler) and its efficiency, and average hourly outdoor relative humidity (%) and wind speed (knot).

$\lambda_{h,d,t}$ represents different possible time-fixed effects. This includes dummies for hours of the day, days of the year, and year-of-sample to control for common hourly

[7] Since the shares in the first and last categories are small, we regroup them into temperature bins of $\leq$ 0°C and $\geq$ 10°C, respectively.

or daily routines and other common factors such as changes in fuel prices. U_{ihdmt} is a stochastic error term. In all specifications, we cluster standard errors at the dwelling level to account for serial correlations within a dwelling.

4 Results and Discussion

Table 3 presents the main results from the specification in Equation 1, with robust standard errors clustered at the dwelling level. Column 1 shows the impact of building energy ratings (BER) on the duration a boiler operates (in minutes) to maintain the thermostat set point temperatures within 0.50°C in a given hour, after controlling for the outdoor temperature, levels of the temperature set point and time fixed effects. In Column 2, we add a vector of other controls for dwelling and main space heating system characteristics.

Our results show that the duration a boiler operates significantly varies across the BER scales, and hourly set point temperatures and outdoor temperatures. The results remain similar when we include controls on characteristics of the dwelling and heating systems, albeit the magnitude of the estimated coefficients increases. Since the specification in Column 2 explains the variations relatively better and captures the potential effects of various dwelling and heating system characteristics, we consider it as our main specification in interpreting the estimated results.

The estimated coefficients on the BER scales are statistically significant and increase along the BER scales relative to the B-rated reference category. The coefficient on the C-rated dwelling implies that on average a boiler in a C-rated dwelling is operational for approximately 3 minutes longer than that in a B-rated property in order to maintain the thermostat set point temperature in a given hour. Keeping all other things constant, a boiler in an E–G rated dwelling, on average, operates about 8 minutes more compared to a boiler in a B-rated dwelling. The results indicate the more energy efficient the dwelling, as measured by BER rating, the less time the boiler operates. The pattern of the gradient in energy performance along the BER scales matches a priori expectations.

Even though our primary interest is in estimating the impacts of BER scales on a boiler operation, it is worth commenting on the estimated coefficients of the set point and outdoor temperatures. A change in hourly thermostat set point temperature is positively related to boiler operation. The higher the thermostat set point temperature, the longer a boiler operates to maintain the indoor temperature within 0.50°C of the set point for an entire hour. Relative to the reference set point temperature of 17°C or less, the estimated additional average duration a boiler operates ranges from about 2 minutes (for set point temperature of 17–18°C) to 21 minutes (for set point temperature of 22°C or above). On the other hand, a change in outdoor temperature is inversely related to a boiler operation. The magnitude of the effects varies from about 3 minutes (for an outdoor temperature of 8–10°C) to 13 minutes (for an outdoor temperature of 0°C or below), compared to the base hourly outdoor temperature of 10°C or above. The estimated effects on the duration a boiler operates show that the variations in boiler operation across dwellings are to a greater extent driven by thermostat set point values and outdoor temperatures than the BER rating of a property.

The estimated coefficients on the set point temperatures are useful in providing estimates on how adjusting the thermostat set point affects energy savings, which is

