Master

Applied Econometrics and Forecasting

Master's Final Work

Dissertation

Energy Efficiency & Housing Prices: An Econometric Analysis at the Municipality Level

Bernardo del Fiume Sarcinelli Leonel

Supervision:

Dr. Esmeralda Lopes Arranhado

July – 2025

"I don't want to believe.
I want to know."

— Carl Sagan

Preface

Issues regarding sustainability have become an imperative for investors, financing entities, supervisors, governments, and the population in general. As climate change intensifies and resource constraints become more pressing, the demand for energy efficiency and responsible buildings has grown rapidly.

According to the World Economic Forum (WEF) and the United Nations Environment Programme Finance Initiative (UNEP FI), buildings account for roughly 40% of global greenhouse gas emissions, which includes emissions from both the operation of buildings (heating, cooling, etc.) and construction and demolition processes. Therefore, addressing these emissions is crucial for climate action.

The European Union has implemented regulations to improve the energy performance and sustainability of buildings, with a focus on decarbonisation and energy efficiency. The Energy Performance of Buildings Directive (EPBD) is a key piece of legislation driving these changes, setting standards for new and existing buildings, and promoting the use of renewable energy sources. All new buildings must be NZEBs (Nearly-Zero Energy Buildings), meaning they have high energy performance and low energy needs, largely covered by on-site or nearby renewable energy. Individual EU Member States are responsible for implementing the EPBD and developing their own national regulations and incentives to support building renovations, as is the case in Portugal.

What is of great importance is to understand how market prices respond to this growing pressure for "green buildings" and whether buyers are willing to develop (and pay) for more sustainable housing. It is also important to recognise the need to reduce emissions and energy consumption from existing buildings, many of which do not comply with recent regulations. Households with lower incomes are often less willing or able to pay a premium for sustainable buildings or to invest in improvements that would enhance the energy efficiency of their current homes.

For this reason, Bernardo's study is very important in bringing clarity regarding the relationship between energy efficiency certifications and housing prices in Portugal. By providing empirical evidence on how the market values energy-efficient buildings, this research contributes to a better understanding of the economic incentives and barriers that shape the transition towards a more sustainable built environment.

Susana Antunes
ESG Director
Santander Portugal

Lisbon, 28 July 2025

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Professor Esmeralda Arranhado. Working under her supervision has been one of the most enriching academic collaborations I have ever experienced. Her solidarity, sharp analytical mind, and remarkable expertise—reflected in the precision of her words and depth of knowledge—have left a lasting impression on me, and I am truly grateful for the opportunity to have learned from her.

I am also deeply grateful to Susana Antunes, ESG Director at Santander Portugal and my former manager, for her unwavering encouragement and insightful contributions throughout the development of this dissertation. Her guidance, both during my internship and in the final stages of this academic journey, was instrumental. I hold great admiration for her professionalism and leadership, and recognise the immense importance of having had the opportunity to work alongside such an inspiring mentor.

Throughout this Master's degree, I have been intellectually enriched by the faculty at ISEG, whose dedication to teaching consistently motivated me to strive for excellence and deepen my understanding. I am equally grateful to my classmates, family, and friends, whose support—whether through thoughtful discussions, insightful feedback on my dissertation, or their steady encouragement—played a crucial role in making this journey both academically rewarding and personally meaningful.

Abstract

This dissertation analyses the relationship between energy efficiency and housing prices in Portugal, using a panel dataset of 275 municipalities for the period from 2020 to 2023. Applying microeconometric techniques, namely the Poisson Pseudo-Maximum Likelihood (PPML) estimator with high-dimensional fixed effects, the study reveals that a one percentage point (pp) increase in the share of high-efficiency Energy Performance Certificates (EPCs, rated A+, A, or B) is associated with a 0.193% increase in median housing prices per square meter at the municipal level. Control variables, such as population density and the share of employees with higher education, further explain price variations, reflecting socioeconomic influences. The results corroborate previous European studies and highlight the market's valuation of energy-efficient buildings, driven by cost savings and growing environmental awareness. The analysis also underscores Portugal's legislative commitment to the Energy Performance of Buildings Directive (EPBD) and the need for public policies that support low-income households in retrofitting their homes, mitigating energy poverty and promoting housing affordability. This municipal-level approach provides a scalable framework for studying the economic impacts of energy efficiency, contributing to the broader discussion on sustainable development.

Keywords: energy efficiency, housing prices, energy performance certificates, municipal-level analysis, Portugal, microeconometrics, sustainable development.

JEL Codes: R31, Q41, Q56, C23, O18.

Resumo

Esta dissertação analisa a relação entre a eficiência energética e os preços de habitação em Portugal, utilizando um painel de dados de 275 municípios no período de 2020 a 2023. Aplicando técnicas microeconométricas, nomeadamente o estimador de Pseudo-Máxima Verosimilhança de Poisson (PPML) com efeitos fixos de alta dimensão, o estudo conclui que um aumento de um ponto percentual (pp) na proporção de Certificados de Desempenho Energético (CDE) de alta eficiência (classificados como A+, A ou B) está associado a um acréscimo de 0,193% nos preços medianos de habitação por metro quadrado, a nível municipal. Variáveis de controlo, como a densidade populacional e a proporção de trabalhadores com ensino superior, explicam variações adicionais nos preços, refletindo influências socioeconómicas. Os resultados corroboram estudos europeus anteriores e evidenciam a valorização de edifícios energeticamente eficientes pelo mercado, impulsionada pelas poupanças de custos e pela crescente consciência ambiental. A análise sublinha também o compromisso legislativo de Portugal com a Diretiva de Desempenho Energético dos Edifícios (DDEE) e a necessidade de políticas públicas que apoiem os agregados familiares de baixos rendimentos na reabilitação das suas habitações, mitigando a pobreza energética e promovendo a acessibilidade à habitação. Esta abordagem a nível municipal oferece um quadro escalável para estudar os impactos económicos da eficiência energética, contribuindo para a discussão sobre o desenvolvimento sustentável.

Palavras-chave: eficiência energética, preços de habitação, certificados de desempenho energético, análise ao nível municipal, Portugal, microeconometria, desenvolvimento sustentável.

Códigos JEL: R31, Q41, Q56, C23, O18.

Contents

List of Figures

List of Tables

Glossary

ADENE Portuguese Energy Agency.

DGEEC Direção-Geral de Estatísticas da Educação e Ciência.

DGPJ Direção-Geral da Política de Justiça.

EPBD Energy Performance of Buildings Directive.

EPC Energy Performance Certificate.

ESG Environmental, Social, and Governance.

EU European Union.

EU Taxonomy European Union Taxonomy for Sustainable Activities.

