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]

ABSTRACT

This paper examines 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 54% 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.

https://doi.org/10.5547/2160-5890.13.2.tmel

1. INTRODUCTION

This paper examines the extent with which building Energy Performance Certificates (EPCs) reflect observed energy used for heating, by isolating building fabric performance from occupant behavioural effects. Improving energy performance of buildings is commonly cited as an important component of a cost-effective decarbonization strategy, with global investment in building energy efficiency reaching approximately US$184 billion in 2020, predominately in the European Union (UNEP, 2021). Public support is economically justified due to the misalignment between private and social benefits, where energy efficiency investments yield private comfort and savings but also reduce carbon emissions as a public benefit. As exemplified by Ireland's Climate Action Plan 2023 (CAP, 2023), which aims to upgrade 500,000 homes to a 'B2' Building Energy Rating (BER; the Irish EPC standard) by 2030, policy targets

[a] Economic and Social Research Institute, Dublin, Ireland
[b] Department of Economics, Trinity College Dublin, Ireland
[c] SFI MaREI Centre for Energy, Climate and Marine, Ireland
* Corresponding author. E-mail: tensay.hadushmeles@esri.ie

are often defined relative to benchmarks provided by EPCs. Similar EPC policy benchmarks are followed in other countries, including the European Union and the USA.[1]

When national-level policies are set relative to the achievement of certain EPC benchmarks, incentives at the household level adjust accordingly. In Ireland, many supports are available to upgrade a dwelling to at least a 'B2' level on the BER scales (SEAI, 2023a). In the UK, private landlords of rented properties are legally required to have their proprieties a minimum of E-rated as part of building decarbonization (Sayce and Hossain, 2020). EPCs can also serve as tools for identifying areas at risk of fuel poverty (Camboni et al., 2021), assessing eligibility for home grants (Semple and Jenkins, 2020), and calibrating property rental and sale prices (e.g., Brounen and Kok, 2011; Hyland et al., 2013; Fuerst et al., 2015; Stanley et al., 2016; Myers et al., 2022; Cassidy, 2023). These implicitly incentivise 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 investments cast doubts about the accuracy of EPCs in capturing actual energy use, citing the contribution of behavioural influence for a discrepancy (e.g., Levinson, 2016; Fowlie et al., 2018; Davis et al., 2020; Coyne and Denny, 2021). Behavioural elements such as the rebound effect—occupants of energy-efficient homes respond to lower energy costs by using more energy (Gillingham et al., 2020; Aydin et al., 2017; Sorrell et al., 2009)—and potential self-selection biases driven by income, family size, preferences for warmer temperatures, and working from home can mask energy savings differences between efficient and less efficient homes and potentially misalign policy incentives. Apart from behaviour, using standardised or default values (DEAP, 2022) instead of dwelling-specific values in EPCs calculations can introduce noise or bias to the estimated energy performance. Christensen et al. (2023) identifies factors affecting the performance gap between projected and realised energy savings, underscoring biases in model projections and workmanship as major contributors, while rebound effects play a minor role.

Our study investigates the extent with which EPCs capture observed energy used for heating, focusing solely on the effect of building fabric while excluding the effect of occupants' behaviour. We exploit variations in boiler operation for home heating while the indoor temperature is within a small threshold of the thermostat's set point during the main winter heating months. This serves as a proxy measure of the variations in energy use across building energy ratings attributable to building fabric alone, isolating it from occupants' behavioural effects. 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 in the neighborhood of the thermostat set point. We model the relationship between the duration a boiler operates and building energy performance ratings, conditioning on hourly outdoor

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.

temperature, relative humidity, wind speed, levels of thermostat set point, and several building characteristics.

We find a considerable deviation between the gradient of EPC-measured energy performance and that measured ex-post. For our sample, improving the EPC from 'E–G' to B-rated decreases boiler operation by 54%, 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. (2023), who conduct a study to isolate the effects of building fabrics from behavioural factors and report observed energy performance falling short of initial projections. Our results also reveal that factors other than variations in EPCs have a more pronounced influence on energy use for heating, particularly highlighting the thermostat set point temperature's greater effect on boiler runtime. Although thermostat adjustments have the potential to achieve greater reductions in heating demand compared to costly energy-efficiency building upgrades, empirical evidence suggests that households are reluctant to change their actual heating behaviours (Brewer, 2023; Brandon et al., 2022).