Table 3: Impacts of building energy performance ratings on heating demand

Variables	(1)	(2)
	Dependent variable: Duration a boiler operates (in minutes at hour h)	
BER scales (Reference: B):		
C	2.64*	3.38**
	(1.35)	(1.46)
D	3.51**	4.49**
	(1.47)	(1.77)
E–G	5.81***	7.53***
	(1.56)	(2.27)
Set point temperature at hour h (Reference: $\leq 17^0$C):		
(17^0C, 18^0C]	2.26***	2.28***
	(0.50)	(0.49)
(18^0C, 19^0C]	5.96***	6.00***
	(0.93)	(0.93)
(19^0C, 20^0C]	10.07***	10.13***
	(0.89)	(0.90)
(20^0C, 21^0C]	13.40***	13.50***
	(0.99)	(1.00)
(21^0C, 22^0C]	15.93***	16.06***
	(0.97)	(0.98)
$> 22^0$C	20.62***	20.73***
	(1.26)	(1.29)
Outdoor temperature at hour h (Reference: $\geq 10^0$C):		
$\leq 0^0$C	12.14***	12.98***
	(1.01)	(1.05)
(0^0C, 2^0C)	10.60***	11.13***
	(0.74)	(0.76)
[2^0C, 4^0C)	8.69***	9.13***
	(0.57)	(0.58)
[4^0C, 6^0C)	6.54***	6.77***
	(0.42)	(0.43)
[6^0C, 8^0C)	4.04***	4.37***
	(0.34)	(0.34)
[8^0C, 10^0C)	2.33***	2.46***
	(0.28)	(0.27)
Hour of a day dummies	Yes	Yes
Day of a year dummies	Yes	Yes
Year dummies	Yes	Yes
Other controls	No	Yes
R^2 (within)	0.206	0.210
Observations	49,149	49,149
Sample dwellings	320	320

Note: The dependent variable is the duration a boiler was operating for heating (in minutes) to maintain the indoor temperature within 0.50°C of the set point temperature in a given hour. The set point temperature bins are based on the hourly level thermostat set point temperatures that range from 15.5°C to 30°C for the sample dwellings. Similarly, the outdoor temperature bins are based on the hourly outdoor temperatures that vary from -5.6°C to +14.1°C during the period of analysis. The other controls included are hourly outdoor relative humidity, wind speed, dwelling area, efficiency of main space heating unit, dummies for fuels of main space heating (gas or oil boiler), dwelling type, number of storeys, and building age band. Robust standard errors clustered at the dwelling level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

a issue of acute policy relevance during the current high energy price crisis. From a policy perspective, adjusting the thermostat set point temperature is easy, and relatively instant in terms of effect compared to energy efficiency retrofits. Similarly, the estimated coefficients on the outdoor temperature are relevant to utilities and policymakers in providing rough estimates of energy demand during cold winter weather.

We check the robustness of our results by both narrowing and widening ϵ, the steady-state threshold around the thermostat set point. Table A.1 in the appendix presents the estimated coefficients with values of ϵ that range from 0.35°C – 0.60°C. Widening the window towards 0.60°C or narrowing it to 0.35°C does not much change the estimated impacts on boiler operation for heating across the BER scales; only a slight change in the magnitude of the coefficients. The small variations with a narrower or wider temperature window are likely due to the change in the number of hourly observations that satisfy the restrictions. The overall pattern remains similar as in the baseline results with 0.50^0C window.

For a meaningful economic interpretation and policy insights, we need to express the duration a boiler operates into energy demand and CO_2 emissions. For this, we convert the estimated effects on the duration a boiler operates in Table 3 into energy use (kWh) and CO_2 emissions (kg) using a typical boiler power capacity in Ireland and the CO_2 emissions factor for the main fuels for heating. The power capacity of a domestic boiler in Ireland generally ranges between 18 kW, 24 kW, or 30 kW, depending on the number of bedrooms.[8] We provide energy and emissions calculations based on these three typical boiler sizes. The emissions factor of natural gas is 0.203

