Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Measuring What Markets Price: LLM-Based Quality Measures and the Energy Efficiency Premium

Tom Gillespie, Ronan C. Lyons, Hannah Ortega-McCormack

TEP Working Paper No. 1226

June 2026

Abstract:

Energy performance certificates are intended to incentivize residential energy efficiency investment by allowing markets to capitalize lower energy costs into higher property values.

Measuring What Markets Price: LLM-Based Quality Measures and the Energy Efficiency Premium[*]

Tom Gillespie[†] Ronan C. Lyons[‡] Hannah Ortega-McCormack[§]

July 2026

Abstract

Energy performance certificates are intended to incentivize residential energy efficiency investment by allowing markets to capitalize lower energy costs into higher property values. A large empirical literature documents price premiums for energy-efficient dwellings, but these estimates may be biased upward if internal dwelling condition, which is correlated with energy ratings, is not adequately controlled for. We address this measurement problem using a large language model to extract a scalable, consistent measure of internal condition from the full text of 374,000 residential listings in Ireland between 2018 and 2025. Because the approach requires only that properties be described in natural language, it is broadly transferable to other markets. Integrating this measure into a standard hedonic framework, we find that a one-step improvement in Ireland's Building Energy Rating is associated with a price premium of approximately 2.4 percent in conventional specifications. Once internal condition is controlled for, this falls to around 1.5 percent – meaning that models that omit internal quality may be overstating the energy efficiency premium by 60 percent. This attenuation holds across alternative efficiency measures, throughout the price distribution, and in matched transaction-price regressions. These findings suggest that market-based incentives for energy efficiency investment are more limited than conventional estimates imply but also that text-based quality measures derived from LLMs offer a practical tool for improving the credibility of hedonic analysis wherever listing data exist.

Keywords: Energy efficiency, internal condition, housing markets.

JEL codes: C55; Q41; R31.

---

[*]We thank James Carroll (Central Bank of Ireland) for helpful discussions and participants in the Trinity Economics Work-in-Progress series for their comments, questions and enthusiasm. We thank Paul Monahan, Neil Mackay and Conall Broughan at Distilled for their cooperation with all aspects of the dataset and LLM prompt. Any errors are those of the authors and theirs alone.

[†]University of Galway; corresponding author: thomas.gillespie@universityofgalway.ie.

[‡]Department of Economics, Trinity College Dublin; Central Bank of Ireland.

[§]Trinity College Dublin and Central Bank of Ireland.

1 Introduction

Residential buildings account for a large share of energy consumption and carbon emissions in advanced economies, placing housing at the center of efforts to achieve climate and energy targets. In response, policymakers have increasingly relied on information-based instruments – most notably energy performance certificates (EPCs) – to encourage investment in energy efficiency through market mechanisms rather than direct regulation. The underlying premise is straightforward: where more energy-efficient homes command higher prices or rents, private incentives will align with social objectives, supporting decarbonization through decentralized investment decisions. Whether housing markets in practice capitalize energy efficiency into prices – and whether observed premiums genuinely reflect the valuation of energy performance rather than correlated attributes of the dwellings that carry high ratings – is therefore a question of first-order importance for both climate policy and housing economics.

This paper addresses a central concern in the empirical literature on residential energy efficiency: the extent to which estimated price premiums reflect the valuation of energy performance itself, rather than correlated but unobserved aspects of housing quality. Energy ratings are intrinsically linked to dimensions such as interior condition, fixtures and overall upkeep – features that are difficult to observe in standard datasets but strongly capitalized into prices. As a result, conventional hedonic estimates of the energy efficiency premium (EEP) may be viewed with skepticism if these quality dimensions are not adequately controlled for. Our central research question is therefore whether existing estimates overstate the market value of energy performance by failing to account for internal dwelling condition. Using a large dataset of nearly 400,000 residential property listings in Ireland between 2018 and 2025, we combine a standard hedonic framework with a novel, scalable measure of internal condition derived from the full text of property advertisements using a large language model (LLM). This approach allows us to directly assess how controlling for internal quality alters estimated EEPs.

A substantial empirical literature shows that residential energy efficiency is capitalized into housing prices, with more energy-efficient dwellings enjoying price premiums across a wide range of institutional settings. This finding has been documented for both owner-occupied and rental markets and underpins the policy emphasis on energy performance certification as a market-based mechanism to incentivize investments in energy efficiency (Brounen and Kok, 2011; Bruegge et al., 2016; Fuerst et al., 2016; Evangelista et al., 2020; Mudgal et al., 2013; Taruttis and Weber, 2022). Existing evidence for Ireland closely mirrors this international consensus (Hyland et al., 2013; Stanley et al., 2016; Carroll et al., 2024). At the same time, a parallel literature emphasizes that internal dwelling condition and quality exert large and independent effects on housing values, raising the possibility that conventional hedonic estimates may conflate energy performance with correlated but unobserved housing attributes (Brounen and Kok, 2011; Chegut et al., 2016; Miller et al., 2018). A recent review identifies housing quality as the single most important source of omitted variable bias in hedonic estimates of the EEP (Ou et al., 2025), and the most direct evidence – from Belgium and Spain – suggests that controlling for condition materially reduces or in some cases eliminates estimated premiums (Reusens et al., 2026; Marmolejo-Duarte and Chen, 2022). Recent advances in large language models (LLMs) offer a scalable route to addressing this measurement problem: property listings routinely contain rich unstructured descriptions of interior quality, and LLMs are well-suited to extracting consistent, ordinal condition measures from such text at scale (Liu et al., 2020; Nowak and Smith, 2020; Baur et al., 2023). Despite this potential, text-derived quality measures have not, to our knowledge, been deployed to directly reassess the magnitude of the EEP itself.

Our contribution is twofold. Methodologically, we develop a scalable, text-based measure of internal dwelling condition by applying a large language model to the full text of nearly 375,000 property listings. Unlike approaches that rely on structured administrative data or market-specific appraisal records – which are available only in a small number of countries and settings – the LLM-based method requires only that properties be described in natural language, a condition met by listing platforms in many housing markets around the world. This makes it broadly applicable as a tool for recovering an otherwise unobservable but economically important dimension of housing quality. We then use this measure to directly reassess the magnitude of the EEP in the Irish housing market. In baseline specifications that mirror the existing literature, a one-step improvement in BER is associated with a price premium of approximately 2.4 percent. Once internal condition is explicitly accounted for, however, these estimates are revised downward markedly. Across a wide range of specifications – including categorical and continuous measures of energy efficiency, controls for dwelling age, quantile regressions, and matched transaction-price regressions – the implied premium falls to around 1.5 percent per BER step, meaning an overstatement of roughly 60 percent in conventional estimates. Equally striking, internal condition is itself strongly and independently capitalized into prices: the estimated premium for a high-quality relative to a mid-range dwelling exceeds 25 percent, a magnitude comparable to the full range of energy efficiency effects in the baseline model. This attenuation is evident throughout the housing price distribution, suggesting that omitted internal quality inflates estimated EEPs not only on average but across market segments. Together, these findings have direct implications for the interpretation of existing estimates and for the design of energy efficiency policies that rely on market-based incentives.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature, while Section 3 describes the data, including the construction of dwelling characteristics, location controls, energy efficiency measures and the LLM-based internal condition score. Section 4 outlines the empirical strategy and identification framework and Section 5 presents the main results and robustness checks, including comparisons across alternative specifications. Section 6 concludes with implications for the interpretation of EEPs and for housing and climate policy.

2 Literature

A large international literature documents a positive relationship between energy efficiency and housing prices, as reviewed in Ou et al. (2025). Early and influential studies show that dwellings with superior energy labels or ratings sell at a premium, even after controlling for detailed structural and locational characteristics (Brounen and Kok, 2011; Bruegge et al., 2016). Subsequent work across multiple European housing markets confirms the existence of statistically significant EEP, although estimated magnitudes vary across institutional settings, market conditions, tenures, and certification regimes (Fuerst et al., 2016; Evangelista et al., 2020; Taruttis and Weber, 2022; März et al., 2022). A recent synthesis of the European literature suggests that, on average, each improvement in energy performance certification is associated with an EEP in the range of 1–3 percent, while also highlighting substantial heterogeneity across countries and submarkets (Ou et al., 2025). While the broader European evidence generally finds that better EPC ratings are associated with higher housing prices, one notable exception is Aydin et al. (2020), who find capitalization of underlying energy efficiency but no additional premium for the EPC grade itself.[1] Taken together, this body of work establishes the EEP as a robust empirical regularity, while leaving open questions regarding its precise magnitude and interpretation.

Evidence for Ireland aligns closely with this international consensus. Using listings data, a series of studies find that more energy-efficient dwellings command statistically significant price premiums under the Building Energy Rating (BER) system, typically in the range of 2–3 percent per rating improvement (Hyland et al., 2013; Stanley et al., 2016). These estimates are robust across national and urban samples and place Ireland squarely within the broader European evidence base on energy efficiency capitalisation (Mudgal et al., 2013). Experimental evidence further shows that improved energy cost information

[1] This result should be interpreted with some caution in cross-country comparison. Aydin et al. (2020) study the Netherlands, where the paper focuses on the informational role of EPCs in a setting with a voluntary label regime for part of their sample period. Related Dutch evidence suggests that voluntary labels had limited information value, while later mandatory labels had more informational value, implying that institutional context may matter for whether EPC grades are capitalized into prices. This may limit the external validity of the Dutch result for settings where EPC disclosure is more comprehensive or more salient.

affects property investment and retrofit decisions (Carroll et al., 2024), underscoring the role of salience and information in shaping household responses to energy performance.

Recent evidence also suggests that housing markets may respond not only to underlying energy performance itself, but to the way that information is summarized and presented. Using French administrative and transaction data, Lu and Spaenjers (2025) document sharp price discontinuities at energy-label thresholds, along with bunching and evidence of misreporting around cutoff values, consistent with buyers and sellers relying on simplified or heuristic interpretations of energy labels rather than a smooth valuation of continuous energy use. This reinforces the broader theme that observed energy premiums may reflect informational salience and labeling design, not just underlying pure efficiency differences.

A persistent challenge in this literature is that standard hedonic datasets lack direct measures of internal property condition, raising the concern that estimated EEPs may partly reflect unobserved housing quality correlated with energy efficiency rather than energy performance *per se*. Higher energy ratings are frequently associated with newer construction, recent refurbishment, and superior materials and finishes, so quality-related price effects risk being misattributed to energy efficiency (Fuerst et al., 2016; Copiello and Donati, 2021; Bruegge et al., 2016). A recent review identifies housing quality as the single most important source of omitted variable bias in hedonic estimates of EEP, while also noting that it is particularly difficult to measure objectively and consistently across markets (Ou et al., 2025). Where condition variables exist, they are typically obtained indirectly through appraisals, seller assessments, or administrative records; are often coarse or highly collinear with energy ratings; and are available only in a small number of countries and settings (Copiello and Donati, 2021; Evangelista et al., 2020). As a result, the magnitude of omitted quality bias remains largely unresolved.