FE Fixed Effects.

GDP Gross Domestic Product.

IEFP Instituto do Emprego e Formação Profissional.

INE Portuguese National Statistics Institute.

LASSO Least Absolute Shrinkage and Selection Operator.

NZEB Nearly Zero-Energy Building.

OLS Ordinary Least Squares.

POLS Pooled Ordinary Least Squares.

PPML Poisson Pseudo Maximum Likelihood.

RE Random Effects.

RESET Ramsey's Regression Equation Specification Error Test.

1 Introduction

The relationship between energy efficiency and housing prices has gained increasing importance in recent years due to growing concerns about climate change and the need for sustainable development. Buildings account for around 50% of gas consumption in the European Union (EU), as highlighted by the European Commission (2024), underscoring the sector's substantial role in energy use and emissions. This has driven governments to implement stringent policies aimed at improving building energy efficiency. Within the EU, the Energy Performance of Buildings Directive (EPBD)[1] provides a regulatory framework to promote energy-efficient buildings, requiring member states to implement energy certification systems.

Portugal follows these regulations through the use of Energy Performance Certificates (EPCs), which rate buildings from A+ (most efficient) to G (least efficient). This dissertation investigates whether buildings with higher EPC ratings are associated with higher housing prices at the municipal level in Portugal. Importantly, the analysis focuses exclusively on residential properties. Although both commercial and residential buildings are subject to energy certification, this study concentrates on the housing market for several reasons.

Residential buildings make up the majority of the building stock and play a key role in achieving the EU's climate and energy targets, particularly in reducing household energy consumption and addressing energy poverty. In addition, real estate transactions in the residential sector tend to be more frequent and transparent than in the commercial sector, providing more reliable and granular price data at the municipal level. The structure of the available dataset also aligns more naturally with residential analysis, as both the energy certification records and housing price indicators are better developed for this segment. Focusing on residential properties therefore ensures greater consistency, data quality, and policy relevance.

This research question is relevant for several reasons. Firstly, it helps determine whether the real estate market recognises and rewards energy-efficient buildings, which could encourage homeowners and investors to prioritise energy-saving improvements. Secondly, it contributes to evaluating energy certification programs as effective and credible instruments for promoting long-term environmental and economic sustainability. Finally, by conducting the analysis at the municipal level, the study captures local variation that may be obscured in broader national-level studies, allowing for a more detailed and nuanced understanding of differences between regions and urban structures.

[1]The Energy Performance of Buildings Directive (EPBD), currently based on Directive 2010/31/EU as amended by Directive (EU) 2018/844. A further revision was adopted in 2024 as part of the EU Green Deal, but its implementation is still ongoing.

A growing body of literature has analysed the link between energy efficiency and property values, with generally positive findings across different countries and contexts. Fuerst and McAllister (2011) found that homes with better EPC ratings consistently sold at a premium in the UK market. Similarly, Brounen and Kok (2011) observed notable price premiums for energy-efficient homes in the Netherlands, though the size of the premium varied significantly by market conditions. Research by Kok and Jennen (2012) also strongly supports the view that energy performance positively influences housing prices, though the effect interacts with other important property and neighbourhood characteristics.

Many existing studies rely on transaction-level data to estimate hedonic price functions that relate the value of individual dwellings to their structural and locational attributes. However, such data are often subject to confidentiality constraints and are not publicly available in many contexts — including Portugal, a limitation highlighted by Evangelista, Ramalho, and Silva (2020). This dissertation adopts a different approach, using municipality-level median housing prices and regional characteristics to explore the association between energy efficiency and market values. While this precludes a hedonic analysis in the traditional sense, it allows for the examination of broader patterns using aggregated indicators and avoids the limitations associated with micro-data access.

The empirical strategy is grounded in microeconometric techniques applied to panel data. The models estimate the association between the share of high-efficiency EPCs and median housing sales prices at the municipal level, while controlling for key socioeconomic and infrastructural factors such as average wages, educational attainment, population density, and local economic activity. Using year and municipality FE, the approach controls for unobserved, time-invariant characteristics and national-level shocks. At this stage of the research, classical and widely established log-linear models appear appropriate given the nature of the price variable, although exponential forms may also be considered to assess robustness and interpretability. The analysis will explore potential heterogeneity in effects across municipalities with different demographic and economic profiles, considering variations in local amenities and infrastructure.

This distinction in the modelling framework is especially relevant and important to highlight. Since the currently available data do not include detailed property-level information, this dissertation therefore does not employ a hedonic pricing model. Instead, the analysis focuses on capturing associations between regional characteristics — such as the prevalence of energy-efficient buildings — and the typical housing prices observed within each municipality. The main goal here is not to estimate the marginal effect of individual property features, but rather to assess whether higher levels of energy efficiency at the local municipal level tends to correspond with higher market values over time.

The dataset used in this study is primarily sourced from Portuguese Energy Agency (ADENE) and complemented with municipal-level information from Portuguese National Statistics Institute (INE) and other official entities. The ADENE dataset includes monthly data on EPCs issued by municipalities from 2014 to 2025, aggregated into annual series from 2020 to 2023 for analytical consistency. Key variables include the share of high-efficiency certificates (rated A, A+, or B) among all residential EPCs — a transformation developed by the author — and the median square meter price per municipality. The properties under analysis are largely consistent across variables, due to the mandatory nature of EPCs for new and renovated properties, expiration rules, and alignment of data collection periods. Controls include the share of the employed population with higher education, firms and schools per capita, population density, and housing stock characteristics.

This dissertation contributes to the existing literature in several ways. Firstly, it provides empirical evidence from Portugal, a context under-explored compared to Northern and Western European countries. Secondly, the use of a rich panel dataset — originally collected monthly but aggregated annually — allows the study to capture long-term dynamics and explore structural shifts in the housing market. Thirdly, the construction of original indicators, such as the EPC share and the integration of education and employment variables, adds analytical dimensions to the relationship between energy efficiency and housing prices.

Moreover, this research aligns with broader European policy efforts, including the European Union Taxonomy for Sustainable Activities (EU Taxonomy)[2], which emphasises decarbonising buildings to mitigate climate change. Understanding how energy efficiency influences market prices contributes to evaluating the economic incentives of environmental regulation, and may inform future policy design to encourage investment in sustainable construction and renovation.

The dissertation is structured as follows: Section 2 reviews the literature on the impact of energy efficiency on housing prices, highlighting key methodological approaches and empirical findings. Section 3 provides institutional background on the Energy Performance Certificate (EPC) system in Portugal and Europe, including the regulatory framework established by EU directives. Section 4 outlines the econometric framework, including the use of panel data models with log-linear and exponential functional forms to explore the relationship between energy efficiency and housing prices. Section 5 presents the dataset used in the analysis, along with the construction of key variables. Section 6 discusses empirical results and robustness checks. Finally, Section 7 summarises the main insights and proposes directions for future research.