[8]For example, see https://www.gasworks.ie/

kg/kWh and 0.272 kg/kWh for kerosene heating oil (DEAP, 2022). The estimated coefficients in Table 3 are the relative difference in effects on the duration a boiler operates (minutes) of the variable of interest, compared to the reference category. The energy use in kWh is obtained by multiplying the relative effects on the duration a boiler operates (in hours) with a boiler power capacity, while the CO_2 emissions is the energy use in kWh multiplied by a fuel emission factor. These calculations are reported in Table 4. These estimates indicate that compared to B-rated dwellings, the hourly energy use for heating is approximately 1.01 – 1.69 kWh higher in C-rated dwellings and 1.35 – 2.25 kWh more in D-rated dwellings, depending on the power capacity of the boiler. At the bottom end of the energy performance rating, the average hourly energy use for heating in E–G rated dwellings is 2.26 – 3.77 kWh higher. If the estimated impacts on boiler operation for heating are measured in CO_2 emissions (kg), the calculated hourly emissions are, on average, 0.21 – 0.35 kg more in C-rated dwellings, 0.28 – 0.47 kg in D-rated, and 0.47 – 0.79 kg in E–G rated dwellings, all relative to B-rated reference category. It is worth reiterating that these estimates relate to boiler operation during 'steady-state' operation and excludes periods when properties are being heated from a 'cold' state to reach thermostat set point values.

The magnitude of the estimates of the impact of BER is much larger than comparable findings by Coyne and Denny (2021) and Meles et al. (2022). The paper by Meles et al. (2022) is based on the same underlying dataset as used here but is based on a larger set of dwellings and also investigates a different metric; variations in heat loss during the early morning hours when the heating unit is confirmed as being turned off. The Meles et al. (2022) study finds a difference in heat loss between

BER scales but the magnitude is less than that expected ex-ante and no evidence of a distinct gradient in performance along the BER scales. Coyne and Denny (2021), in an analysis that incorporates behavioral response, document little difference in actual energy use across Irish residential building energy performance certificates. Unlike our hourly-level smart thermostat analysis, the methodology of Coyne and Denny (2021) does not control for unobserved occupant behavior and uses bi-monthly energy consumption data. Similar differences in research conclusions based on low and high-frequency data also arise in the analysis of electricity savings in California (Novan et al., 2022; Levinson, 2016). This highlights the importance of high-frequency, household-level data in evaluating energy efficiency investments.

The results covering the influence of the thermostat set point on energy consumption are also of interest. Compared to the impacts of BER scales on energy use and CO_2 emissions, the effects of thermostat set point temperatures and outdoor temperatures are substantially greater in magnitude. For example, the hourly energy used at a set point temperature of 22°C or above is as much as 10.37 kWh higher than the reference set point of 17°C or below. The corresponding hourly CO_2 emissions are 2.18 kg higher. The impact on energy use and emissions from a reduction in the set point temperature by 2°C where the set point is 20°C and above is not practically different than that associated with a BER rating improvement from E–G to B. This result is not evidence that energy efficiency upgrades do not work, rather it is evidence that the energy efficiency label, i.e., the DEAP standard, is not a strong indicator of actual energy use.

Finally, we attempt to relate the relative differences of the ex-post estimates (regression results) along the BER scales with the relative differences in ex-ante primary

Table 4: Calculated average hourly energy use and CO_2 emissions for different power capacity boilers

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Energy use in kWh for a boiler:			CO_2 emissions in kg for a boiler:		
	18 kW	24 kW	30 kW	18 kW	24 kW	30 kW
BER scales (Reference:B):						
C	1.01	1.35	1.69	0.21	0.28	0.35
D	1.35	1.80	2.25	0.28	0.38	0.47
E–G	2.26	3.01	3.77	0.47	0.63	0.79
Set point temperature at hour h (Reference: $\leq 17^0$C):						
(17^0C, 18^0C]	0.68	0.91	1.14	0.14	0.19	0.24
(18^0C, 19^0C]	1.80	2.40	3.00	0.38	0.50	0.63
(19^0C, 20^0C]	3.04	4.05	5.07	0.64	0.85	1.06
(20^0C, 21^0C]	4.05	5.40	6.75	0.85	1.13	1.42
(21^0C, 22^0C]	4.82	6.42	8.03	1.01	1.35	1.69
$> 22^0$C	6.22	8.29	10.37	1.31	1.74	2.18
Outdoor temperature at hour h (Reference: $\geq 10^0$C):						
$\leq 0^0$C	3.89	5.19	6.49	0.82	1.09	1.36
(0^0C, 2^0C)	3.34	4.45	5.57	0.70	0.93	1.17
[2^0C, 4^0C)	2.74	3.65	4.57	0.58	0.77	0.96
[4^0C, 6^0C)	2.03	2.71	3.39	0.43	0.57	0.71
[6^0C, 8^0C)	1.31	1.75	2.19	0.28	0.37	0.46
[8^0C, 10^0C)	0.74	0.98	1.23	0.15	0.21	0.26
Sample dwellings	320	320	320	320	320	320