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 the paper.

2. DATA

The data for our analysis combines high-frequency smart thermostat data with hourly weather variables and building characteristics data. The smart thermostat data is a panel dataset that provides high-frequency observations on 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. The data is obtained from a 'Hub Controller', an automatic energy manager device with a smart thermostat functionality, which we refer to as a 'smart thermostat', installed in the main living area of the dwelling by Hub Controls Ltd.[2] The data covers a 2-year sample period from the 1st October 2019 to the 30th September 2021 for more than 10,000 dwellings in Ireland built before 2006. The smart thermostats were installed as part of a government programme to improve energy efficiency (i.e., Energy Efficiency Obligation Scheme), which had qualifying criteria for buildings constructed before 2006. For our analysis, we construct relevant variables at an hourly level from the raw smart thermostat data and merge them with building and weather data.

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 anonymised form.[3] Following the introduction of the European Union Energy Performance of Buildings Directive (EPBD) in 2002 (European Commission, 2002), Ireland adopted a mandatory BER certificate, starting in January 2009. By law, all new homes and homes for sale or rent are required to display the building's BER information for prospective buyers and renters in any advertisement media (SEAI, 2023b). The BER certificate is valid for up to 10 years unless a major structural change, like a house extension or a new heating system, that affects the dwelling energy performance is made (SEAI,

2. See https://thehubcontroller.com/ for more information.
3. The Building Energy Rating (BER) data is publicly available at https://ndber.seai.ie/BERResearchTool/ber/search.aspx

2023b). BER assessments are based on the Dwelling Energy Assessment Procedure (DEAP), the official procedure for calculating and rating the energy performance of dwellings (DEAP, 2022), and undertaken by registered assessors who have completed training under the National Framework of Qualifications. The DEAP 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.

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 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. Of the more than 10,000 dwellings in the original sample, only 798 included variables that facilitate matching the smart thermostat data with BER and building characteristic information.

Weather data for individual properties is not available. Instead, we use hourly weather data from Ireland's National Meteorological Service, Met Éireann.[5] The 798 dwellings are from the Dublin region, making the weather station located at Dublin Airport the most relevant. Weather data variables include outdoor temperature, relative humidity, and wind speed. Given that the dwellings in our study are predominantly within 30km of Dublin Airport with broadly similar weather across such a small geographical area, potential measurement errors for dwellings further away from Dublin Airport are likely to be small.

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

4. Note that BER does not include electricity used for running home appliances like cookers, fridges, and washing machines.
5. https://www.met.ie

this approach, we contend that the time a boiler is firing is directly related to the performance of the building fabric and isolates it from occupant behaviour.

We focus on 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) to further ensure that we exclude indoor temperature readings around the set point threshold, particularly when a boiler is solely operating for water heating, not space heating. In the Appendix, we include sensitivity estimates based on data covering the months of November to March inclusive. We can not distinguish whether a boiler is operating for space or water heating, though both are incorporated with the BER assessment. On implementing those conditions, we have 49,149 hourly observations of 'steady-state' operation across 320 dwellings. The key limiting factor that reduces the data from 789 to 320 dwellings is that the minimum data requirement is that a dwelling's heating system is operational for a full hour within ϵ°C of the thermostat set point temperature. This excludes hours when the heating system has been just turned on and internal temperatures are increasing to the set point value. It is customary in many homes to set the heating system to operate for a period in the morning and again in the evening rather than operate continuously or just manually switch on the heating system as required. Consequently, the number of hours per day where the heating system maintains the internal temperature within ϵ°C of the thermostat set point are few, and in many instances, completely excludes some dwellings.