[2]Regulation (EU) 2020/852 of the European Parliament and of the Council on the establishment of a framework to facilitate sustainable investment, also known as the EU Taxonomy Regulation.

2 Literature Review

In recent decades, there has been a growing interest in the intersection between energy efficiency and housing economics, particularly in the context of residential buildings. The use of EPCs has become increasingly widespread in the EU, driven by regulatory requirements and the push toward decarbonisation of the built environment. These certificates are now a standard requirement in property transactions across most European countries. In parallel, microeconometric methods have been increasingly applied to investigate the relationship between energy efficiency and property values, often using detailed datasets at the household or municipal level. Recent research on the topic has followed two complementary lines: one focusing on the economic impact of EPCs on housing markets, and another exploring methodological approaches in environmental and housing economics.

2.1 EPCs and the Value of Energy Efficiency in Residential Housing

EPCs provide an official assessment of a property's energy efficiency and are a central feature of EU climate policy. Several studies have assessed whether residential buildings with higher energy efficiency ratings determine price premiums in the housing market. Broadly, literature suggests a positive effect of EPC ratings on housing prices, although magnitudes vary by country and local market conditions. One reason for this variation lies in the differences in consumer awareness, energy costs, and trust in the certification systems.

For instance, Ou et al. (2025) conduct a scoping review of 68 peer-reviewed studies across Europe to evaluate whether EPCs are reflected in housing prices. Their findings confirm that most studies identify a positive price premium for energy-efficient homes, although the size of the premium varies across regions and housing segments. These variations appear more pronounced in competitive, high-demand markets. While hedonic pricing models dominate the literature, common challenges include omitted variable bias and limited property-level data. The authors stress the need for stronger EPC transparency, harmonisation across EU countries, and retrofit policies adapted to local contexts.

Similarly, Koengkan and Fuinhas (2022) analyse EPC effects on housing prices across 289 Portuguese municipalities from 2014 to 2019. Using panel data and quantile regressions, they find that energy-efficient dwellings tend to sell at a premium, while lower-rated properties are associated with discounts. Their analysis also highlights that local factors—such as municipal GDP, energy efficiency incentives, and housing supply—play a role in shaping transaction values. These findings reinforce the importance of considering both certification levels and the broader economic and policy environment when interpreting EPC impacts on pricing.

On a broader scale, Huang and Du (2021) analyse how air pollution influences investor behaviour and land valuation in the Chinese land market. Using transaction-level data and an instrumental variable strategy based on the Huai River policy[3], they find that higher pollution levels reduce land bidding prices, especially for high-value parcels. While the study does not focus on carbon emissions directly, it demonstrates that environmental quality—through investor perception and cognitive responses—can be capitalised into property values, reinforcing the role of non-financial factors in shaping housing markets.

Although a positive effect of EPC ratings on property values is widely reported, concerns remain regarding potential selection bias, omitted variable bias, and measurement errors. These issues can affect the robustness of the estimated relationships and require careful and appropriate model specification. Additionally, the extent to which energy efficiency is capitalised into housing prices may be influenced by factors such as buyers' awareness, local market conditions, and the overall credibility of the energy certification system.

2.2 Microeconometric Methods in Housing Market Analysis

Microeconometric models have become essential for estimating the impact of energy efficiency on housing prices. This approach encompasses a range of techniques potentially suitable for dealing with administrative datasets or repeated observations over time, allowing researchers to control for unobserved heterogeneity and account for dynamic effects. Moreover, it includes hedonic pricing models that relate property characteristics to observed market values, offering a flexible framework for empirical evaluation.

Hedonic price models, first formalised by Rosen (1974), remain a standard approach to estimate how property attributes—including EPC ratings—affect market prices. In these models, the log of house price is regressed on a vector of structural and locational characteristics. However, the basic hedonic model assumes functional form and no omitted variable bias. To address this, researchers have increasingly turned to Fixed Effects (FE) models, instrumental variable techniques, and semi-parametric specifications.

Kuethe and Coggeshall (2010) use panel data models with time and spatial fixed effects to account for persistent unobserved heterogeneity in property-level data. Their work demonstrates that including fixed effects significantly reduces upward bias in the estimated premium for energy-efficient homes. In a similar spirit, Ayala, Galarraga, and Spadaro (2016) employ a spatial fixed effects panel model using

[3]The Huai River policy refers to a Chinese government regulation that historically provided winter heating to cities north of the Huai River but not to the south, creating a natural experiment for air pollution exposure.

regional data from Italy to estimate the effects of green housing incentives. Their approach accounts for policy heterogeneity across provinces and confirms the robustness of the estimated premium.

Wooldridge (2010) emphasises the importance and usefulness of panel data estimators in effectively controlling for unobserved heterogeneity in empirical research. By exploiting within-unit variation over time, fixed effects estimators are capable of eliminating time-invariant omitted variables, thereby providing more credible and reliable estimates of causal relationships. This methodological approach is especially relevant for studies using repeated sales or administrative records tracking properties across time. Moreover, panel data techniques help account for structural changes, local policy shifts, macroeconomic shocks, or other time-varying influences that could potentially affect housing prices beyond individual property characteristics.

Furthermore, recent advancements in high-dimensional econometrics and causal machine learning have progressively been incorporated into housing economics. The work of Belloni, Chernozhukov, and Hansen (2014) regarding the application of post-double selection techniques using the LASSO estimator introduces a rigorous framework for conducting variable selection procedures within hedonic price regressions that involve an extensive array of control variables, minimising overfitting while maintaining interpretability. This is especially useful when dealing with rich real estate datasets containing numerous interaction terms and spatial indicators. As real estate datasets become increasingly complex, these tools offer practical ways to handle dimensionality while preserving transparency.

Beyond the conventional framework of hedonic pricing models, quantile regressions have been used to explore distributional effects in more detail. Yoshida and Sugiura (2016) examine whether the observed premium associated with superior EPC ratings remains uniform across the entire housing price distribution spectrum. Their findings reveal that the effect is notably stronger in the upper quantiles, suggesting that energy efficiency is valued more highly in high-end and luxury market segments. These results reinforce the idea that willingness to pay for energy-efficient attributes is not homogeneous across buyers but rather varies substantially depending on household characteristics, individual consumer preferences, and the specific housing segment being considered.