Note: Table 4 shows the calculated hourly average energy use in kWh and CO_2 emissions of the estimated coefficients in Table 3 for boilers with 18 kW, 24 kW and 30 kW power capacity, ranges of typical domestic boilers in Ireland. The emissions factor of the fuel for main gas is 0.203 kg/kWh and 0.272 for heating oil. In our data, the average emissions factor in each of the four BER scales is 0.21 kg/kWh.

energy use (BER in kWh/m^2/year), using B-rated dwellings as a reference. Direct comparison of the ex-ante and ex-post estimates is problematic for a couple of reasons. First, our empirical analysis is based on the variations in a boiler operation around the steady state indoor temperature (set point temperature) for the three main winter heating months (December–February) while the ex-ante estimates of primary energy demand are based on 8 months (October–May), with 8 hours a day of heating periods (DEAP, 2022). Thus, our paper does not capture a boiler operation or energy use for heating for the entire period, even within the three main winter heating months. Second, our ex-ante measure of building energy efficiency, the BER certificate, is broader in scope than the energy use for heating that we are primarily interested in analyzing. Beyond energy use for heating, BER includes energy use for ventilation and lighting and accounts for other aspects of a building such as dwelling dimensions, building fabric, type of fuel, and renewable energy technologies. It is, therefore, an aggregate indicator of dwelling energy performance.

Considering those limitations, we provide insights into the ex-post and ex-ante estimates by looking at the relative differences in ex-ante and ex-post energy performance across the BER scales. Table 5 shows the relative differences in the average values of the BER in kWh/m^2/year (ex-ante estimates) and regression results (ex-post estimates) across BER scales relative to B-rated dwellings. Compared to B-rated dwellings, the average BER in kWh/m^2/year (ex-ante primary energy demand) is 41% higher for C-rated dwellings, 92% for D-rated dwellings and 183% for E-rated or below dwellings. Based on the 14.42 minutes average duration of a boiler operation to maintain the indoor temperature around the thermostat set point in an hour in B-rated dwellings (see Table 2), the estimated average boiler operation for a typical

C-rated dwelling is 23% more, 31% more for D-rated dwelling and 52% more for E-rated or below dwellings. This indicates variations in boiler operation along the BER scales are substantially proportionately less than projected primary energy demand (kWh/m^2/year) along the BER scales. While the results presented earlier in Table 3 and again in Table 5 show a clear gradient of performance across BER scales, that gradient is substantially less than that implied from the official BER ratings, as measured in kWh/m^2/year. Consequently, directly linking policy targets to BER ratings is likely to lead to performance outcome substantially less than anticipated.

Table 5: Average ex-ante and ex-post estimates across BER scales

	Ex-ante estimates:		Ex-post estimates:	
BER scales	Average BER (in kWh/m^2/year)	Relative change	Average boiler operation (in minutes at hour h)	Relative change
B	136.35	Reference	14.42	Reference
C	192.68	+41%	17.80	+23%
D	262.14	+92%	18.91	+31%
E	386.40	+183%	21.95	+52%
Sample dwellings	320		320	

Note: The average BER in kWh/m^2/year is based on the BER database for our sample dwellings. The average boiler operation in minutes is constructed using the mean duration a boiler operates for B-rated dwelling in our data (see Table 2) and relative differences of the estimated coefficients in Column 2 of Table 3.