Recent advances in machine learning and natural language processing offer a scalable and broadly applicable route to addressing this and other measurement problems in real estate economics. Property listings routinely contain rich, unstructured textual descriptions – including of internal condition, through references to finishes, fittings, maintenance and renovation history – that are typically absent from administrative or structured datasets. A growing literature demonstrates that such descriptions contain economically meaningful information about dwelling quality and that incorporating text-derived attributes into hedonic models improves fit and reduces omitted variable bias (Liu et al., 2020; Nowak and Smith, 2020; Baur et al., 2023; Seo et al., 2020; Hebdzyński, 2023). Large language models (LLMs), in particular, offer substantial advantages for this task: their capacity for contextual understanding, nuanced interpretation, and flexible extraction from free-form text makes them well-suited to constructing consistent condition measures at scale from heterogeneous listing data. Despite these possibilities, most existing applications of textual analysis to property markets focus on price prediction or index construction rather than causal interpretation. To the best of our knowledge, none have directly deployed text-derived quality measures to reassess the magnitude of the EEP.

The evidence that emerges when condition *is* adequately measured is instructive. Studies that directly control for internal quality consistently find that it has substantial and economically meaningful effects on prices – often comparable in magnitude to estimated energy efficiency effects – that persist after controlling for structural and neighborhood characteristics (Brounen and Kok, 2011; Chegut et al., 2016; Kain and Quigley, 1970; Miller et al., 2018; Ooi et al., 2014). Using detailed quality controls in Spanish cities, Marmolejo-Duarte and Chen (2022) show that estimated EEPs can disappear entirely once architectural quality (rather than internal condition) is adequately accounted for, while related work finds that energy performance matters primarily in lower-quality or lower-income housing segments (Marmolejo-Duarte and Chen, 2019). Most directly relevant to the present paper, Reusens et al. (2026) show that controlling for reported building condition materially reduces the estimated EEP in Flanders (Belgium), providing the clearest existing evidence that a portion of the observed capitalization of energy performance reflects correlated dwelling quality rather than energy efficiency *per se*. Their approach relies on structured condition information drawn from listings and administrative records in a setting where such data are available. Our contribution is to construct an equivalent measure directly from unstructured listing text using an LLM, making the approach applicable in the far more common setting where condition is not otherwise observed, and demonstrating its potential as a general-purpose tool for resolving omitted quality bias in hedonic analyses of energy efficiency.

3 Data

The analysis focuses on the Republic of Ireland, which comprises the majority of the island of Ireland. Ireland is a small open economy and a member of the European Union, with a population of approximately five million, around 1.5 million of whom reside in the metropolitan area of its capital city, Dublin. The sample period, 2018–2025, spans a phase of sustained housing market tightness characterised by strong demand and persistent supply constraints, alongside significant macroeconomic and policy shifts. These include continued population growth, rising construction costs, Covid-19 lockdowns and increasing policy attention to housing affordability and energy efficiency. While housing market conditions evolved over this period, prices generally exhibited upward pressure across most regions, making it a suitable setting in which to examine how energy performance and internal dwelling quality are reflected in residential prices.

Housing listings data The housing data used in this paper come from a national dataset of residential property listings drawn from `www.daft.ie`, Ireland's leading property website. Daft.ie captures the vast majority of residential listings nationwide, with coverage estimated at over 95% of the market, and has been widely used in academic research on Irish housing markets (Gillespie, Lyons, and Hynes, 2024; Gillespie, Lyons, and McDermott, 2025). The dataset includes just under 374,000 advertised residential sales listings between 2018 and 2025 and provides detailed information on property characteristics, location, and descriptive text.

Key property-level attributes for inclusion in a hedonic housing price model are available at a high rate of coverage. These include dwelling type, number of bedrooms and bathrooms, and precise geographic location. Consistent with Irish market practice, property size is most commonly reported in terms of bedrooms and bathrooms, with internal floor area available for a large subset of listings. The availability of rich unstructured text descriptions for each listing makes the dataset particularly well suited to the analysis of internal dwelling condition using natural language processing techniques.

Energy efficiency measures Since 2013, all property advertisements in Ireland are required to display a Building Energy Rating (BER), providing standardized information on a dwelling's energy performance (Carroll et al., 2024). The BER summarizes the estimated energy use for space and water heating, ventilation, and lighting, expressed in kilowatt-hours per square meter per year. This metric is mapped onto a 15-grade, colour-coded scale ranging from A1 (most efficient) to G (least efficient). Ratings are produced by independent, registered assessors using a standardized software application that incorporates detailed information on the dwelling's physical characteristics, including dimensions, orientation, insulation, and the efficiency of heating and hot water systems. The BER is based solely on these technical attributes and does not reflect household behavior or realized energy consumption; this makes the BER a fixed characteristic of the dwelling, rather than a usage measure, as required by the logic of hedonic regressions. In the majority of property advertisements, only the dwelling's individual BER grade is displayed, rather than the full comparative scale. In the empirical analysis, we exploit this information by measuring energy efficiency using both categorical indicators for BER grades and a continuous index derived from the underlying 15-point scale; for continuous measures, properties that are exempt from having a BER rating or whose BER is unknown are excluded. The BER rating is stated for just over 344,000 listings, with a further 6,550 stating that they are exempt from having a BER (typically due to the property being designated a protected structure). For just over 23,000 listings, the BER rating is unknown.

Dwelling characteristics Dwelling type is captured using a categorical variable constructed from the property type reported in the listing. We exclude sites and classify each observation into one of eight mutually exclusive categories: apartment, bungalow, duplex [2], detached house, end-of-terrace house, semi-detached house (the reference category), terraced house and townhouse. This classification reflects the standard typology of residential properties in the Irish housing market and allows for flexible differences in pricing across fundamentally distinct dwelling forms. In addition to dwelling type, the dataset includes standard structural controls such as floor area and site size (where available), bedroom/bathroom configurations and other listing-based attributes. These dwelling features are included as categorical variables in all specifications, ensuring that comparisons of energy efficiency and internal condition are made within, rather than across, broadly comparable housing forms.

Environmental characteristics and amenities To account for the role of local amenities and environmental characteristics, we construct a rich set of proximity measures capturing access to both natural and built features. For each property, we compute distances to canals, rivers, lakes, golf courses, beaches, cliffs and power lines, as well as areas classified as forest (mixed, deciduous or coniferous), nature reserves and transitional waters. Rather than imposing a linear distance gradient, we discretize each distance measure into a set of categorical indicators based on four distance bands: within 100 meters, 100–250 meters, 250–500 meters, and 500–1,000 meters, with properties located beyond one kilometer serving as the reference category. This approach allows for flexible, non-linear effects of proximity to both amenities and disamenities, accommodating the possibility that very close proximity may have different price implications than moderate distance.

Housing markets and time controls To allow for both level and trend differences across space, we use fixed effects at the level of the Census district (Electoral Division) and year pair. Such fine-grained spatial fixed effects ensure that estimated energy efficiency and condition effects are not driven by unobserved variation in environmental quality or local spatial context. These local market indicators are interacted with calendar year and included as a full set of categorical variables. This structure permits flexible, district-specific time trends, allowing house prices to evolve differently across districts and years in response to local demand and supply conditions. These controls ensure that identification of energy efficiency and internal condition effects comes from within-district, within-year variation, net of both persistent local attributes and housing cycles.

[2] Duplex refers to a single apartment or flat that spans two floors internally, typically connected by a private staircase

Measuring internal condition To construct a measure of internal property condition, we employ a large language model (LLM) to analyze the free-text descriptions in online property advertisements. Each listing is processed individually, and the model assigns a condition score on an ordinal scale from 1 (highest quality) to 5 (poorest condition). The classification is based entirely on the descriptive language used in the advertisement, without reference to photographs or structured attributes. This approach provides a consistent, scalable measure of interior condition that captures the nuance and tone of agent descriptions, which are typically difficult to incorporate into hedonic models.

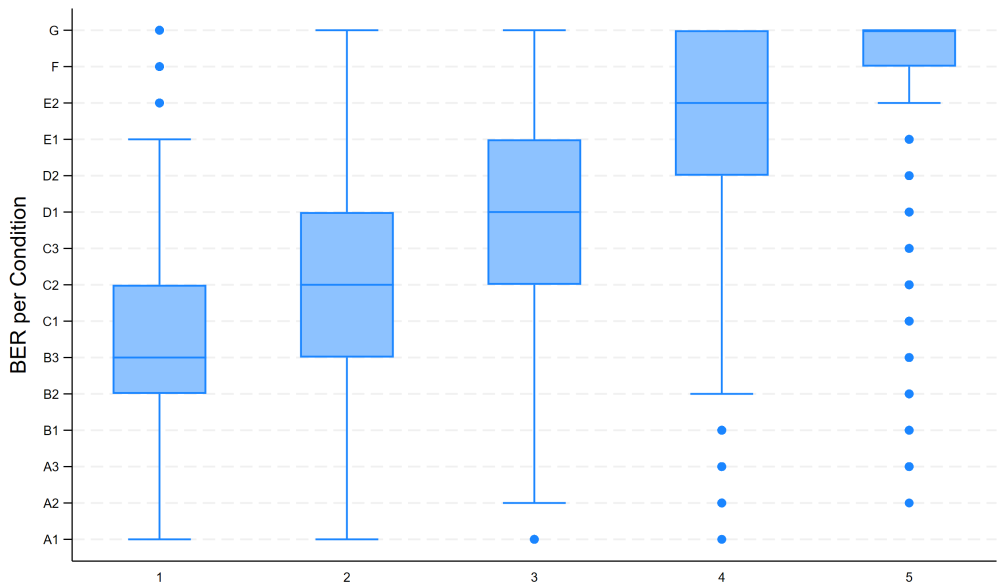
The model identifies linguistic indicators of quality, finish, and maintenance status. Descriptions associated with the highest category often include terms relating to luxury finishes or architectural distinction (e.g. "elegant," "architect-designed," "bespoke interiors," or "premium fixtures"). Mid-range properties are commonly described as "well maintained," "turnkey," or "recently refurbished," while those at the lower end of the scale include phrases such as "in need of modernization," "requiring refurbishment," or "offering potential to improve." The model does not rely on any single keyword but instead evaluates the broader linguistic and contextual framing of the property's condition, as expressed by the listing agent. The LLM used, Claude 3.7 Sonnet, was guided by a structured classification schema to ensure that condition scores are consistent and reproducible across listings.

Of the 374,000 listings, all but 5,820 (1.6%) were able to be assigned an internal condition. The most frequently occurring condition was 3, representing 44.3% of listings, followed by 2 (40.3%). A classification of 5 was the least frequently occurring (1.7%). A simple box-plot of estimated internal condition and energy rating reveals a positive correlation between the two, supporting our main research hypothesis; this is shown in Figure 1.