Although greenwashing—a phenomenon where organisations misleadingly market themselves as environmentally responsible—is more frequently studied in corporate finance, some parallels can be drawn in housing markets. For example, Cajias and Bienert (2014) discuss the role of certification credibility, suggesting that if EPCs are perceived as weak or unreliable, their impact on housing prices may be limited. This points to the need for strong institutional frameworks and public trust. Regulatory quality and enforcement mechanisms are crucial to maintaining the signalling power of energy labels in real estate markets.

Another important contribution to the literature comes from Evangelista, Ramalho, and Silva (2020), who analyse over 250,000 residential transactions in Portugal. Their study finds that homes with high EPC ratings (A or B) generally sell at higher prices, with a pronounced premium for apartments—reaching up to approximately 13%—and smaller premiums for houses, typically between 5% and 6%. The paper also examines differences across building types and regions, providing valuable insights into the Portuguese housing market. By using micro-level data and controlling for various property features, the authors offer robust evidence of how energy efficiency influences property values. These findings are particularly relevant for studies focusing on Portugal, where energy certification has gained increasing prominence in recent years.

While most of the literature relies on micro-level data with individual property transactions, this methodological approach is not always feasible. Several challenges arise, particularly due to stringent confidentiality restrictions, legal privacy concerns, or the inherent difficulty in gaining access to sufficiently detailed and disaggregated datasets. In the context of this dissertation, rather than using individual-level data, the author analyses the relationship between energy efficiency and housing prices using municipality-level median prices and regional characteristics. This methodological choice allows the study to effectively circumvent the limitations associated with privacy regulations and examine broader patterns across different local contexts. Although this represents a more aggregated level of analysis, it still provides valuable insights into how regional characteristics relate to energy efficiency and property values.

It is also important to note that, in contrast to the traditional framework of hedonic models that use property-specific attributes, the present study does not estimate a hedonic price function. Instead, it establishes a relationship between housing prices and a set of broader contextual factors such as population density, educational attainment, and the share of energy-efficient dwellings in each municipality. Therefore, the analytical approach adopted here focuses more on structural regional determinants than on the valuation of specific housing attributes. While this choice limits the granularity of the analysis, it provides a complementary perspective that is especially relevant for policy discussions at the municipal or regional level.

The growing availability of EPC data across Europe, combined with the rising policy emphasis on achieving sustainability goals, creates new and significant opportunities for further research on how such certifications interact with local housing markets and broader environmental objectives. As climate targets continue to become more ambitious, understanding the dynamic interplay between real estate prices and regulatory instruments such as EPCs becomes essential. The next section offers a closer look at the institutional framework that shapes the EPC system in the EU and Portugal, setting the stage for the empirical strategy adopted in this dissertation.

3 Institutional Background: The EPC System in Europe and Portugal

Energy performance in buildings has emerged as a fundamental and pressing issue within the global response to the challenges posed by climate change and the pursuit of sustainable development goals. In the EU, the building sector stands out as one of the largest and most significant consumers of energy, being directly responsible for nearly 40% of total final energy use and approximately 36% of greenhouse gas (GHG) emissions, as emphasised by the European Commission (2024). This high share reflects not only the large number of buildings but also the typically long life cycle of construction, which makes the sector slow to adapt to newer energy standards. As such, improving the energy performance of residential and commercial buildings has become a central pillar of the EU's energy and climate agenda.

To address this issue, the EU has developed and implemented a robust regulatory framework that is designed to enhance transparency and energy efficiency within the real estate sector. Among the various policy instruments introduced under this framework, one of the most impactful and widely recognised measures is the EPC, a label designed to inform buyers and tenants about the energy characteristics of a property. EPCs assess key technical aspects such as energy demand (kWh/m^2/year), insulation levels, heating and cooling system efficiency, natural ventilation, solar orientation, and types of windows and glazing. By making energy performance a visible and comparable attribute, EPCs aim to influence market behaviour, guide investment in building retrofits, and support the EU's broader decarbonisation goals.

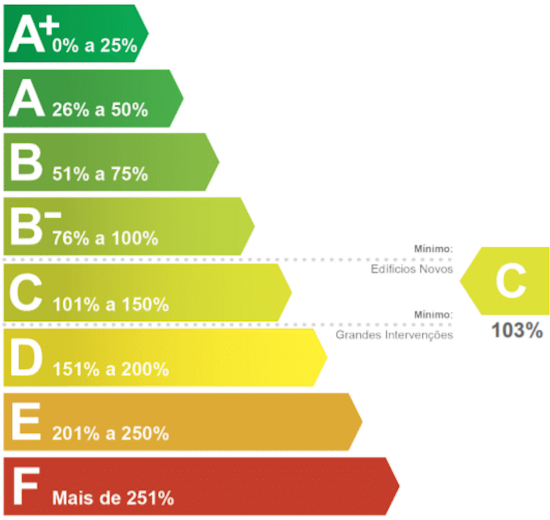
3.1 Evolution of the EPC System in Europe

The legal basis for EPCs was first established through the EPBD, formally adopted in 2002 (Directive 2002/91/EC). This directive required all EU Member States to develop a standardised certification scheme to assess and disclose the energy efficiency of buildings at the time of sale or rental. Since its introduction, the EPBD has undergone three major revisions: in 2010 (Directive 2010/31/EU), in 2018 (Directive 2018/844), and most recently in 2024, as revised by the European Parliament and Council (2024). Each revision progressively strengthened the regulatory framework by expanding the scope of implementation, tightening performance thresholds, and aligning the policy more closely with the EU's climate neutrality goals under the European Green Deal[4].

[4]The European Green Deal is the EU's flagship policy package aiming to achieve climate neutrality by 2050. It includes a broad legislative package covering energy, buildings, transport, agriculture, and finance. The EPBD revision of 2024 is part of this package.

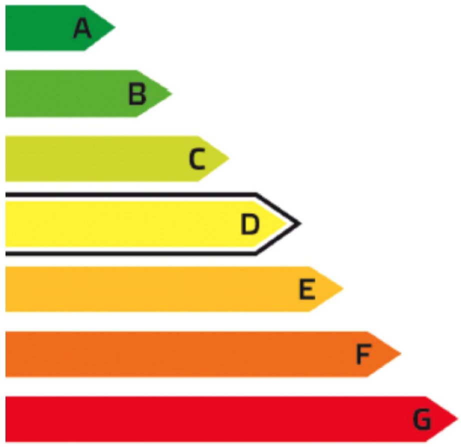
Importantly, while the EPBD sets general principles, each Member State retains autonomy in defining methodology, benchmarks, and compliance mechanisms. As a result, EPC classifications, although mandatory and widely adopted, are not directly comparable across countries. As the Buildings Performance Institute Europe (2014) explains, discrepancies in calculation methods, climatic zones, and reference building definitions reduce cross-country comparability. For example, a property rated “B” in Portugal may not meet the same technical standards as a similarly rated dwelling in Italy or Sweden. These differences are also reflected in the design and layout of EPC labels, as shown in Figure 1. Additional examples from Estonia, Germany, and Italy in Figure 4, Appendix A, reinforce this diversity in presentation and structure.