5 Conclusions

Improving building energy efficiency is advocated as one of the most cost-effective approaches to address climate change, with building Energy Performance Certificates (EPC) serving as a benchmark of performance in many markets. For example, the Government of Ireland's 2021 Climate Action Plan has targeted to upgrade half a million homes to B2 rating by 2030, with €8 billion retrofit scheme. In this paper, we ask the extent to which residential building energy performance certificates, an ex-ante measure of home energy efficiency, predict observed energy and emissions savings.

To answer this, we use high-frequency smart thermostat panel data in combination with building characteristics and local weather information and exploit variations in boiler operation for heating in the neighborhood of thermostat set point temperatures during the main winter heating months. We find a distinct gradient of building energy performance along the Irish EPC metric, the Building Energy Rating (BER), but with a gradient substantially less than implied by ex-ante energy use associated with BER scales.

In the context of policy ambition to reduce both fossil energy use and emissions, the three most important factors explaining length of boiler operation are weather, thermostat set point temperature, and BER scale. Unlike weather, both BER scale and temperature set points are within homeowners' influence but represent a complex choice. Adjusting internal temperature set point values is simple to implement, has an immediate impact, and saves money via lower fuel expenditures but potentially also reduces home comfort and well-being. Achieving the B2 standard by 2030 is the government policy target but improving a property's BER scale is expensive (Collins and Curtis, 2017; Moran et al., 2020), with a move from an E–G to a B rating potentially costing €30–50,000/property. In addition to the financial costs, retrofits are relatively slow to procure plus entail considerable disruption to occupants. The policy question is whether the B2 policy target is the most cost-effective approach to reducing fossil energy use and emissions.

This paper provides evidence to suggest that EPCs may misguide energy efficiency investments. There are many conceivable extensions to this work. The dwellings in the analysis are not a representative sample of the national housing stock so while they provide a suitable sample to estimate whether, and to what extent, a gap exists, the results cannot be easily extrapolated to represent all dwellings. Nonetheless, the sample size for the analysis is relatively large and substantially greater than several existing studies examining energy efficiency within the Irish housing stock (e.g., Beagon et al., 2018; Rau et al., 2020).

Our findings are unambiguously illustrative of two things. First, energy use in residential buildings declines with improvement in energy efficiency, as measured by the BER standard. However, the relative differences of our estimated results along the BER scales are substantially lower than that of the relative differences in the projected primary energy demand (BER in kWh/m^2/year). Second, the difference in energy use attributable to the BER scales is modest and not substantially different in magnitude than minor behavioral interventions. These results do not imply that upgrading a dwelling's energy efficiency would not result in a significant change in energy use and emissions; rather it is evidence that BER certificates are broader in scope than energy use for heating. The findings underscore that BERs are poor predictors of actual energy use and consequently cast doubt on the efficacy of public energy efficiency retrofit targets that are aligned to B2 BER standard.

References

Aydin, E., Kok, N., Brounen, D., 2017. Energy efficiency and household behavior: The rebound effect in the residential sector. The RAND Journal of Economics 48, 749–782.

Beagon, P., Boland, F., O'Donnell, J., 2018. Quantitative evaluation of deep retrofitted social housing using metered gas data. Energy and Buildings 170, 242–256. https://doi.org/10.1016/j.enbuild.2018.04.022.

CAP, 2021. Climate Action Plan 2021: Securing our Future. Government of Ireland. Accessed July 25, 2022.

Chong, H., 2012. Building vintage and electricity use: Old homes use less electricity in hot weather. European Economic Review 56, 906–930.

Christensen, P., Francisco, P., Myers, E., Souza, M., 2021. Decomposing the Wedge between Projected and Realized Returns in Energy Efficiency Programs. The Review of Economics and Statistics , 1–46.

Collins, M., Curtis, J., 2017. Value for money in energy efficiency retrofits in Ireland: grant provider and grant recipients. Applied Economics 49, 5245–5267. https://doi.org/10.1080/00036846.2017.1302068.