Other LLM-derived variables In addition to internal condition, the system extracts a small number of ancillary variables from listing text, such as indicators related to construction period or site characteristics. However, the primary variable of interest for this study is the condition category. To mitigate the risk of strategic manipulation of listing language, the precise formulation of the prompt and classification logic are not disclosed. The resulting condition measure serves as a text-derived proxy for internal dwelling quality, enabling an explicit examination of how energy performance interacts with subjective assessments of housing quality in determining residential property prices.

Matched transactions subsample By construction, the listings dataset does not contain the final transaction price. The use of listed prices is nonetheless well established in the hedonic housing literature where transaction prices are unavailable, including in prior

Figure 1: Box-plot of estimated internal condition and energy efficiency



Note: Figure 1 shows the distribution of BER ratings within each LLM-derived internal condition category, where 1 denotes the best condition and 5 the poorest.

work on Ireland (Gillespie, Lyons, and McDermott, 2025; Hyland et al., 2013; Malpezzi, 2003). While listed prices may exceed transaction prices in some settings, such deviations are typically costly for sellers in terms of longer time on the market and lower final sale prices, making persistent mispricing of housing attributes unlikely (Knight, 2002; Merlo and Ortalo-Magné, 2004; Faller et al., 2009). Moreover, listed and transaction prices are generally found to share the same underlying trends, particularly once hedonic controls are applied, including in the Irish housing market during periods of elevated volatility (Lyons, 2019). Existing evidence, in other words, suggests that differences between listing and transaction prices are not systematically related to observed housing characteristics.

Rather than assume no bias, however, we directly assess the sensitivity of our results to the choice of price measure by using a subsample of approximately 170,000 listings that are matched to Ireland's Property Price Register (PPR) using the unique Eircode identifier. The PPR records transaction prices and dates but contains no information on housing characteristics, making the listings data essential for hedonic analysis. The matched transactions sample is not random relative to the full set of listings: successful matches are more likely in urban areas and locations where address information can be precisely aligned across data sources. Nonetheless, this subsample provides a useful robustness check, allowing us to compare estimates obtained using listed prices with those based on final transaction prices. The results of this comparison are presented in Section 5.2.

Table 1: Summary statistics

	Mean	SD	Median	Min	Max	N
List price (€)	323,695	218,729	275,000	30,000	2,000,000	373,943
Log list price	12.51	0.60	12.52	10.31	14.51	373,943
Internal condition	2.61	0.82	3.00	0.00	5.00	373,943
BER band A-G (1-7)	3.67	1.42	3.00	1.00	7.00	344,103
Bedrooms	3.11	0.96	3.00	1.00	5.00	373,943
Bathrooms	2.05	0.97	2.00	1.00	8.00	373,943
Site size (acres)	2.69	23.01	0.75	0.00	3051.75	60,814
Age of Property (years)	41.08	37.37	24.00	1.00	850.00	202,380
Has garden	0.70	0.46	1.00	0.00	1.00	373,943
Sea view	0.05	0.22	0.00	0.00	1.00	373,943
Green view	0.27	0.45	0.00	0.00	1.00	373,943

Note: Table 1 presents summary statistics for the main variables used in the empirical analysis. For Site size and Age of property, a lower N reflects incpmlete coverage in the dataset.

Summary statistics Summary statistics for the main variables used in the analysis are reported in Table 1. The full sample comprises 373,943 residential listings. Listed prices range from 30,000 to 2,000,000, with a mean of 323,695 and a median of 275,000. A BER rating is available for 344,103 listings; the mean BER band of 3.67 on the condensed A–G scale (mapped to 1–7) corresponds to a rating in the C range. The LLM-derived internal condition score has a mean of 2.61 and a median of 3.00 on the scale from 1 to 5, with, as noted, a small number of listings not assigned an internal conditions core. The typical property in the sample has three bedrooms and two bathrooms. Site size and property age are available for subsets of 60,814 and 202,380 listings respectively, reflecting variation in how fully these attributes are reported; where available, the median property age is 24 years, though the mean is 41 years, reflecting the presence of older protected structures in the sample. Approximately 70 percent of listings report a garden, while sea views and green views are recorded for 5 and 27 percent of properties respectively.

4 Empirical Strategy

Our empirical strategy builds on a standard hedonic pricing framework to estimate the EEP in the housing market, extending existing specifications by explicitly controlling for internal dwelling condition derived from unstructured listing text. Throughout, the dependent variable is the logarithm of the listed sale price, except where we switch to transacted prices in a subsample, for robustness.

Baseline hedonic specification For both the full listings sample, and the matched transactions sample, we begin by estimating a baseline hedonic regression that relates log price to energy efficiency, measured using the BER alongside a rich set of control variables. These controls include detailed dwelling characteristics (such as property type, size and age) and proximity-based measures of local amenities and environmental features, as well as Electoral Division (ED) by year fixed effects, which allow housing prices to follow distinct regional trends over time. Standard errors are clustered at the ED level.

In the baseline specification, energy efficiency is measured using the BER system in categorical form, with a full set of indicators for BER grades from A1 to G, as well as exempt and unknown, with one category (the 'D1' rating) omitted as the reference group. This flexible specification allows for non-linear price gradients across energy ratings and mirrors the approach used in much of the existing literature. Econometrically, our baseline is given by the following equation:

$$\ln P_{it} = \beta\,\mathrm{BER}_i + \mathbf{X}_i'\boldsymbol{\delta} + \lambda_{m(i)t} + \varepsilon_{it}, \tag{1}$$

where P_{it} denotes the listed sale price of property i at time t. BER_i captures the dwelling's energy efficiency, as a set of categorical indicators for BER grades. $\mathbf{X}_i$ is a vector of observable dwelling characteristics and proximity-based amenity controls, while $\lambda_{m(i)t}$ denotes district–year fixed effects and ε_{it} is an idiosyncratic error term.

Including internal condition The central extension to the baseline model is the inclusion of an internal condition score derived from a large language model applied to the full text of each listing. This variable enters the hedonic regression as an ordinal categorical control, capturing variation in interior quality and maintenance that is not observed in standard housing datasets. Comparing estimates from specifications with and without this condition measure allows us to assess the extent to which conventional estimates of the EEP reflect unobserved internal quality that is correlated with BER ratings.

$$\ln P_{it} = \beta\,\mathrm{BER}_i + \gamma\,\mathrm{Condition}_i + \mathbf{X}_i'\boldsymbol{\delta} + \mu_{a(i)} + \lambda_{m(i)t} + \varepsilon_{it}, \tag{2}$$

where $\mathrm{Condition}_i$ is the LLM-based internal condition score.

Alternative measures of energy efficiency To examine whether the estimated effects depend on the functional form used to capture energy efficiency, we re-estimate the model using a continuous measure of BER. In these specifications, the categorical BER scale is mapped onto a numeric index where 1 refers to a BER of G (least energy efficient) and 15 refers to A1; here, the sample is restricted to dwellings with non-missing energy ratings as well as excluding the 'exempt' category. While this imposes a linear relationship between adjacent BER categories, it provides a parsimonious summary of the energy efficiency gradient and facilitates direct comparison of effect sizes across specifications.

Robustness to age and floor area Two key robustness check concern the role of dwelling age and floor area (in square meters). In our baseline, we include age (vintage) and exclude square meters (floor area). Here, we outline that our results do not depend on these choices.

We include age in our baseline, because its may mean important dwelling attributes that impact prices, including vintage effects, are omitted. Where these are correlated with energy efficiency, their omission may lead to omitted variable bias (OVB). However, because age is a strong determinant of energy efficiency, its inclusion as a control may mechanically absorb part of the variation associated with BER. Therefore, we re-estimate all main specifications excluding age, while retaining the full set of other dwelling characteristics, location controls and fixed effects. This exercise tests whether the attenuation of the EEP associated with controlling for internal condition persists when age-related variation is not explicitly conditioned out.

A similar issue arises with floor area, measured in square meters. This is excluded in the baseline, as it is a direct input into the BER calculation, and controlling for it risks absorbing part of the variation in energy performance that the BER is intended to capture. Again, we test this in robustness checks, by including square meters.

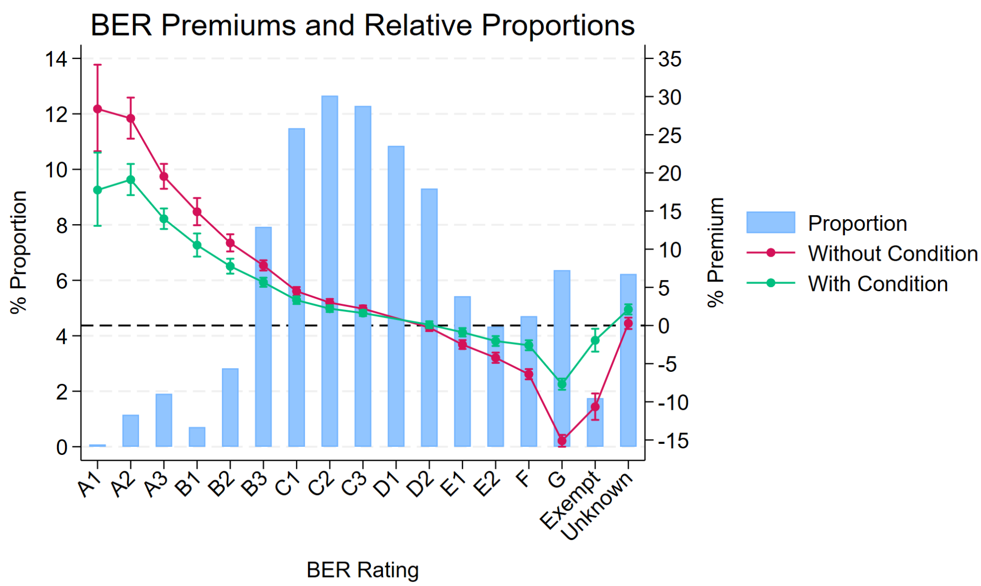
Across all specifications, identification of the EEP and the condition premium comes from within-cell variation in energy ratings and internal condition scores, after absorbing two distinct sources of unobserved heterogeneity. Fixed effects are at the level of the Electoral Division/year pair ($\lambda_{m(i)t}$), capturing both persistent differences in local amenity quality and neighborhood mix that would otherwise confound estimated energy efficiency effects and allowing each district to have its own trend over time. This ensures that the estimated BER and condition coefficients reflect comparisons among dwellings that are observationally similar, located in the same neighborhood and transacting in the same year. The stepwise structure of the specifications then allows us to directly attribute any attenuation of the estimated EEP to the absorption of correlated internal quality, rather than to changes in the sample or control set.