Figure 1: Comparison of EPC Label Designs in Portugal and Sweden



(a) EPC Label in Portugal

Source: Agência para a Energia (ADENE), Portuguese Energy Agency.



(b) EPC Label in Sweden

Source: Buildings Performance Institute Europe (BPIE).

This divergence is often seen as a limitation, but it reflects a strategic choice to let national systems account for climatic, technological, and construction-specific characteristics. For instance, Northern European countries prioritise heating efficiency, while Southern nations focus on cooling and solar protection. As highlighted in Kranzl et al. (2020), methodologies differ not only in technical calculations but also in how local conditions are weighted. Some countries base EPC ratings on calculated energy performance, while others integrate operational consumption data, depending on building typology or status. This flexibility helps maintain technical relevance and cost-effectiveness within each context, though it reduces comparability and complicates EU-wide analysis.

Despite these differences, EPCs across the EU share a common underlying logic. Buildings are typically classified on a scale from A+ or A (most efficient) to G (least efficient), with ratings that must follow standardised validity periods and be made publicly accessible. According to Kranzl et al. (2020), nearly all EU Member States store core elements such as general building information, energy performance data, and the current EPC rating in their databases. However, only a limited number of countries collect additional information, such as the potential EPC rating or indicators related to indoor environmental quality. Table 1 illustrates this heterogeneity across a selection of EU countries.

Table 1: Information included in EPC databases across selected EU Member States

Country	General building info	Energy performance data	Current EPC rating	Potential EPC rating	Indoor env. quality
Austria	★	★	★		
Belgium	★	★	★	★	
Denmark	★	★	★	★	
Estonia	★	★	★		
Greece	★	★	★		★
Italy	★	★	★		
Poland	★	★	★		
Portugal	★	★	★	★	★
Romania	★	★	★	★	
Germany	★				
Ireland	★	★	★	★	
France	★	★	★		
Spain	★	★	★		
Sweden	★	★	★		

Source: Adapted from Kranzl et al. (2020), with minor modifications.

The data reveal a notable degree of heterogeneity, but Portugal stands out as the only country that includes all five categories of information considered in the assessment. This suggests a relatively high level of regulatory commitment to both the transparency and comprehensiveness of energy performance certification frameworks. The availability of detailed data not only enhances the monitoring and enforcement capacity of national authorities, but also significantly increases the system's potential for supporting market-based mechanisms, such as informing real estate valuations or guiding renovation incentives.

This level of institutional engagement is particularly relevant in light of recent regulatory developments and growing environmental concerns. From 2020 onwards, new provisions under the revised EPBD have mandated that all new buildings must be constructed as Nearly Zero-Energy Buildings (NZEBs), and that major renovations must bring buildings closer to these ambitious standards[5]. As a consequence, there has been a notable increase in the issuance of EPCs with higher energy classes (A, A+ and B), particularly in densely populated urban areas where construction and renovation activity is more intense and economically significant.

3.2 EPC System in Portugal

Portugal implemented its national EPC framework in 2009, with mandatory certification for all buildings subject to sale or rental from 2013 onwards. This regulatory shift marked a significant milestone in aligning national energy policy with EU directives on building performance and sustainability. The system is administered by ADENE, which oversees certification procedures and maintains a national EPC database. The Portuguese rating scale ranges from A+ (most efficient) to G (least efficient), remaining consistent with broader EU conventions and serving as a critical tool for promoting energy transparency in the real estate market.

The criteria used in Portuguese EPCs include insulation quality, ventilation, efficiency of heating and cooling systems, window and glazing types, solar exposure, and other relevant architectural features that influence overall energy consumption and comfort levels. While these categories are common across EU Member States, their specific evaluation is carefully adapted to national climatic and technical conditions. For example, Portuguese EPCs are tailored to Mediterranean climates, placing relatively greater weight on passive solar gains and summer thermal comfort indicators, reflecting the region's unique environmental challenges and building traditions.

In addition to aligning with EU directives, Portugal has developed a comprehensive national legislative framework to support the implementation and enforcement of its EPC system. This includes the Long-Term Strategy for Building Renovation (ELPRE), which defines decarbonisation targets for the national building stock, and a set of legal instruments such as *Decree-Law No. 101-D/2020*, which establishes energy performance requirements and regulates the certification process, and *Decree-Law No. 102/2021*, which governs the qualifications of energy certifiers. Together, these instruments ensure the system's technical robustness, promote energy efficiency in both new constructions and renovations, and support Portugal's climate and energy goals.

[5]Under Article 7 of Directive (EU) 2018/844, NZEBs are buildings with very high energy performance, where the nearly zero or very low amount of required energy is covered to a significant extent by renewable sources.

3.3 EPCs in the Context of ESG and Regulatory Reporting

EPCs have gained increased importance in the financial sector, particularly as part of Environmental, Social, and Governance (ESG) disclosure requirements. Under the EU's Pillar 3 disclosure regulation, banks must report energy efficiency exposure in their real estate portfolios using EPC data[6]. The granularity of EPCs—especially those that include postcode-level data—enables institutions to assess physical risk exposures in mortgage portfolios and to align with green lending criteria.

While ADENE's public datasets support academic research, more detailed datasets with higher granularity—down to the postal code and individual building characteristics—are being developed to serve regulatory compliance. These datasets are critical for financial institutions seeking to assess transition and physical climate risks embedded in housing portfolios and improve the accuracy and reliability of ESG reporting.

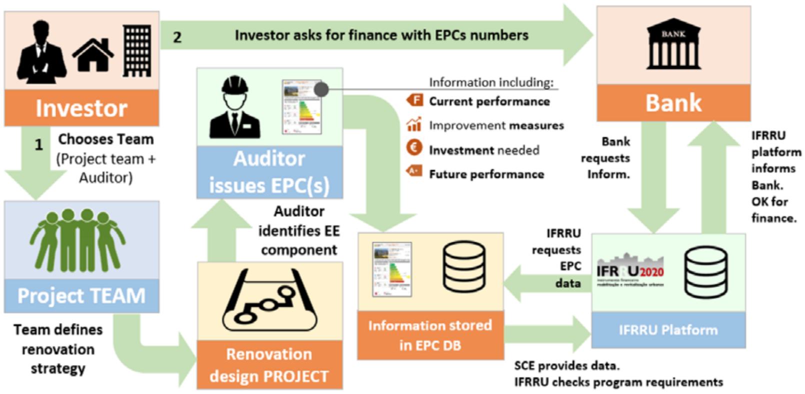
Moreover, the integration of environmental and climate risk assessments by financial institutions, particularly large institutions, is now relatively well established. In the case of mortgage lending, for example, the presentation of an energy performance certificate by the client has become a mandatory condition for loan approval. Institutions have also developed specific indicators and metrics to monitor environmental and climate factors—such as transition and physical risks—allowing them to track the evolution of their portfolios, especially housing credit, as highlighted by the European Central Bank (2020).