Coyne, B., Denny, E., 2021. Mind the energy performance gap: testing the accuracy of building energy performance certificates in Ireland. Energy Efficiency 14, 1–28. https://doi.org/10.1007/s12053-021-09960-1.

Davis, L.W., Martinez, S., Taboada, B., 2020. How effective is energy-efficient housing? Evidence from a field trial in Mexico. Journal of Development Economics 143, 102390.

DEAP, 2022. Dwelling Energy Assessment Procedure (DEAP) software and methodology. Sustainable Energy Authority of Ireland. https://www.seai.ie/home-energy/building-energy-rating-ber/support-for-ber-assessors/software/deap/. Accessed May 20, 2022.

Deschênes, O., Greenstone, M., 2011. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. American Economic Journal: Applied Economics 3, 152–85.

Fowlie, M., Greenstone, M., Wolfram, C., 2018. Do energy efficiency investments deliver? Evidence from the Weatherization Assistance Program. The Quarterly Journal of Economics 133, 1597–1644.

Ge, Q., Ho, B., 2019. Energy use and temperature habituation: Evidence from high frequency thermostat usage data. Economic Inquiry 57, 1196–1214.

Gillingham, K., Rapson, D., Wagner, G., 2020. The rebound effect and energy efficiency policy. Review of Environmental Economics and Policy .

Hancevic, P.I., Sandoval, H.H., 2022. Low-income energy efficiency programs and energy consumption. Journal of Environmental Economics and Management 113, 102656.

Jacobsen, G.D., Kotchen, M.J., 2013. Are building codes effective at saving energy?

Evidence from residential billing data in Florida. Review of Economics and Statistics 95, 34–49.

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, 135–153.

Levinson, A., 2016. How much energy do building energy codes save? Evidence from California houses. American Economic Review 106, 2867–94.

Liang, J., Qiu, Y., James, T., Ruddell, B.L., Dalrymple, M., Earl, S., Castelazo, A., 2018. Do energy retrofits work? Evidence from commercial and residential buildings in Phoenix. Journal of Environmental Economics and Management 92, 726–743.

Meles, T.H., Farrell, N., Curtis, J., 2022. How well do building energy performance certificates predict heat? Economic and Social Research Institute (ESRI) Working Paper Series 733. https://www.esri.ie/pubs/WP733.pdf.

Moran, P., O'Connell, J., Goggins, J., 2020. Sustainable energy efficiency retrofits as residential buildings move towards nearly zero energy building (NZEB) standards. Energy and Buildings 211, 109816. https://doi.org/10.1016/j.enbuild.2020.109816.

Novan, K., Smith, A., Zhou, T., 2022. Residential building codes do save energy: Evidence from hourly smart-meter data. The Review of Economics and Statistics , 1–18.

Rau, H., Moran, P., Manton, R., Goggins, J., 2020. Changing energy cultures? Household energy use before and after a building energy efficiency retrofit. Sustainable Cities and Society 54, 101983. https://doi.org/10.1016/j.scs.2019.101983.

Sayce, S.L., Hossain, S.M., 2020. The initial impacts of minimum energy efficiency standards (MEES) in England. Journal of Property Investment & Finance 38, 435–447.

SEAI, 2023a. Home energy grants. Sustainable Energy Authority of Ireland. https://www.seai.ie/grants/home-energy-grants/. Accessed March 06, 2023.

Sorrell, S., Dimitropoulos, J., Sommerville, M., 2009. Empirical estimates of the direct rebound effect: A review. Energy policy 37, 1356–1371.

UNEP, 2021. 2021 Global status report for buildings and construction: Towards a zero-emission, efficient and resilient buildings and construction sector. United Nations Environment Programme, Nairobi. https://www.globalabc.org. Accessed April 29, 2022.

Zivin, J.G., Novan, K., 2016. Upgrading efficiency and behavior:Electricity savings from residential weatherization programs. The Energy Journal 37.