5 Results

Baseline EEP. We begin by estimating the baseline hedonic specification without controlling for internal condition. Consistent with the existing literature, energy efficiency is associated with a statistically and economically significant price premium. Relative to a

D1-rated property, a property rated A is estimated to enjoy a price premium of roughly 12%. Conversely, a G-rated home is estimated to have a listed price approximately 15% below that of an otherwise comparable D1-rated property. These estimates, which follow those reported in earlier studies of the Irish housing market, provide a useful benchmark for subsequent analysis. Figure 2 shows the baseline EEPs (and associated confidence intervals), relative to the reference category D1, in green (on the right-hand axis).[3] The distribution of observations across BER categories is shown by blue bars using the left-hand axis.

Figure 2: Energy efficiency premiums, with and without internal condition



Note: Figure 2 plots estimated energy efficiency premiums relative to the D1 reference category across all BER grades, derived from the baseline hedonic specification (equation 1) and the augmented specification that includes the LLM-based internal condition score (equation 2). Point estimates are shown with 95% confidence intervals on the right-hand axis. Blue bars (left-hand axis) show the share of listings in each BER category, read against the left-hand axis. All specifications include Electoral Division by year fixed effects, and the full set of dwelling characteristics and amenity controls described in Section 3. Standard errors are clustered at the Electoral Division level. The dashed horizontal line denotes zero.

Including internal condition. The same figure also reports estimates from an augmented specification that includes the LLM-based internal condition score. The inclusion of condition leads to a substantial attenuation of the estimated EEP across the full range of BER categories, reflecting the strong positive correlation between the two variables. The absolute value of the coefficients on all energy efficiency labels, relative to the reference category, declines once internal condition is controlled for. In particular, the estimated

[3]The coefficient on the unknown category is similar to the base (and most common) rating. This assuages concerns about sample selection in the reporting of energy labels.

premium for A-rated homes falls from approximately 12% to around 9%, while the estimated discount for G-rated homes narrows from roughly 15% to about 8%. These results strongly support the hypothesis that part of the EEP identified in conventional hedonic models reflects unobserved variation in internal dwelling quality that is correlated with BER ratings.

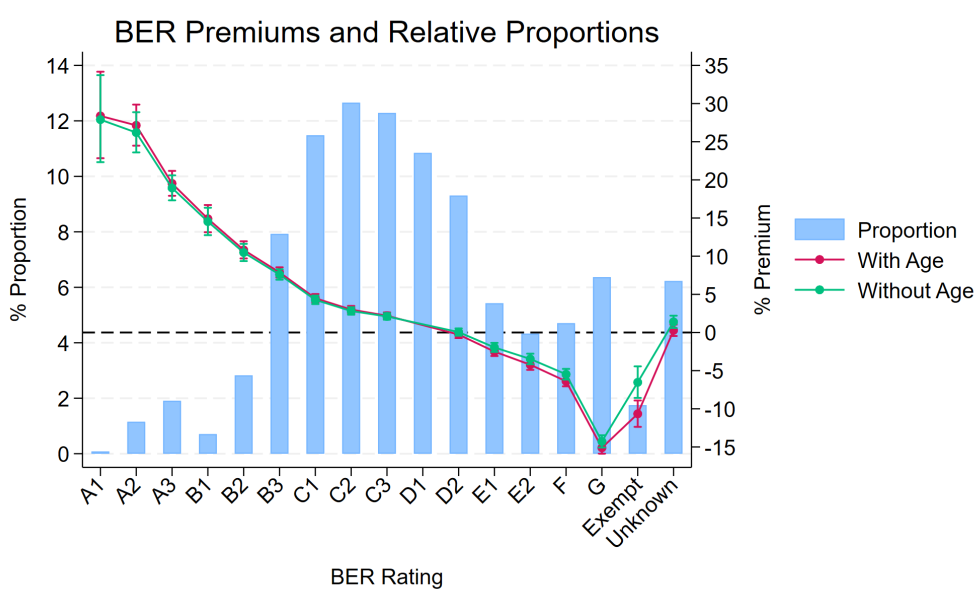
Equally striking, however, is the independent capitalization of condition itself. The coefficient on the condition score is large and precisely estimated: relative to the reference category (Condition $= 4$), a property assessed as having substantially higher internal quality (Condition $= 2$) commands a price premium of nearly 25%, holding energy efficiency and all other characteristics constant. This magnitude is comparable to the price effect of having an A rating compared to having an F rating, reinforcing the view that internal quality is a first-order determinant of residential prices in its own right. That such a large premium was previously unaccounted for in standard hedonic models of the Irish housing market underscores both the severity of the omitted variable problem and the practical value of the LLM-based approach: the condition measure captures a substantial and independently priced dimension of dwelling quality that conventional datasets do not observe.

The attenuation of the EEP is not uniform across the condition distribution. Appendix Figure A1 reports interaction coefficients between the continuous BER measure and each condition category, relative to the $Condition = 4$ base. $Condition = 5$ is not reported, as there properties are derelict and thus will not have an energy rating. For properties in conditions 2, 3 and 4, the BER-price gradient is statistically indistinguishable from each other. For the highest-condition properties ($Condition = 1$), however, the estimated BER slope is significantly flatter — approximately 0.008 lower per BER step — suggesting that energy efficiency is least salient as a pricing factor precisely where internal quality is highest. This is consistent with buyers of premium-condition properties placing relatively greater weight on interior quality and relatively less on energy performance, and reinforces the finding in the Spanish literature that the EEP is concentrated in lower-quality market segments (Marmolejo-Duarte and Chen, 2019).

The role of vintage. One of the larger gaps between baseline and augmented estimates arises for properties that are exempt from BER requirements. These dwellings are protected structures, such as Georgian terraces or Victorian townhouses that are, on average, substantially older than the rest of the housing stock. Their exemption from BER requirements makes them an informative case: because they carry no energy rating, any price differential relative to the D1 reference category in the baseline specification must reflect other correlated attributes. Protected structures disproportionately combine poor energy performance with high internal quality (through, for example, cornicing, stonework or joinery), features that the LLM condition score is well-placed to capture. The larger attenuation for this category therefore reflects the LLM measure doing exactly what it is designed to do: absorbing a dimension of interior quality that is simultaneously distinctive, economically valuable, and entirely absent from standard hedonic datasets.

More generally, this pattern illustrates how the link between energy efficiency, age and internal condition need not run neatly across the housing stock. In older properties, high quality and poor energy performance can coexist, making the case for explicit condition controls all the more important. To assess whether age-related factors affect our conclusions, we re-estimate the main specifications with and without explicit controls for dwelling age. As shown in Figure 3, the inclusion or exclusion of age does not materially alter our main findings. Once internal condition is included, estimated EEPs are very similar regardless of whether age is controlled for. The principal exception concerns BER-exempt properties, for which the inclusion of age results in a somewhat larger estimated discount. Overall, these results support the interpretation that the downward revision of EEPs reflects independent variation in internal quality rather than vintage effects.

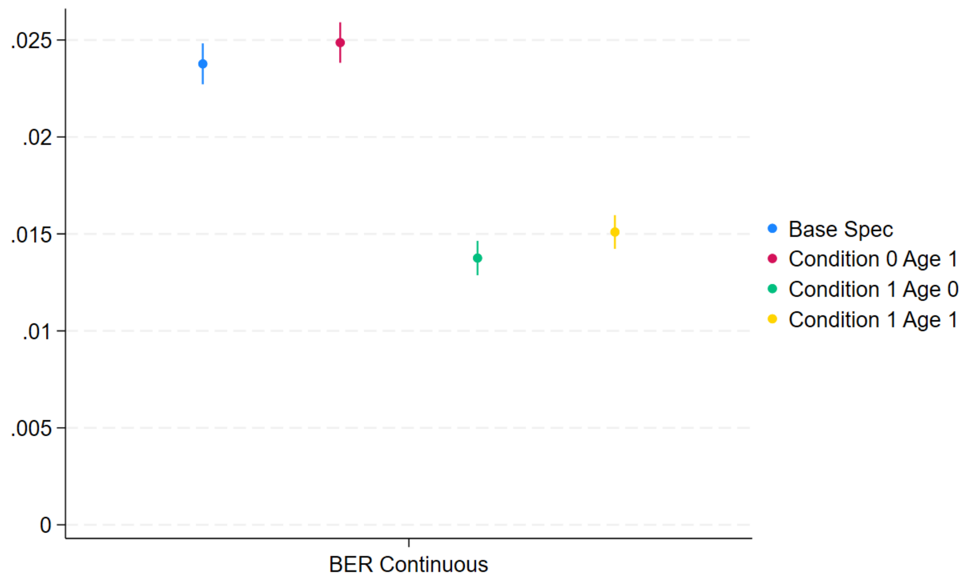
Figure 3: **Energy efficiency premiums, with and without age**



Note: Figure 3 plots estimated energy efficiency premiums relative to the D1 reference category, comparing specifications that include and exclude explicit controls for dwelling age. Age controls comprise LLM-derived period-of-construction categories and regex-based indicators for architectural period (Georgian, Victorian, Edwardian) and associated features. Point estimates are shown with 95% confidence intervals on the right-hand axis. Blue bars (left-hand axis) show the share of listings in each BER category, read against the left-hand axis. All specifications include Electoral Division by year fixed effects, and the full set of dwelling characteristics and amenity controls described in Section 3. Standard errors are clustered at the Electoral Division level. The dashed horizontal line denotes zero.

Continuous measures of energy efficiency. Estimates based on continuous measures of energy efficiency reinforce the conclusion that accounting for internal condition leads to a substantial revision of the estimated EEP. Our primary continuous specification converts the 15-point BER scale into a numeric index ranging from 1 (G) to 15 (A1). Implementing this approach requires restricting the sample to properties with non-missing and non-exempt BER ratings, reducing the sample size from just under 374,000 listings to approximately 330,000 observations.

Figure 4: **Energy efficiency premiums (continuous specifications)**



Note: Figure 4 reports the estimated price effect of a one-step improvement on the 15-point BER scale, where the scale runs from 1 (G, least efficient) to 15 (A1, most efficient). Four specifications are shown: a baseline excluding both age controls and internal condition; a specification including age controls only; a specification including internal condition only; and a full specification including both. Point estimates are shown with 95% confidence intervals. All specifications include Electoral Division by year fixed effects, and the full set of dwelling characteristics and amenity controls. The sample is restricted to listings with a non-missing, non-exempt BER rating (approximately 330,000 observations). Standard errors are clustered at the Electoral Division level.

Figure 4 presents the estimated relationship between a one-step improvement in BER rating on the 15-point scale and the listed price, holding constant a comprehensive set of dwelling characteristics and location controls. Four distinct estimates are shown: a baseline specification excluding both age and internal condition; specifications alternately including age and internal condition; and a full specification including both.

Consistent with the categorical results, including age alone does not materially affect the estimated impact of a one-step improvement in BER rating, which remains at approximately 2.4% – a finding at odds with earlier Irish evidence (Stanley et al., 2016). By contrast, including internal condition leads to a pronounced reduction in the estimated effect, lowering it to around 1.5%. Failing to account for internal condition, in other words, would lead to an overstatement of the EEP by more than half. These estimates are summarized in Table 2, where Condition $= 4$ serves as the reference category.