Since the share of each BER scale (e.g., B2, B3, C1, etc.) in the final sample is relatively small, we regroup them into four BER categories: B, C, D, and E–G.[6] With just 320 dwellings, we are not asserting that the sample is representative of the housing stock in Ireland, or even within the Dublin region. However, as illustrated in Table A.1 the dwellings in the sample broadly follow the shares of the national building stock in terms of BER rating, building age, and size, though there is an underrepresentation of properties at the upper tail of the true BER distribution and a tendency for our sample dwellings to be older and smaller in size. We also observe differences in dwelling types and the types of home heating units between the sample dwellings in our analysis and those in the Dublin region. Consequently, the precise estimates of building performance cannot be extrapolated to the wider building stock. Considering that the focus here is to evaluate building fabric performance in isolation from the influence of occupant behaviour, the final sample of dwellings is sufficiently broad to be illustrative of the scale of building fabric performance. While the estimates will illustrate the relationship between building fabric performance and BER under specific conditions (i.e., maintaining set point temperature), they have little practical use due to their reliance on a narrow window of boiler operating time. Even if a truly representative sample of the housing stock were available, the number of minutes a boiler operates per hour to maintain internal temperatures at set point values serves negligible practical or policy purposes. Rather the estimates reveal a better estimate of the real gradient of energy use across the BER scale than that implicit in the BER itself. The 320 sample dwellings across a broad range of building characteristics are substantially greater than several existing studies examining energy efficiency within the Irish housing stock (e.g., Beagon et al., 2018; Rau et al., 2020).

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.

Table 1 presents descriptive statistics of the final 320 sample dwellings across the BER scales. All the dwellings are located in the Dublin region, with an average ex-ante primary energy usage of about 242 kWh/m²/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 m² 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 assessment is 6 years. Considering the variation in the years since BER assessment across the BER scales, we include a year of BER assessment fixed effects to account for potential change in BER assessment over the years, as a robustness check in our analysis. Table A.1 in the Appendix reports the corresponding characteristics of all BER assessed dwellings in the Dublin region, as well as all 967,608 residential properties with a BER assessment as of February 2022.

TABLE 1
Descriptive statistics of sample dwellings across BER scales

	BER scales:				
Variables	All	B	C	D	E–G
BER (in kWh/m²/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 (m²)	99.08	114.08	102.31	97.04	86.34
Area of a living room (m²)	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 heater (%)	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

Note: Table 1 presents average values or percentage of each category.

Table 2 reports summary statistics of the hourly smart thermostat readings and weather variables for the 320 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. 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. Despite using solely weather data from Dublin airport, slight differences in the weather variables (outdoor temperature, relative humidity, and wind speed) across BER scales in Table 2 are reported due to variations in the hours when the restriction of the temperature being within 0.50°C of the set point is satisfied. To account for this, we introduce dummies for hours of the day, days of the year, and years in our regression specifications.

TABLE 2
Summary of hourly smart thermostat readings and weather variables of the sample dwellings across BER scales

	BER scales:				
Variables	All	B	C	D	E–G
Duration a boiler was operating (in minutes)	16.66	14.42	16.11	17.78	20.31
	(11.93)	(10.37)	(12.17)	(11.77)	(12.94)
Thermostat set point temperature (°C)	19.78	18.84	19.93	20.04	20.40
	(2.29)	(2.06)	(2.32)	(2.26)	(2.17)
Average indoor temperature (°C)	19.82	18.87	19.97	20.06	20.40
	(2.27)	(2.03)	(2.30)	(2.23)	(2.16)
Average indoor relative humidity (%)	46.44	46.48	46.04	46.08	48.78
	(7.77)	(7.60)	(7.28)	(8.04)	(8.79)
Average outdoor temperature (°C)	5.50	5.40	5.44	5.54	5.76
	(3.46)	(3.42)	(3.48)	(3.40)	(3.62)
Average outdoor relative humidity (%)	85.12	85.12	85.23	85.03	84.94
	(8.53)	(8.68)	(8.45)	(8.44)	(8.77)
Average wind speed (knot)	10.30	10.14	10.32	10.41	10.22
	(4.56)	(4.55)	(4.58)	(4.56)	(4.42)
Observations	49,149	10,070	20,800	12,965	5,314
Number of dwellings	320	32	138	92	58

Note: Table 2 displays average values with standard deviations in parentheses.