Appendix

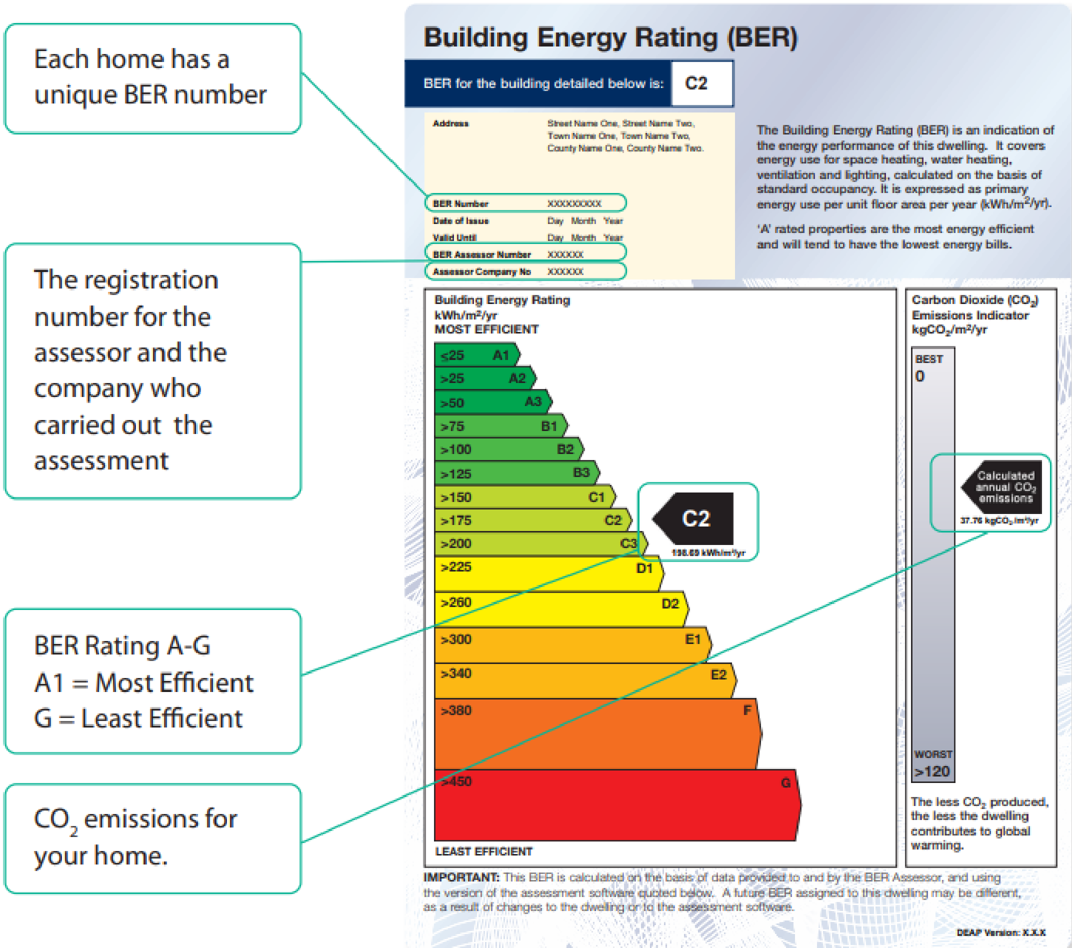

Figure A.1: Example of BER certificate
Source: https://www.seai.ie/publications/BER-dwellingowner-Leaftlet.pdf

Table A.1: Sensitivity of the main results to different thresholds within a thermostat set point temperature