Table 2: **Energy Efficiency Premium: Continuous Specification**

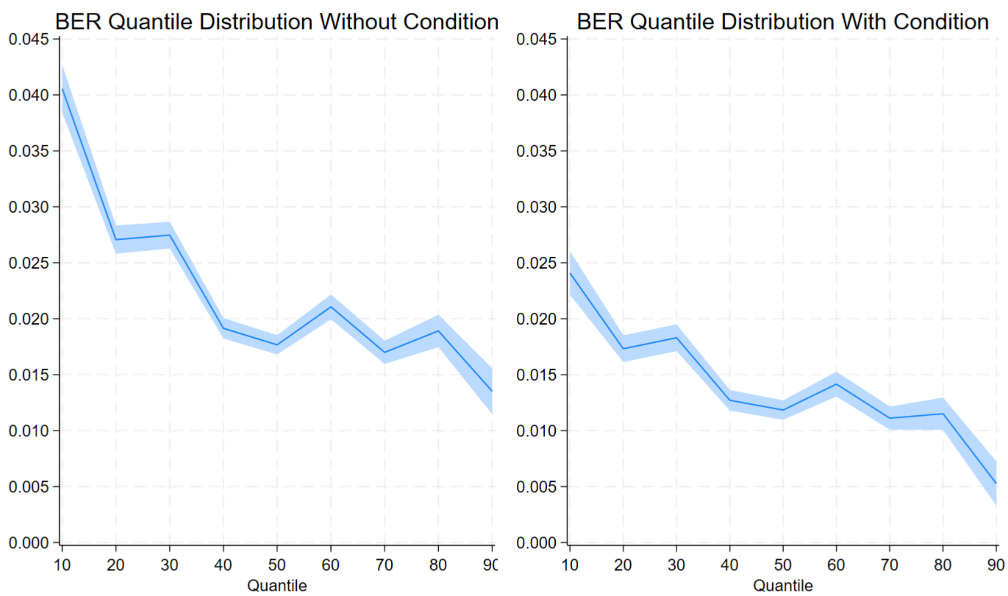
	(1) All	(2) Base	(3) Age	(4) Condition
BER Continuous	0.015	0.024	0.025	0.014
	34	44	47	30
Internal Condition: 1	0.411			0.429
	62			63
Internal Condition: 2	0.218			0.222
	50			50
Internal Condition: 3	0.141			0.141
	38			37
Internal Condition: 4 (Base)	0.0			0.0
	0			0
Internal Condition: 5	-0.264			-0.264
	-29			-29
Constant	12.666	12.823	12.779	12.700
	55	54	54	55
ED Year Fixed Effects	Y	Y	Y	Y
Other Controls	Y	Y	Y	Y
Age (LLM + regex)	Y	N	Y	N
Condition (LLM)	Y	N	N	Y
Observations	330,122	330,122	330,122	330,122
R-squared	0.89476	0.87717	0.88077	0.89207
RMSE	0.193	0.209	0.206	0.196
Absorbed Regressors	22314	22314	22314	22314

Note. The table reports OLS estimates of the price effect of a one-step improvement on the 15-point BER scale, where the scale runs from 1 (G, least efficient) to 15 (A1, most efficient). Column (1) includes both internal condition and age controls; column (2) is the baseline specification excluding both; column (3) includes age controls only; and column (4) includes internal condition only. The internal condition score is measured on an ordinal scale from 1 (highest quality) to 5 (poorest condition), with Condition $= 4$ serving as the reference category. All specifications include Electoral Division by year fixed effects, and the full set of dwelling characteristics and amenity controls described in Section 3. The sample is restricted to listings with a non-missing, non-exempt BER rating. Absorbed regressors refers to the number of fixed effect parameters absorbed by the within-estimator. Standard errors, clustered at the Electoral Division level, are reported below each coefficient. The dependent variable is the natural logarithm of the listed sale price.

5.1 Heterogeneity

A natural question is whether the energy efficiency premium varies systematically across the housing price distribution. If energy efficiency is valued differently in lower- versus higher-priced segments of the market, this would have implications both for interpretation of average EEP estimates and for the distributional effects of energy efficiency policies. To examine this, we estimate unconditional quantile regressions using the continuous BER measure, allowing the marginal price effect of a one-step improvement in energy efficiency to differ across the conditional distribution of prices.

Figure 5: **Heterogeneity in the energy efficiency premium: quantile regressions**



Note: The two panels of Figure 5 plot the estimated coefficient on a one-step improvement in the 15-point BER scale, where the scale runs from 1 (G, least efficient) to 15 (A1, most efficient), across deciles of the unconditional price distribution. The left-hand panel reports estimates from a specification excluding internal condition; the right-hand panel includes the LLM-based internal condition score as an additional control. Shaded areas denote 95% confidence intervals. All specifications include Electoral Division by-year fixed effects, and the full set of dwelling characteristics and amenity controls described in Section 3. The sample is restricted to listings with a non-missing, non-exempt BER rating. Standard errors are clustered at the Electoral Division level.

The left-hand panel of Figure 5 shows that, in specifications that do not control for internal condition, the estimated EEP is strongly decreasing in the price distribution. At the lower end of the market, a one-step improvement in BER is associated with a price premium of approximately 4% at the 10th percentile, falling to around 1.4% at the 90th percentile. This monotonic gradient suggests that energy efficiency is relatively more salient for lower-priced dwellings. While income is unobserved, this is consistent with the hypothesis that energy costs represent a larger share of total housing expenditure for these households.

The right-hand panel shows the corresponding estimates once internal dwelling condition is included. While the qualitative pattern of declining EEPs across quantiles remains, the magnitude of the premium is substantially reduced at every point in the distribution. At the 10th percentile, the estimated EEP falls from roughly 4% to just under 2.5%, while at the 90th percentile it declines from about 1.4% to approximately 0.5%. These results reinforce our headline finding: omitted internal quality inflates estimated energy efficiency premiums not only on average, but throughout the housing price distribution.

A further dimension of heterogeneity concerns the urban/rural composition of the market. Appendix Figure A4 and Figure A3 examine this in more detail. The first plots the estimated BER premium across the CSO's six-category urban/rural typology, with and without internal condition controls. The results reveal a striking gradient: the EEP is approximately 0.7 percent per BER step in cities, rising steadily to around 4.5 percent in highly rural areas. Controlling for internal condition reduces the premium in every category, confirming that omitted quality bias is not confined to any single market type. However, the proportional attenuation is broadly similar across the urban/rural spectrum, so the gradient itself is not an artefact of differential condition correlations – it persists after condition is absorbed.

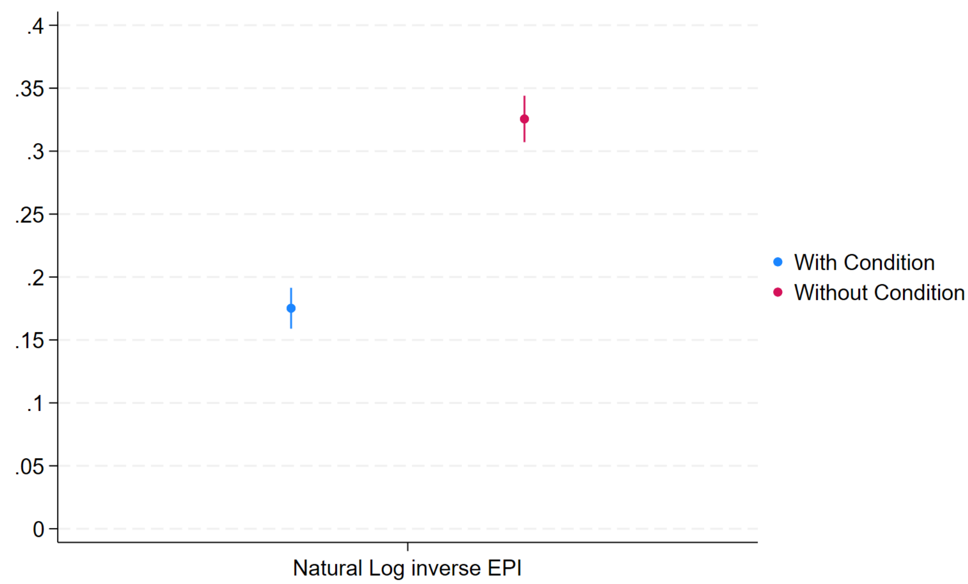
This pattern is consistent with rural buyers placing greater weight on energy running costs, which represent a larger share of total occupancy costs in less energy-efficient older rural stock, and with thinner rural markets offering less price discovery around individual property attributes. The second figure shows a closely related pattern across the five broad regions of Ireland, with Dublin recording the lowest EEPs and Connacht-Ulster the highest.

5.2 Robustness

We perform two additional robustness exercises to assess whether our conclusions depend on the specific measure of energy efficiency used or on the reliance on listed rather than transacted prices. In both cases, the results strongly support the interpretation that failing to control for internal dwelling condition leads to a substantial overstatement of the energy efficiency premium.

Energy Performance Indicator The BER grading system maps a continuous measure of modeled energy consumption—kilowatt-hours per square meter per year—onto a discrete 15-point categorical scale. However, the width of these categories varies considerably, meaning that a one-step improvement does not correspond to a uniform change in underlying energy use. To address this, we re-estimate the core specifications using the underlying Energy Performance Indicator (EPI) directly. We transform the EPI by taking its natural logarithm and rescaling it so that coefficients can be interpreted as elasticities, with higher values corresponding to greater energy efficiency. This allows for a more precise assessment of how prices respond to proportional improvements in modeled energy performance.

Figure 6: **Energy efficiency premiums using the continuous EPI measure**



Note: Figure 6 reports the estimated price effect of a 10% improvement in energy efficiency on the log of the listed price. These are derived from specifications using the natural log of the inverse Energy Performance Indicator (EPI), rescaled so that higher values correspond to greater energy efficiency. Two estimates are shown: a specification without internal condition and a specification with the LLM-based internal condition score. Point estimates are shown with 95% confidence intervals. All specifications include Electoral Division by year fixed effects, and the full set of dwelling characteristics and amenity controls. The sample is restricted to listings with an EPI indicator. Standard errors are clustered at the Electoral Division level.

As shown in Figure 6, the raw specification implies that a 10% improvement in energy efficiency is associated with a price premium of approximately 3.3%. Once internal condition is included, this estimate falls to around 1.8%. The proportional attenuation is at least as large as in the BER-based specifications, suggesting that conventional models substantially overstate the valuation of energy performance when internal quality is omitted.

This finding strengthens our central conclusion by showing that the result is not an artifact of the discretization inherent in the BER label system. It also addresses a related concern raised by Lu and Spaenjers (2025), who show that discrete energy labels can drive heuristic pricing behavior – with house prices responding discontinuously at label thresholds – in ways that may be partly independent of underlying energy performance. Because the EPI measure bypasses the discrete label entirely, the persistence of attenuation in this specification indicates that the omitted quality bias we document is not simply a product of how the label system discretizes continuous energy consumption.