In addition to regulatory compliance, EPCs play a growing role in sustainable finance strategies. Financial institutions are increasingly incorporating EPC ratings into credit risk assessments, loan pricing models, and the eligibility criteria for green mortgages and green bonds. Buildings with higher EPC ratings (A, A+, B) are typically associated with lower energy costs and reduced climate-related risks, which can improve borrowers' repayment capacity and enhance asset values. This alignment with environmental performance enhances the credibility of ESG-labelled financial products and facilitates capital allocation towards more sustainable assets.

Beyond their role in risk assessment and credit evaluation, EPCs also underpin national energy efficiency policies through financial instruments. In Portugal, programmes like the Financial Instrument for Urban Rehabilitation and Revitalisation (IFRRU) and Energy Efficiency Fund (FEE) rely on EPCs to verify renovation compliance, estimate investment needs, and provide key information to financial institutions. These initiatives direct investments into decarbonising and modernising Portugal's building stock. Figure 2 illustrates how EPC data supports financing decisions by connecting investors, renovation teams, auditors, and financial institutions.

[6] European Banking Authority (EBA) Implementing Technical Standards on ESG disclosure (Regulation 2022/2453), which mandates the inclusion of EPC-based energy performance data in Templates 2 and 3.

Figure 2: Use of EPCs as a financial tool in Portugal



Source: Kranzl et al. (2020), Section 8.2.4.

Furthermore, the integration of EPC data into financial decision-making is reinforced by broader EU-level initiatives, such as the EU Taxonomy, which requires financial market participants to disclose the extent to which their investments align with environmental objectives, including energy efficiency in the built environment. In this context, EPCs serve as a practical and standardised proxy for evaluating the environmental performance of real estate assets.

To ensure consistency and relevance, the dataset employed in this study includes only EPC records issued from 2020 onwards. This restriction is essential to align with the most recent certification criteria and to avoid distortions caused by regulatory and methodological changes over time. For instance, a building rated "B" in 2015 might no longer meet the same threshold in 2024 due to updates in calculation methods, emission factors, or thermal performance benchmarks. Such regulatory and methodological changes directly affect EPC ratings, even within the same country and for the same property.

The analysis is further restricted to residential buildings, consistent with the dissertation's main research question and policy focus. Concentrating on the residential sector allows for a more precise examination of how energy performance information influences consumer behaviour, investment decisions, and price formation. Residential properties are also the most impacted by recent EU-level sustainability regulations, making them an ideal subject for studying the intersection between energy efficiency and housing markets. By focusing on the post-2020 period—when the latest revisions of the EPBD and national transpositions began to take full effect—this study enhances comparability, policy relevance, and empirical robustness. Section 4 presents the econometric framework applied to investigate these relationships, detailing the panel data models used and their methodological rationale.

4 Methodology

This section outlines the econometric strategy adopted to investigate the relationship between energy efficiency and housing prices across Portuguese municipalities. The analysis focuses exclusively on mainland Portugal, comprising a panel dataset with 278 municipalities observed over four years (2020 to 2023). To account for both observed and unobserved heterogeneity, several panel data estimators are considered, including Pooled Ordinary Least Squares (POLS), Random Effects (RE), FE, and the Poisson Pseudo Maximum Likelihood (PPML) with fixed effects. Each model is introduced below with its technical properties and microeconometric implications.

4.1 Log-linear Models

The estimation strategy is built upon a log-linear model that serves as a common framework for the subsequent estimators. The general specification for panel data can be expressed as:

$$\ln(y_{it}) = \alpha + \mathbf{x}_{it}'\boldsymbol{\beta} + c_i + u_{it}, \quad i = 1, \ldots, N; \; t = 1, \ldots, T \tag{1}$$

In this formulation, $\ln(y_{it})$ denotes the logarithm of the median housing price per square meter (`ln_sqmprice`) for municipality i in year t; $\mathbf{x}_{it}$ is a vector of explanatory variables, including the share of highly energy-efficient buildings and other relevant controls; c_i represents unobserved, time-invariant municipal characteristics (e.g., geographical features or persistent local policies); and u_{it} is the idiosyncratic error term.

Pooled Ordinary Least Squares (POLS)

The estimation begins with the application of a POLS model, which serves as a baseline approach for the analysis. The POLS method ignores the panel structure of the data and treats the dataset as a simple cross-sectional time-series sample, effectively pooling all observations across municipalities and years. This simplification facilitates coefficient estimation and an initial assessment of the variable relationships.

In this setting, the composite error term defined in Equation 1 is treated as a single, undifferentiated component (v_{it}), which does not separate the unobserved, time-invariant municipal effects (c_i) from the idiosyncratic error term (u_{it}) that varies across both municipalities and years. As a result, if c_i is correlated with the regressors $\mathbf{x}_{it}$, the POLS estimates will be biased and inconsistent. This issue motivates the use of panel estimators that explicitly account for the dependence structure in the error term, ensuring more reliable inference and robust estimation of the effects of energy efficiency on housing prices.

Cluster-robust errors are computed at the municipality level, providing valid inference under arbitrary forms of heteroskedasticity and within-cluster correlation (White 1980; Arellano 1987). These adjustments are essential when using POLS in panel data settings.

Despite its limitations, the POLS model offers an intuitive starting point and enables a preliminary assessment of whether energy efficiency is significantly associated with property prices without yet controlling for municipality-specific effects. It also serves as a useful reference when comparing the performance of more advanced models, such as RE, FE, and PPML, which account for panel-specific characteristics and address different types of endogeneity and functional form concerns.

Random Effects Model

Following the estimation of the baseline POLS model, which serves as a preliminary benchmark, a RE model is subsequently estimated as a potentially more efficient alternative specification, particularly in contexts where the unobserved heterogeneity across municipalities is assumed to be uncorrelated with the explanatory variables included in the regression equation. The RE specification explicitly accommodates the panel structure of the dataset by modelling a composite error term that captures both municipality-specific and idiosyncratic components of the error, just as shown in Equation 1.