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.

$$Y_{ihdmt} = \beta BER_i + \gamma_k \sum_k Temp_{k,hdmt} + \delta_j \sum_j Setpoint_{j,ihdmt} + \theta' X + \lambda_h + \lambda_d + \lambda_t + U_{ihdmt} \quad (1)$$

Where the outcome variable Y_{ihdmt} 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 runtime 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 effect of BER scales on boiler operation (a proxy for energy use), with B-rated dwellings as a reference. We specify BER_i as a categorical to easily capture any non-linearity effects in a more easily interpretable manner (than for example using a polynomial specification) plus the alpha-numeric BER scale (i.e., B2, B3, C1, etc.) is the standard metric used in policy discussions and in the housing market (i.e., sales and rental) rather than relying on the continuous $kWh/m^2/year$ scale.

$Temp_{k,hdmt}$ 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 for 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 at a given hourly outside temperature depends on the level of the set point temperatures. 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 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 years to control for common hourly or daily routines and other common factors such as changes in fuel prices and policies. 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 effect 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. Column 2 additionally includes controls for the outdoor temperature, temperature set point, and time-fixed effects. In Column 3, we add a vector of other controls for dwelling and main space heating system characteristics. For the sake of space, we have provided the full regression results in Table A.2 in the appendix and robustness checks of the main results in Appendix B.

Our results show that the duration a boiler operates significantly varies across the BER scales, thermostat set points, and outdoor temperatures. The results remain similar when we include controls on characteristics of the dwelling and heating systems, albeit the magnitude

7. Since the shares in the first and last categories are small, we regroup them into temperature bins of ≤0°C and ≥10°C, respectively.

of the estimated coefficients increases. Since the specification in Column 3 better explains the variations 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 dwelling in order to maintain the thermostat set point temperature within 0.50°C 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.[8]

TABLE 3
Effects of building energy performance ratings on heating demand

	(1)	(2)	(3)
Variables	Dependent variable: Duration a boiler operates (in minutes at hour *h*)		
BER scales (Reference: B):			
C	3.69***	2.64*	3.36**
	(1.33)	(1.35)	(1.45)
D	4.49***	3.51**	4.49**
	(1.46)	(1.47)	(1.76)
E–G	4.80***	5.81***	7.73***
	(1.54)	(1.56)	(2.31)
Set point temperature bins	No	Yes	Yes
Outdoor temperature bins	No	Yes	Yes
Additional controls	No	No	Yes
Hours of a day dummies	No	Yes	Yes
Days of a year dummies	No	Yes	Yes
Year dummies	No	Yes	Yes
R^2 (within)	0.00	0.21	0.21
Observations	49,149	49,149	49,149
Sample dwellings	320	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 additional 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 (full results are reported in Table A.2). Robust standard errors clustered at the dwelling level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Even though our primary interest is in estimating the effects of BER scales on boiler operation, it is worth commenting on the estimated coefficients of the set point and outdoor temperatures (see Column 3 of Table A.2). 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

8. This discussion of the results is relative to the reference BER scale (i.e., B-rated dwellings). When we further conduct post-estimation tests to assess the equality of coefficients among the BER scales, the difference between C and D-rated dwellings is not statistically significant (p-value=0.41). However, we observe significant differences when comparing C with E-G-rated dwellings (significant at the 5% level) and D with E-G-rated dwellings (significant at the 10% level), with p-values of 0.02 and 0.05, respectively.

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 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 dwelling.

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 an 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. However, in practice, households may be reluctant to follow such actions (Brewer, 2023; Brandon et al., 2022). 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.