Internal Floor Area Appendix Figure A2 shows that including square meters attenuates the estimated EEP modestly but does not alter the main conclusions. This holds whether or not the sample is restricted observations with non-missing floor area. This confirms that the omission of square meters from the baseline is not driving the core results.

Listed versus Transaction Prices Finally, we examine whether the use of listed prices rather than final transaction prices affects our results. For a large subsample of approximately 170,000 dwellings, we are able to match listings to Ireland's Property Price Register (PPR) using the unique Eircode identifier. This allows us to re-estimate the core continuous-BER specification using transaction prices, as well as to explore the role of timing by aligning prices with either the listing date or the transaction date.

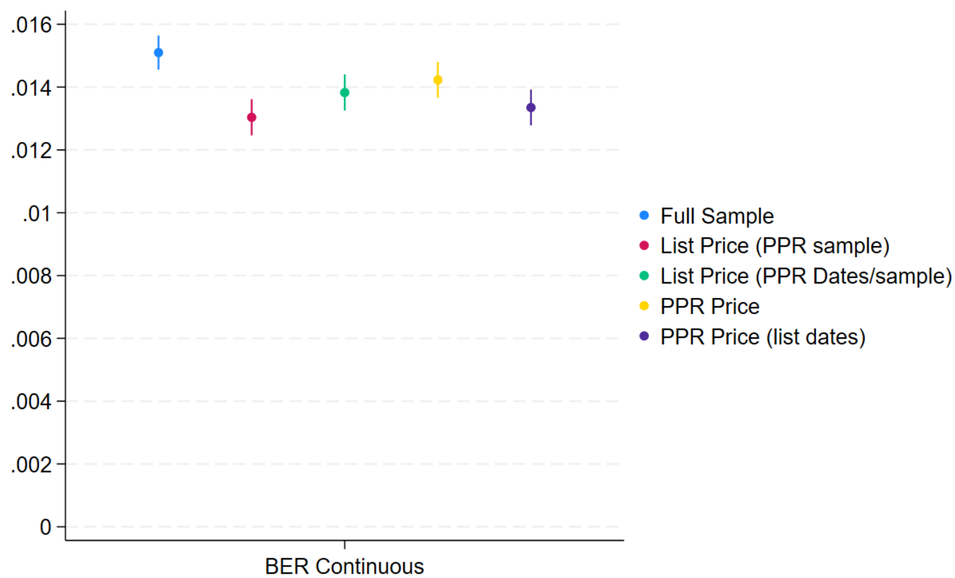
The first estimate in Figure 7 reproduces our baseline result from the full sample: an energy efficiency premium of approximately 1.5% per one-step improvement in the 15-point BER scale. Restricting the sample to matched properties and continuing to use listed prices yields a slightly smaller estimate of around 1.3%, reflecting non-random sample selection, particularly the under-representation of rural properties in the matched data.[4]

Adjusting the timing of listed prices to the transaction date produces only a minor change. When transaction prices are used directly, with prices dated at the transaction date, the estimated premium rises slightly to just over 1.4%. Importantly, the confidence interval for this estimate overlaps with that of the baseline result, indicating no statistically meaningful difference between list-price and transaction-price estimates once sample composition and timing are accounted for. Overall, these results suggest that the choice between listed and transaction prices does not drive our findings.

Summary Taken together, the heterogeneity and robustness analyses provide consistent evidence that internal dwelling condition is strongly capitalized into housing prices and that failing to control for it leads to a substantial overstatement of the energy efficiency premium. This conclusion holds across the housing price distribution, alternative measures of energy efficiency, and both listed and transacted price outcomes. The implied magnitude of omitted variable bias is economically meaningful and consistent across

[4] In the full sample, the EEP for properties in fully rural areas is over twice what it is in the most urban categories of settlement: about 1% for cities, vs. about 2.5% for rural areas.

Figure 7: **Energy efficiency premiums: listed vs. transaction prices**



Note: Figure 7 compares estimated energy efficiency premiums across five specifications that vary the price measure and sample. From left to right: the baseline continuous-BER estimate using listed prices for the full sample; listed prices for the PPR-matched subsample; listed prices for the matched subsample dated to the transaction date; transaction prices from the Property Price Register (PPR) dated to the transaction date; and PPR transaction prices dated to the listing date. All specifications use the continuous BER measure and include internal condition, Electoral Division fixed effects, market–year fixed effects, and the full set of dwelling characteristics and amenity controls. Point estimates are shown with 95% confidence intervals. Standard errors are clustered at the Electoral Division level.

specifications: in the continuous BER specification, controlling for condition reduces the estimated per-step premium from approximately 2.4% to 1.5% – an overstatement of around 60%; using the underlying EPI directly, the equivalent figures are approximately 3.3% and 1.8% per 10% improvement in energy performance, implying a similar degree of upward bias. In the categorical specification, the premium for A-rated properties relative to D1 falls from roughly 12% to 9%, while the discount for G-rated properties narrows from approximately 15% to 8%.

The condition premium is not a second-order correction: the estimated price difference between a high-quality and a mid-range property – nearly 25%, holding energy efficiency constant – is itself comparable in magnitude to the energy efficiency effects documented in the baseline specification. This reinforces the broader conclusion that internal dwelling quality is an independently important determinant of residential prices, and that its omission from conventional hedonic models meaningfully distorts inference about the valuation of energy performance.

6 Conclusion

This paper revisits the estimation of residential energy efficiency premiums by explicitly accounting for internal dwelling condition, a key dimension of housing quality that is typically unobserved in transaction and listings data but strongly capitalized into prices. Using a large sample of Irish residential property listings from 2018 to 2025, we combine a standard hedonic pricing framework with a novel, scalable measure of internal condition derived from the full text of property advertisements using a large language model (LLM). In baseline specifications that replicate the existing literature, we find economically meaningful energy efficiency premiums: highly efficient dwellings command substantial price premiums relative to mid-rated homes, while inefficient dwellings sell at sizable discounts. Once internal condition is introduced, however, these premiums are substantially attenuated. Across a wide range of specifications – including alternative representations of energy efficiency, continuous energy performance indicators, unconditional quantile regressions, controls for dwelling age, and matched transaction-price regressions – we find that the raw EEP associated with a one-step improvement in energy efficiency may be overstated by roughly 60 percent relative to a specification that accounts for internal quality.

At the same time, internal condition itself is shown to be strongly capitalized into prices, with differences in interior quality associated with price effects that are considerably larger than those attributable to energy performance. The attenuation of the EEP is evident not only on average but throughout the housing price distribution, indicating that omitted internal quality inflates estimated energy efficiency premiums across market segments. Taken together, these findings suggest that while housing markets do reward energy efficiency, the magnitude of this reward is materially smaller than implied by conventional hedonic estimates that abstract from internal dwelling condition.

Several limitations of the analysis warrant discussion. First, while the LLM-based condition measure captures rich information embedded in listing text, it remains an indirect proxy for internal quality and reflects how properties are described by agents rather than directly observed physical characteristics. Second, the primary analysis relies on listed prices rather than final transaction prices, although robustness checks using matched transaction data suggest that the main conclusions are not sensitive to this choice. Third, the results are specific to the Irish institutional context, where energy performance certification is standardized, widely disclosed and based on a detailed continuous measure; the magnitude of attenuation may differ in housing markets where coverage is not as high, or where ratings are not as standardized or do not have mandatory disclosure. Finally, as with most hedonic analyses, the estimates reflect equilibrium price differentials across dwellings observed at a point in time rather than the causal effect of upgrading any individual property. A dwelling that receives a higher BER rating through retrofitting will not necessarily realize a price gain equal to the cross-sectional premium we estimate. Retrofitting may involve broader internal works that also improve condition; further, the cross-sectional premium reflects the preferences of buyers who selected into properties at different points in the quality and efficiency distribution, and these preferences may not generalize to the marginal retrofit.

The findings have important implications for policymakers relying on market signals to incentivize energy efficiency improvements. The mechanism we identify is specific: a substantial share of the price premium conventionally attributed to energy performance reflects correlated variation in internal dwelling quality – higher quality tends to accompany higher BER ratings but is independent of energy performance itself. Once this channel is isolated and controlled for, the estimated premium associated with a one-step improvement in energy efficiency falls from around 2.4 percent to 1.5 percent per BER grade. This implies that roughly two-fifths of the market signal on which retrofit incentive arguments are typically based reflects dwelling quality rather than energy performance.

A back-of-envelope calculation illustrates the practical significance of this revision. At a price of €350,000, roughly in line with the median, a one-step BER improvement is associated with a house price premium of around €8,400 in conventional specifications, falling to approximately €5,250 once internal condition is controlled for. The annual energy saving from a one-step improvement at the reference grade – roughly 30 kWh per square meter on a 100 square meter dwelling, at a blended residential energy cost of approximately €0.12 per kWh – amounts to around €360 per year. Capitalized over a 25-year horizon, at a 3 percent discount rate, the present value of that saving is approximately €6,300. The premium, adjusted for internal condition, is therefore broadly consistent with the fundamental financial value of the energy saving, while the conventional estimate overstates it by roughly a third.[5]

The key point is not that the market wildly misprices energy efficiency, but that conventional estimates substantially overstate the degree to which market signals support the private financial case for retrofitting – and that once internal quality is properly accounted for, the implied return from energy improvement alone is modest at best. To the extent that homeowners and policymakers look to observed price premiums as evidence of the private financial case for retrofitting, our results suggest that case is weaker than the existing literature implies. Policies that rely heavily on price capitalization to drive retrofit activity may therefore need to be complemented by other measures, such as targeted subsidies or financing supports.

More broadly, this paper demonstrates that a specific and transferable methodological contribution – using an LLM to extract a scalable, consistent, ordinal measure of internal dwelling quality from unstructured listing text – can materially improve the credibility of hedonic estimates in settings where condition is otherwise unobserved. Unlike approaches that rely on structured administrative data, hand-coded appraisals, or market-specific training sets, the LLM-based method requires only that properties be described in natural language, a feature that is common globally. The Irish case illustrates that even in settings with unusually rich energy performance data, credible estimation of EEPs remains sensitive to how internal quality is measured, suggesting that the scope for applying this approach elsewhere is wide and the potential for revisiting existing estimates substantial. Future research could extend the method by incorporating visual data from listing photographs, examining dynamic responses to retrofitting, or deploying similar tools to recover other dimensions of housing quality that remain difficult to observe in standard datasets. Beyond energy efficiency, the results point to a wider agenda in which large language models help close the gap between the rich qualitative information routinely generated by property markets and the structured data that empirical housing economics has traditionally relied upon.

[5] Under an alternative assumption of perpetual discounting, both estimates imply some degree of under-capitalization, consistent with the broader international literature.

References

Aydin, Erdal, Dirk Brounen, and Nils Kok (2020). "The capitalization of energy efficiency: Evidence from the housing market". In: *Journal of Urban Economics* 117, p. 103243.