A critical assumption underpinning the RE framework is that the unobserved effect c_i is uncorrelated with the regressors $\mathbf{x}_{it}$ across all time periods t, which, if satisfied, enables consistent estimation of the model parameters. Under this assumption, estimation can proceed through the Generalised Least Squares (GLS) method, which adjusts for the intra-municipality correlation structure inherent in the composite error term. The GLS estimation procedure involves a transformation of the data commonly referred to as *quasi-demeaning*, a process by which a fraction λ of the time-specific average is subtracted from each observation, thereby mitigating the influence of time-invariant unobserved heterogeneity on the estimation results:

$$v_{it}^{*} = v_{it} - \lambda \bar{v}_i. \tag{2}$$

Fixed Effects Model

The FE estimator is specifically designed to control for unobserved heterogeneity that remains constant over time and may be correlated with the explanatory variables included in the model specification. It relies on the same general model in Equation 1, but follows a different estimation approach.

A fundamental assumption underlying the FE model is that the unobserved effect c_i may, in fact, be correlated with the included regressors, which distinguishes it from the assumptions made in the RE framework. In order to eliminate the influence of c_i and obtain consistent parameter estimates, the model undergoes a transformation process known as the "within transformation", where individual means are subtracted from each observation in the panel dataset:

$$\ddot{y}_{it} = \ddot{\mathbf{x}}'_{it}\boldsymbol{\beta} + \ddot{u}_{it} \tag{3}$$

where $\ddot{y}_{it} = y_{it} - \bar{y}_i$, $\ddot{\mathbf{x}}_{it} = \mathbf{x}_{it} - \bar{\mathbf{x}}_i$, and so on. This transformation eliminates c_i, allowing consistent estimation of $\boldsymbol{\beta}$ even when c_i is endogenous.

A matrix formulation of the FE estimator is:

$$\hat{\boldsymbol{\beta}}_{\text{FE}} = (\ddot{X}'\ddot{X})^{-1}\ddot{X}'\ddot{y} \tag{4}$$

While the FE estimator offers a robust approach for dealing with time-invariant heterogeneity, it comes with a notable limitation: it cannot estimate coefficients of time-invariant covariates, as these are removed by the within transformation that subtracts unit-specific means from each observation.

In line with the preceding estimation strategies, the FE estimators are implemented with cluster-robust standard errors, clustered at the municipality level. This correction technique allows for arbitrary forms of heteroskedasticity and within-cluster serial correlation, while maintaining the crucial assumption of independence across municipalities.

Model selection between the RE and FE estimators is guided primarily by the Hausman test, which evaluates the consistency of the RE estimator under the null hypothesis that the individual effects c_i are uncorrelated with the regressors:

$$H_0 : \text{Cov}(c_i, \mathbf{x}_{it}) = 0 \tag{5}$$

$$H_1 : \text{Cov}(c_i, \mathbf{x}_{it}) \neq 0 \tag{6}$$

If the null is not rejected, both RE and FE are consistent but RE is efficient; otherwise, FE is preferred. Formally, the test statistic is given by

$$H = (\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE})' \left[\text{Var}(\hat{\boldsymbol{\beta}}_{FE}) - \text{Var}(\hat{\boldsymbol{\beta}}_{RE})\right]^{-1} (\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE}) \tag{7}$$

following a chi-squared distribution. As a robustness check, the Chamberlain–Mundlak device augments the RE model with the means of time-varying covariates $\bar{\mathbf{x}}_i$, estimating

$$y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + \bar{\mathbf{x}}'_i\boldsymbol{\theta} + c_i + u_{it} \tag{8}$$

where joint significance of $\boldsymbol{\theta}$ indicates correlation between regressors and c_i, violating RE assumptions.

4.2 Exponential Models

The exponential conditional mean model provides a flexible functional form to model strictly positive dependent variables, ensuring non-negativity of predictions while accommodating various forms of heteroskedasticity. In this context, the FE Poisson estimator proposed by Hausman, Hall, and Griliches (1984) is particularly suitable for panel data settings with unobserved heterogeneity, allowing for consistent estimation without requiring distributional assumptions on the error term.

Unlike traditional log-linear FE models estimated by Ordinary Least Squares (OLS) on log-transformed outcomes, the Hausman FE Poisson models the conditional mean directly via an exponential form, avoiding retransformation bias and heteroskedasticity-related inconsistency common in log-linear OLS estimators. These features are especially relevant in housing markets, where price distributions are skewed and variance heterogeneity is frequent.

The model specification assumes that the conditional expectation of the dependent variable follows an exponential form:

$$E[y_{it}|\mathbf{x}_{it}, c_i] = \exp(\mathbf{x}'_{it}\boldsymbol{\beta} + c_i), \tag{9}$$

where y_{it} represents the strictly positive outcome variable, $\mathbf{x}_{it}$ is a vector of covariates, and c_i captures time-invariant, municipality-specific unobserved effects. Unlike RE Poisson models requiring distributional assumptions on c_i and independence from regressors, the FE Poisson estimator remains consistent despite correlation, offering robustness for this empirical setting.

Given the considerable computational challenges associated with estimating high-dimensional FE, the practical implementation of this estimator leverages specialised computational techniques. In particular, the PPML estimator with high-dimensional fixed effects is employed, which utilizes advanced iterative algorithms to efficiently absorb unit-level fixed effects, thereby facilitating estimation. Importantly, while the estimator draws from the Poisson family, it does not require the dependent variable to follow a Poisson distribution, focusing instead on the correct specification of the conditional mean function.

The Hausman FE Poisson approach employed in this dissertation is further complemented by the incorporation of cluster-robust standard errors, with clustering performed at the municipality level. This methodological choice enhances the robustness of the estimation framework, providing reliable inference even under complex error structures. Consequently, this estimator constitutes an appropriate and consistent modelling strategy for analysing relationships involving continuous, strictly positive outcomes in panel data settings marked by unobserved heterogeneity, making it a particularly suitable choice for the empirical investigation of the relationship between energy efficiency and housing prices.

5 Data Description

This study relies on a panel dataset comprising municipality-year observations for mainland Portugal, covering the period from 2020 to 2023. The primary objective is to explore the relationship between energy performance ratings of buildings and the housing price, particularly the median square meter transaction price of residential properties. The unit of observation is the municipality-year, meaning that each data point represents aggregate values for a given municipality in a specific year.

The dataset was constructed by merging publicly available information from multiple official sources. The primary data on energy performance was obtained from ADENE, which provides detailed building-level records of EPCs issued in Portugal. Socioeconomic and demographic indicators were retrieved from INE, complemented by data from DGEEC, IEFP and DGPJ. All variables were harmonised to the municipality-year level, resulting in a panel structure with 278 municipalities observed over four years—yielding 1,112 potential observations.