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 of 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.[9] We provide energy and emissions calculations based on these three typical boiler sizes. The emissions factor of natural gas is 0.203 kg/kWh and 0.272 kg/kWh for kerosene heating oil (DEAP, 2022). The estimated coefficients in Table 3 are relative differences in a boiler operation (in minutes), compared to the reference category (B-rated dwellings). The energy use in kWh is obtained by multiplying the relative effects on the duration a boiler operates (in hours) by 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. Compared to B-rated dwellings, the hourly energy use for heating is approximately 1.01–1.68 kWh higher in C-rated dwellings, 1.35–2.25 kWh more in D-rated dwellings, and 2.32–3.87 kWh higher in E–G rated dwellings, depending on the power capacity of the boiler. 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 0.49–0.81 kg in E–G rated dwellings, all relative to B-rated dwellings. It is worth reiterating that these estimates relate to boiler operation during 'steady-state' operation, excluding periods when properties are being heated from a 'cold' state to reach thermostat set point values.

The magnitude of the estimates of the effects of BER is larger than comparable findings by Coyne and Denny (2021) and Meles et al. (2023). The study by Meles et al. (2023) is based on the same underlying dataset used here but with 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. Meles et al. (2023) find 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

9. For example, see https://www.gasworks.ie/

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

	(1)	(2)	(3)	(4)	(5)	(6)
	Energy use in kWh for a boiler:			CO_2 emissions in kg for a boiler:		
Variables	18 kW	24 kW	30 kW	18 kW	24 kW	30 kW
BER scales (Reference: B):						
C	1.01	1.34	1.68	0.21	0.28	0.35
D	1.35	1.80	2.25	0.28	0.38	0.47
E–G	2.32	3.09	3.87	0.49	0.65	0.81
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.

incorporates behavioural 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 behaviour 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.

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 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 54% more for E-rated or below dwellings. This indicates variations in boiler operation along the BER scales are substantially smaller in proportion than that of the projected primary energy demand (kWh/m^2/year). 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 outcomes 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)	Relative change
B	136.35	Reference	14.42	Reference
C	192.68	+41%	17.78	+23%
D	262.14	+92%	18.91	+31%
E–G	386.40	+183%	22.15	+54%
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 3 of Table 3.

5. CONCLUSIONS

Improving building energy efficiency is advocated as one of the most cost-effective approaches to address climate change, with EPCs serving as benchmarks. Ireland's 2023 Climate Action Plan, for instance, aims to upgrade approximately one-quarter of the housing stock to a 'B2' rating by 2030. This paper examines how well residential building EPCs predict observed energy and emissions savings and finds a distinct gradient of building energy performance along the Irish EPC metric (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 fossil energy use and emissions, the key 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 thermostat settings is simple and instantaneous in terms of immediate fuel cost savings, but it may also compromise home comfort, and evidence indicates that households are slow to change their actual heating behaviours. Whereas achieving the B2 standard is expensive, with a move from an E–G to a B rating potentially costing €30–50,000/property (Collins and Curtis, 2017; Moran 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. Second, the difference in energy use attributable to the BER scales is modest and not substantially different in magnitude than minor behavioural interventions. These results do not imply that upgrading a dwelling's energy efficiency is not beneficial. 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 a B2 BER standard.

ACKNOWLEDGMENTS

We are grateful for the funding and support of the Energy Policy Research Centre (EPRC) at the Economic and Social Research Institute. In addition, we would like to thank Hub Controls Ltd. for providing us access to the high-frequency smart thermostat panel data. We are also grateful to the editor, Charles Mason, and two anonymous referees for their helpful comments and suggestions.

References

Aydin, Erdal Nils Kok, and Dirk Brounen. Energy efficiency and household behavior: The rebound effect in the residential sector. *The RAND Journal of Economics*, 48(3):749–782, 2017. https://doi.org/10.1111/1756-2171.12190.

Brandon, Alec, Christopher M Clapp, John A List, Robert D Metcalfe, and Michael Price. The human perils of scaling smart technologies: Evidence from field experiments. National Bureau of Economic Research, Working Paper 30482. https://doi.org/10.3386/w30482.

Brewer, Dylan. Household responses to winter heating costs: Implications for energy pricing policies and demand-side alternatives. *Energy Policy*, 177:113550, 2023. https://doi.org/10.1016/j.enpol.2023.113550.