Baur, Kilian, Michael Rosenfelder, and Benedikt Lutz (2023). "Automated real estate valuation with machine learning models using property descriptions". In: *Expert Systems with Applications* 213, p. 119147.

Brounen, Dirk and Nils Kok (2011). "On the economics of energy labels in the housing market". In: *Journal of Environmental Economics and Management* 62.2, pp. 166–179.

Bruegge, Christian, Carmen Carrión-Flores, and Jaren C. Pope (2016). "Does the housing market value energy efficient homes? Evidence from the Energy Star program". In: *Regional Science and Urban Economics* 57, pp. 63–76.

Carroll, James, Eleanor Denny, Ronan C. Lyons, and Ivan Petrov (2024). "Better energy cost information changes household property investment decisions: Evidence from a nationwide experiment". In: *Energy Economics* 139, p. 107909.

Chegut, Andrea, Piet Eichholtz, and Rogier Holtermans (2016). "Energy efficiency and economic value in affordable housing". In: *Energy Policy* 97, pp. 39–49.

Copiello, Sergio and Elena Donati (2021). "Is investing in energy efficiency worth it? Evidence for substantial price premiums but limited profitability in the housing sector". In: *Energy and Buildings* 251, p. 111371.

Evangelista, Rita, Esmeralda A. Ramalho, and Joaquim A. e Silva (2020). "On the use of hedonic regression models to measure the effect of energy efficiency on residential property transaction prices". In: *Energy Economics* 86, p. 104699.

Faller, Bernhard, Christoph Helbach, and Alexander Vater (2009). *Möglichkeiten zur Bildung eines Regionalindex Wohnkosten unter Verwendung von Angebotsdaten*. Tech. rep. Research Note 34. Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD).

Fuerst, Franz, Patrick McAllister, Anil Nanda, and Peter Wyatt (2016). "Energy performance ratings and house prices in Wales: An empirical study". In: *Energy Policy* 92, pp. 20–33.

Gillespie, Tom, Ronan C Lyons, and Thomas KJ McDermott (2025). "Estimating the Flood Risk Discount: Evidence From a One-off National Information Shock". In: *Environmental and Resource Economics*, pp. 1–18.

Gillespie, Tom, Ronan C. Lyons, and Stephen Hynes (2024). "The value of a sea view: Hedonic estimates using 3D simulation and natural language processing". In: *Land Economics* 100.2, pp. 278–295.

Hebdzyński, Mateusz (2023). "Quality information gaps in housing listings: Do words mean the same as pictures?" In: *Journal of Housing and the Built Environment* 38.4, pp. 2399–2425.

Hyland, Marie, Ronan C. Lyons, and Seán Lyons (2013). "The value of domestic building energy efficiency: Evidence from Ireland". In: *Energy Economics* 40, pp. 943–952.

Kain, John F. and John M. Quigley (1970). "Measuring the value of housing quality". In: *Journal of the American Statistical Association* 65.330, pp. 532–548.

Knight, John R. (2002). "Listing Price, Time on Market, and Ultimate Selling Price: Causes and Effects of Listing Price Changes". In: *Real Estate Economics* 30.2, pp. 213–237.

Liu, Conghui, Adam D. Nowak, and Peter S. Smith (2020). “Asymmetric or incomplete information about asset values?” In: *Review of Financial Studies* 33.7, pp. 2898–2936.

Lu, Xinyu and Christophe Spaenjers (2025). “Energy labels, misreporting, and heuristic hedonic models”. In: *Real Estate Economics* 53.6, pp. 1200–1222.

Lyons, Ronan C. (2019). “Can List Prices Accurately Capture Housing Price Trends? Insights from Extreme Market Conditions”. In: *Finance Research Letters* 30, pp. 228–232.

Malpezzi, Stephen (2003). *Hedonic Pricing Models: A Selective and Applied Review.* Oxford, UK: Blackwell Science.

Marmolejo-Duarte, Carlos and Ai Chen (2019). “The Uneven Price Impact of Energy Efficiency Ratings on Housing Segments: Implications for Public Policy and Private Markets”. In: *Sustainability* 11.2, p. 372. DOI: 10.3390/su11020372.

— (2022). “The effect of energy performance ratings over residential prices or how an insufficient control of architectural-quality may render spurious conclusions”. In: *Cities* 126, p. 103674.

März, Stefan, Isabel Stelk, and Florian Stelzer (2022). “Are tenants willing to pay for energy efficiency? Evidence from a small-scale spatial analysis in Germany”. In: *Energy Policy* 161, p. 112753.

Merlo, Antonio and François Ortalo-Magné (2004). “Bargaining over Residential Real Estate: Evidence from England”. In: *Journal of Urban Economics* 56.2, pp. 192–216.

Miller, Norman, Vikas Sah, and Michael Sklarz (2018). “Estimating property condition effect on residential property value”. In: *Journal of Real Estate Research* 40.2, pp. 179–198.

Mudgal, Shailendra, Lorcan Lyons, François Cohen, Ronan C. Lyons, and Davide Fedrigo-Fazio (2013). *Energy performance certificates in buildings and their impact on transaction prices and rents in selected EU countries.* Tech. rep. ENER/C3/2010-578. European Commission, Directorate-General for Energy.

Nowak, Adam D. and Peter S. Smith (2020). “Quality-adjusted house price indexes”. In: *American Economic Review: Insights* 2.3, pp. 339–356.

Ooi, Joseph T. L., Tien T. Le, and Neil J. Lee (2014). “The impact of construction quality on house prices”. In: *Journal of Housing Economics* 26, pp. 126–138.

Ou, Yunbei, Nick Bailey, David Philip McArthur, and Qunshan Zhao (2025). “The price premium of residential energy performance certificates: A scoping review of the European literature”. In: *Energy and Buildings* 332, p. 115377.

Reusens, P, F Vastmans, J Vandenbergh, T van Kempen, and S Damen (2026). “Impact of the Flemish energy renovation obligation on house prices”. In: *Economic Review*, pp. 1–44.

Seo, Youngmin, Junghoon Im, and Brian Mikelbank (2020). “Does the written word matter? The role of uncovering and utilizing information from written comments in housing ads”. In: *Journal of Housing Research* 29.2, pp. 133–155.

Stanley, Séamus, Ronan C. Lyons, and Seán Lyons (2016). “The price effect of building energy ratings in the Dublin residential market”. In: *Energy Efficiency* 9.4, pp. 875–885.

Taruttis, Lukas and Christian Weber (2022). “Estimating the impact of energy efficiency on housing prices in Germany: Does regional disparity matter?” In: *Energy Economics* 105, p. 105750.

A Online Appendix

Table A1: **Energy Efficiency Premium: Categorical Specification**

	(1) All	(2) Baseline	(3) +Age	(4) +Condition
BER = 1, A1	0.163	0.246	0.250	0.151
	7.9	11	11	7.1
BER = 2, A2	0.175	0.233	0.240	0.160
	20	22	22	18
BER = 3, A3	0.131	0.173	0.178	0.120
	22	25	26	20
BER = 4, B1	0.100	0.136	0.139	0.093
	14	17	18	13
BER = 5, B2	0.075	0.100	0.103	0.070
	16	19	20	15
BER = 6, B3	0.055	0.073	0.076	0.051
	19	23	24	17
BER = 7, C1	0.033	0.042	0.044	0.030
	13	15	16	12
BER = 8, C2	0.022	0.028	0.030	0.020
	9.8	12	13	8.7
BER = 9, C3	0.016	0.021	0.022	0.015
	8	9.6	10	7.4
BER = 11, D2	0.001	0.001	-0.003	0.004
	.43	.23	-1.1	1.7
BER = 12, E1	-0.009	-0.020	-0.025	-0.003
	-3.1	-6.4	-8.3	-1.2
BER = 13, E2	-0.021	-0.035	-0.043	-0.013
	-6.1	-9.7	-12	-3.7
BER = 14, F	-0.026	-0.056	-0.066	-0.017
	-7.7	-15	-18	-4.7
BER = 15, G	-0.080	-0.154	-0.164	-0.070
	-20	-32	-35	-17

Continued on next page

Table continued from previous page

	(1) All	(2) Baseline	(3) +Age	(4) +Condition
BER = 16, Exempt	-0.020	-0.068	-0.113	0.019
	-2.5	-6	-11	2.1
BER = 17, Unknown	0.021	0.014	0.003	0.030
	6.3	3.4	.74	8.4
Internal Condition LLM = 0	-0.025			-0.029
	-2.7			-3.1
Internal Condition LLM = 1	0.421			0.446
	65			66
Internal Condition LLM = 2	0.232			0.237
	52			52
Internal Condition LLM = 3	0.152			0.153
	41			40
Internal Condition LLM = 5	-0.295			-0.299
	-36			-36
"Period" property	0.124		0.143	
	20		22	
Edwardian	0.081		0.106	
	3.6		4.4	
Victorian	0.059		0.075	
	5.8		6.6	
Georgian	0.051		0.079	
	3.6		5.4	
French Doors	0.017		0.027	
	9.8		14	
High Ceiling	0.055		0.077	
	14		17	
Period of Construction LLM = 0, pre-1400	0.078		0.136	
	.86		1.2	
Period of Construction LLM = 1, 1400–1699	0.326		0.413	
	3.9		4.6	

Continued on next page

Table continued from previous page

	(1)	(2)	(3)	(4)
	All	Baseline	+Age	+Condition
Period of Construction LLM = 2, 1700–1849	0.144		0.170	
	13		14	
Period of Construction LLM = 3, 1850–1899	0.128		0.137	
	15		15	
Period of Construction LLM = 4, 1900–1919	0.059		0.059	
	5.9		5.6	
Period of Construction LLM = 5, 1920s	0.083		0.076	
	6.4		5.7	
Period of Construction LLM = 6, 1930s	0.098		0.088	
	11		8.9	
Period of Construction LLM = 7, 1940s	0.081		0.070	
	6.7		5.7	
Period of Construction LLM = 8, 1950s	0.070		0.048	
	8.6		5.8	
Period of Construction LLM = 9, 1960s	0.051		0.036	
	7.1		4.9	
Period of Construction LLM = 10, 1970s	0.029		0.019	
	5		3	
Period of Construction LLM = 11, 1980s	0.015		0.008	
	2.3		1.1	
Period of Construction LLM = 12, 1990s	0.012		0.010	
	2.9		2	
Period of Construction LLM = 14, 2010s	-0.022		-0.020	
	-3.9		-3.3	
Period of Construction LLM = 15, 2020s	-0.010		0.003	
	-.85		.25	
Period of Construction LLM = 99, Unknown	-0.002		-0.006	
	-.93		-2	
ED Year Fixed Effects	Y	Y	Y	Y
Other Controls	Y	Y	Y	Y