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. var: Duration a boiler operates (minutes) at hour h					
		Temperature thresholds within a set point:				
	0.35^0C	0.40^0C	0.45^0C	0.50^0C	0.55^0C	0.60^0C
BER scales (Reference: B):						
C	3.47***	3.12**	3.21**	3.38**	3.38**	3.81**
	(1.34)	(1.36)	(1.38)	(1.46)	(1.49)	(1.53)
D	4.20**	4.32***	4.59***	4.49**	4.31**	4.82***
	(1.64)	(1.66)	(1.71)	(1.77)	(1.82)	(1.85)
E–G	6.92***	7.45***	7.54***	7.53***	7.00***	7.37***
	(2.09)	(2.14)	(2.20)	(2.27)	(2.32)	(2.33)
Set point temperature at hour h (Reference: $\leq 17^0$C):						
(17^0C, 18^0C]	2.11***	2.15***	2.20***	2.28***	2.38***	2.46***
	(0.44)	(0.47)	(0.46)	(0.49)	(0.50)	(0.51)
(18^0C, 19^0C]	5.35***	5.58***	5.79***	6.00***	6.24***	6.39***
	(0.89)	(0.92)	(0.92)	(0.93)	(0.95)	(0.96)
(19^0C, 20^0C]	9.17***	9.56***	9.83***	10.13***	10.47***	10.77***
	(0.82)	(0.85)	(0.86)	(0.90)	(0.93)	(0.96)
(20^0C, 21^0C]	12.21***	12.81***	13.10***	13.50***	14.02***	14.40***
	(0.90)	(0.93)	(0.94)	(1.00)	(1.04)	(1.08)
(21^0C, 22^0C]	14.56***	15.25***	15.67***	16.06***	16.69***	17.10***
	(0.91)	(0.93)	(0.95)	(0.98)	(1.01)	(1.04)
$> 22^0$C	18.93***	19.69***	20.20***	20.73***	21.40***	21.80***
	(1.19)	(1.21)	(1.24)	(1.29)	(1.28)	(1.31)
Outdoor temperature at hour h (Reference: $\geq 10^0$C):						
$\leq 0^0$C	12.14***	12.75***	12.77***	12.98***	13.11***	13.25***
	(1.00)	(1.02)	(1.02)	(1.05)	(1.06)	(1.06)
(0^0C, 2^0C)	10.43***	10.95***	11.03***	11.13***	11.20***	11.23***
	(0.76)	(0.78)	(0.78)	(0.76)	(0.77)	(0.78)
[2^0C, 4^0C)	8.64***	9.03***	9.06***	9.13***	9.29***	9.33***
	(0.56)	(0.58)	(0.57)	(0.58)	(0.58)	(0.58)
[4^0C, 6^0C)	6.39***	6.66***	6.66***	6.77***	6.83***	6.85***
	(0.42)	(0.43)	(0.42)	(0.43)	(0.42)	(0.42)
[6^0C, 8^0C)	3.98***	4.34***	4.33***	4.37***	4.42***	4.33***
	(0.33)	(0.35)	(0.34)	(0.34)	(0.33)	(0.32)
[8^0C, 10^0C)	2.17***	2.45***	2.41***	2.46***	2.48***	2.46***
	(0.27)	(0.28)	(0.27)	(0.27)	(0.25)	(0.25)
Hour of a day dummies	Yes	Yes	Yes	Yes	Yes	Yes
Day of a year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (within)	0.213	0.214	0.212	0.210	0.209	0.209
Observations	39,739	43,948	46,700	49,149	51,856	53,798
Sample dwellings	299	311	317	320	325	332

Table A.1 shows the sensitivity of the main results in Column 3 of Table 3 to different thresholds within temperature set point (from 0.35°C to 0.60°C). The dependent variable is the duration a boiler operates for heating (in minutes) to maintain the indoor temperature within 0.50°C of the thermostat set point temperature in a given hour. The set point temperature bins are based on the hourly thermostat set point temperature that ranges from 15.5°C – 30°C for the sample dwellings. Similarly, the outdoor temperature bins are based on the hourly outdoor temperature that varies from -5.6°C to +14.1°C during the period of analysis. The other controls included are hourly outdoor relative humidity, wind speed, dwelling area, efficiency of main space heating unit, dummies for fuels of main space heating (gas or oil boiler), dwelling type, number of storeys, and building age band. Robust standard errors clustered at the dwelling level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$