Camboni, Riccardo, Alberto Corsini, Raffaele Miniaci, and Paola Valbonesi. Mapping fuel poverty risk at the municipal level. A small-scale analysis of Italian Energy Performance Certificate, census and survey data. *Energy Policy*, 155:112324, 2021. https://doi.org/10.1016/j.enpol.2021.112324.

CAP. *Climate Action Plan 2023 (CAP23): Changing Ireland for the Better*. Government of Ireland, 2023. https://assets.gov.ie/256997/b5da0446-8d81-4fb5-991e-65dd807bb257.pdf.

Christensen, Peter, Paul Francisco, Erica Myers, and Mateus Souza. Decomposing the Wedge between Projected and Realized Returns in Energy Efficiency Programs. *The Review of Economics and Statistics*, 105(4):798–817, 2023. https://doi.org/10.1162/rest_a_01087.

Collins, Matthew and John Curtis. Value for money in energy efficiency retrofits in Ireland: grant provider and grant recipients. *Applied Economics*, 49(51):5245–5267, 2017. https://doi.org/10.1080/00036846.2017.1302068.

Collins, Matthew and John Curtis. Bunching of residential building energy performance certificates at threshold values. *Applied Energy*, 211:662–676, 2018. https://doi.org/10.1016/j.apenergy.2017.11.077.

Coyne, Bryan and Eleanor Denny. Mind the energy performance gap: testing the accuracy of building energy performance certificates in Ireland. *Energy Efficiency*, 14(6):1–28, 2021. https://doi.org/10.1007/s12053-021-09960-1.

DEAP. 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/, 2022. Accessed May 20, 2022.

Deschênes, Olivier and Michael Greenstone. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics*, 3(4):152–85, 2011. https://doi.org/10.1257/app.3.4.152.

European Commission. Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex-%3A32002L0091, 2002. Accessed May 16, 2023.

Ge, Qi and Benjamin Ho. Energy use and temperature habituation: Evidence from high frequency thermostat usage data. *Economic Inquiry*, 57(2):1196–1214, 2019. https://doi.org/10.1111/ecin.12744.

Gillingham, Kenneth,David Rapson, and Gernot Wagner. The rebound effect and energy efficiency policy. *Review of Environmental Economics and Policy*, 10(1), 2020. https://doi.org/10.1093/reep/rev017.

Levinson, Arik. How much energy do building energy codes save? Evidence from California houses. *American Economic Review*, 106(10):2867–94, 2016. https://doi.org/10.1257/aer.20150102.

Meles, Tensay Hadush, Niall Farrell, and John Curtis. How well do building energy performance certificates predict heat? *Energy Efficiency*, 16(74), 2023. https://doi.org/10.1007/s12053-023-10146-0.

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

Novan, Kevin, Aaron Smith, and Tianxia Zhou. Residential building codes do save energy: Evidence from hourly smart-meter data. *The Review of Economics and Statistics*, 104(3):483–500, 2022. https://doi.org/10.1162/rest_a_00967.

Sayce, Sarah Louise and Syeda Marjia Hossain. The initial impacts of minimum energy efficiency standards (MEES) in England. *Journal of Property Investment & Finance*, 38(5):435–447, 2020. https://doi.org/10.1108/JPIF-01-2020-0013.

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

SEAI. Building Energy Rating Certificate (BER).Sustainable Energy Authority of Ireland. https://www.seai.ie/home-energy/building-energy-rating-ber/, 2023b. Accessed January 10, 2023.

Semple, Sally and David Jenkins. Variation of energy performance certificate assessments in the European Union. *Energy Policy*, 137:111127, 2020. https://doi.org/10.1016/j.enpol.2019.111127.

Sorrell, Steve, John Dimitropoulos, and Matt Sommerville. Empirical estimates of the direct rebound effect: A review. *Energy policy*, 37(4):1356–1371, 2009. https://doi.org/10.1016/j.enpol.2008.11.026.

UNEP. 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, 2021. Accessed April 29, 2022.

Zivin, Joshua Graff and Kevin Novan. Upgrading efficiency and behavior:Electricity savings from residential weatherization programs. *The Energy Journal*, 37(4), 2016. https://doi.org/10.5547/01956574.37.4.jziv.