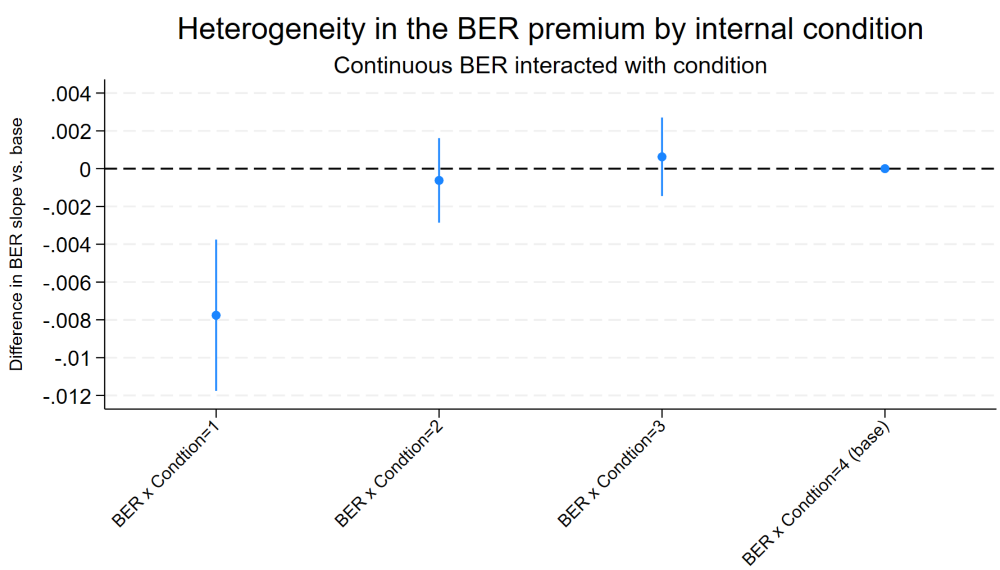
Continued on next page

Table continued from previous page

	(1)	(2)	(3)	(4)
	All	Baseline	+Age	+Condition
Age (LLM + regex)	Y	N	Y	N
Condition (LLM)	Y	N	N	Y
Observations	364,119	364,119	364,119	364,119
R-squared	0.889	0.868	0.872	0.886
RMSE	0.206	0.225	0.221	0.209
Absorbed regressors	22,966	22,966	22,966	22,966

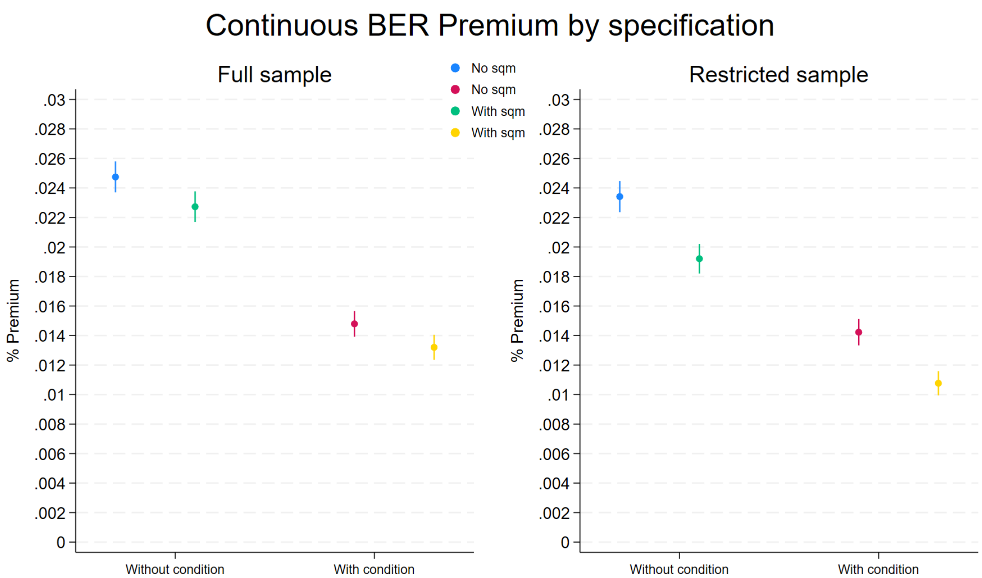
Note. Regression output for the baseline specification, with categorical indicators for each label on the BER scale, and for exempt and unknown BER categories. The reference property is a rating of D1. Standard errors are given below coefficients.

Figure A1: **Heterogeneity in the BER premium by internal condition**



Note: Figure A1 plots the interaction coefficients between continuous BER and the LLM-derived internal condition categories from the hedonic price regression. Coefficients are expressed relative to the omitted base condition category, so each estimate shows how the BER-price gradient differs from that base group rather than the total BER effect for that category. Positive values indicate a steeper BER premium than in the base condition category, while negative values indicate a weaker gradient. All specifications use the continuous BER measure and include internal condition, Electoral Division by year fixed effects, and the full set of dwelling characteristics and amenity controls. Point estimates are shown with 95% confidence intervals. Standard errors are clustered at the Electoral Division level.

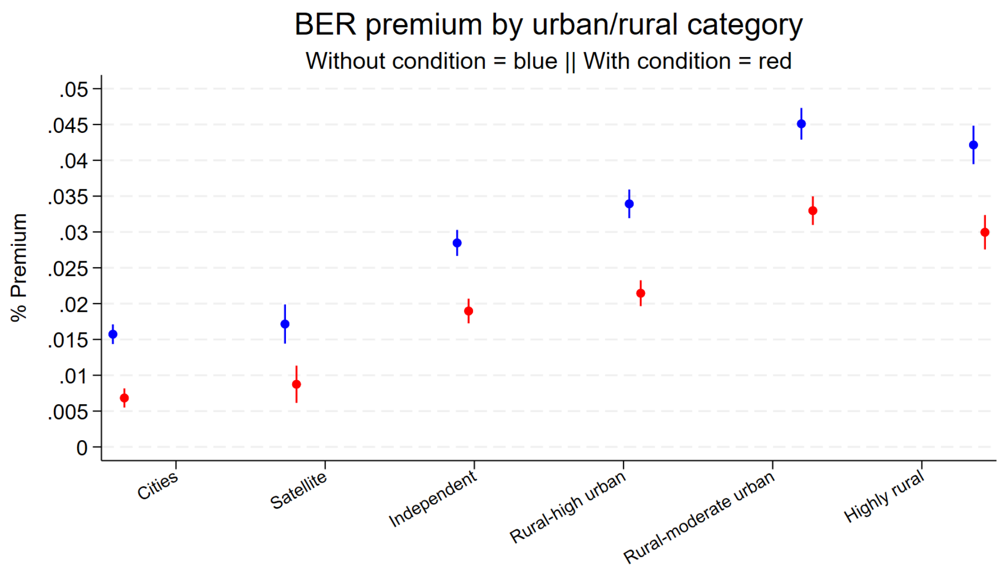
Figure A2: **Heterogeneity in the BER premium by internal condition and the inclusion of square metres**



Note: Figure A2 reports the estimated premium associated with the continuous BER measure across alternative model specifications. The figure compares estimates from the continuous BER sample (N=330,122) and from a restricted sample that excludes observations with missing floor area information (N=256,937). Within each sample, the BER premium is shown both without and with controls for property condition, and each version is estimated either excluding or including square metres controls.

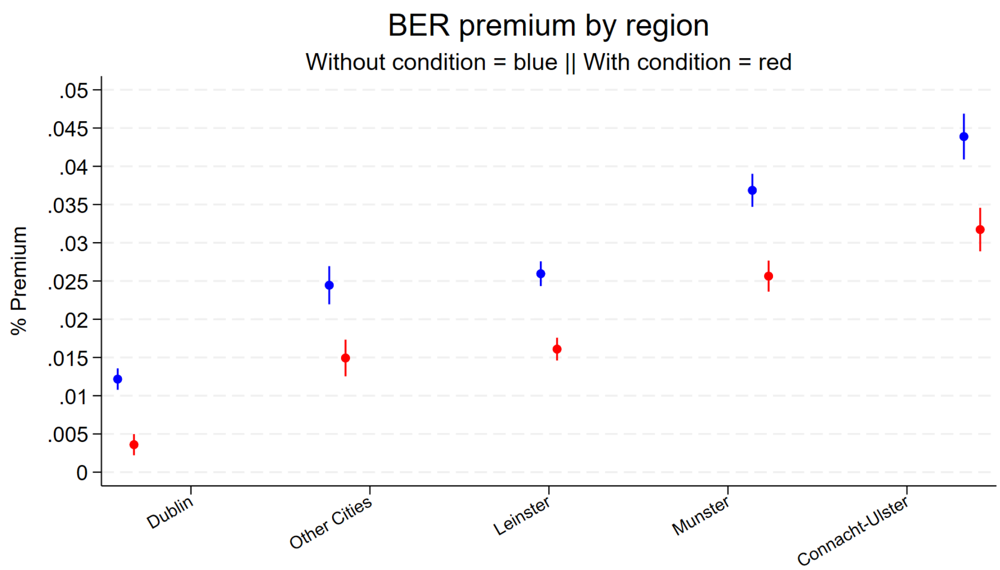
Since BER is calculated using information on the dwelling's energy performance relative to its size, controlling directly for square metres risks absorbing part of the same underlying variation that BER is intended to capture. Adding square metres lowers the estimated premium somewhat, but does not eliminate it, and restricting the sample to observations with non-missing floor area produces very similar patterns. This suggests that missing floor area is not driving the core results, and that the baseline estimates without square metres are not being distorted by selective omission of those observations. Instead, the main conclusions about the positive association between BER and price remain intact across both the full and restricted samples.

Figure A3: **Heterogeneity in the BER premium by internal condition and urban/rural areas**



Note: Figure A3 presents the interaction coefficients between continuous BER and the CSO's six-category urban-rural typology. Under this classification, areas are separated into cities, satellite towns, independent urban towns, and three rural categories defined by differing degrees of urban influence, using population thresholds and commuting-based measures of access to urban centres.
The estimates show a clear relationship: BER premiums are lowest in cities and satellite towns and generally rise as locations become more rural. Controlling for dwelling condition reduces the premium in each category, but the attenuation is broadly similar across areas. This indicates that housing condition explains part of the BER-price relationship everywhere, while a stronger rural premium remains a persistent feature of the data

Figure A4: **Heterogeneity in the BER premium by internal condition and region**



Note: Figure A4 presents the interaction coefficients between continuous BER and the 5 broad regions in Ireland: Dublin, the capital city; Other cities (Cork, Limerick, Galway, Waterford); The rest of the province of Leinster; the rest of the province of Munster; and the combination of the provinces of Connacht and Ulster (excluding Galway city)

The results show a similar relationship to Figure A3: BER premiums are lowest in cities and generally rise as locations become more rural. Controlling for dwelling condition reduces the premium in each region, but the attenuation is broadly similar across them. This indicates that housing condition explains part of the BER-price relationship everywhere, while a stronger rural premium remains a persistent feature of the data