Although the panel is structurally complete in terms of identifiers and time periods, missing values in the variable of interest—the median square meter price—reduce the effective sample. Specifically, this variable is missing for 18 municipalities in 2020, 7 in 2021, 4 in 2022, and 6 in 2023, resulting in 35 missing observations. All other variables are fully observed. As a result, and since this issue affects only the outcome variable, the final analytical sample comprises 1,077 municipality-year observations across 275 municipalities[7].

This implies that the panel dataset becomes unbalanced due to outcome-specific missingness. Such instances of minor unbalancedness are a frequent occurrence in empirical research that relies on administrative data sources, given that data collection processes are often subject to reporting delays or occasional gaps in record-keeping. Nevertheless, the econometric estimators employed in this dissertation cope with this type of irregularity, assuming that the missingness is not systematically related to the error term (Arellano 2003, Ch. 2–3). This assumption ensures that the estimators remain consistent and unbiased, preserving the validity of the empirical analysis.

It is important to contextualize this analysis within Portugal's broader housing challenges. Portugal ranks second in the EU for the share of population living in poor housing conditions (PORDATA 2023), highlighting persistent housing quality issues and the relevance of studying residential energy performance. In this context, EPCs are likely to become increasingly widespread. As discussed in Section 3, EPCs are mandatory for all major renovations and new constructions, supporting efforts to improve housing standards and energy efficiency nationwide.

[7]The municipalities omitted from the analysis due to missing data were: Barrancos, Freixo de Espada à Cinta, and Penedono.

5.1 Main Variables and Transformations

The main outcome variable analysed in this study is the median price per square meter of residential dwellings (`sqmprice`), obtained from INE. This specific variable serves as a comprehensive summary indicator of the housing market conditions at the municipality level, providing an aggregated measure that effectively mitigates the distorting influence of extreme or atypical transaction values. The use of medians aligns with standard practice when relying on aggregated administrative datasets, which are typically constructed to suppress or smooth out extreme values to ensure statistical confidentiality. Although this aggregation limits the granularity of the analysis, it offers a valuable alternative that preserves representativeness and reduces the impact of outliers.

The primary independent variable of interest is `greenepcs`, a measure of the share of highly energy-efficient buildings in a given municipality-year. This variable was constructed by the author as the ratio of residential EPCs rated A+, A, or B to the total residential EPCs issued in each municipality. These top-rated certificates indicate strong energy performance, typically linked to lower energy use and better thermal comfort. As such, the variable captures the relative prevalence of energy-efficient housing within each local market. A related metric—the CO_2 emissions index (`CO2index`)—was also computed as the average emissions per EPC, based on municipality-level totals from ADENE. However, `greenepcs` was preferred over this index to avoid multicollinearity and because it aligns more directly with the rating-based policy instruments studied in the literature (Fuerst et al. 2016).

Given the dataset's structure and large variation in municipality size—from under 1,500 to over 567,000 residents—several transformations were applied to improve model quality and interpretation. Absolute variables (e.g., number of firms or students), highly correlated with population size, were converted into proportions (e.g., `firmsprop`, `studentsprop`) to reduce bias and multicollinearity. Right-skewed variables like `sqmprice`, `meanwage`, and `popdens` were log-transformed to correct skewness and enable elastic coefficient interpretation. A metropolitan area dummy was also considered but remained constant over time. As FE models drop time-invariant variables and RE models include them with limited added value, it was excluded from the analysis.

The author expects a positive relationship between `greenepcs` and the dependent variables, similar to findings in previous studies suggesting that buyers are willing to pay price premiums for energy efficiency—either due to the expected cost savings from reduced energy bills or due to environmental awareness. This relationship is also supported by the broader European literature on housing markets and energy labelling, where certificate-based efficiency ratings have been shown to affect property valuations under varying market and institutional conditions (Fuerst et al. 2015).

5.2 Control Variables

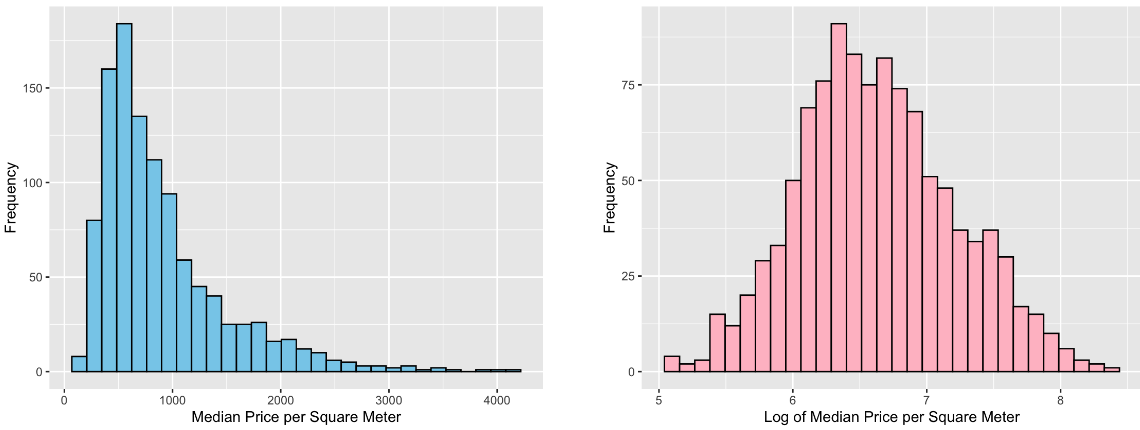
To account for other factors influencing housing prices, several control variables were carefully selected and included in the dataset. These variables capture a broad range of socioeconomic, demographic, and infrastructural characteristics that may affect housing prices across different municipalities and over time. A comprehensive description of each control variable follows, including any applied transformations, the expected direction of the relationship based on economic theory and empirical evidence, as well as relevant supporting literature that motivates their inclusion in the analysis and helps interpret the results within a broader economic context.

- **`ln_meanwage`**: Log of average gross monthly earnings. Higher wages are expected to be associated with higher house prices, reflecting greater purchasing power (Liang, Lu, and Zhang 2016).
- **`empeduhighratio`**: Share of the employed population with post-secondary or tertiary education. Higher educational attainment is expected to correlate positively with housing prices (Fack and Grenet 2010).
- **`ln_popdens`**: Log of population density. Denser areas usually face supply constraints and higher demand, pushing prices up (Zhou et al. 2019).
- **`firmsprop`**: Number of firms per capita. This variable serves as a proxy for economic activity and local employment opportunities (Gibbons, Overman, and Sarvimäki 2021).
- **`repcrimesprop`**: Reported crimes per capita. A negative relationship with housing prices is anticipated, as higher crime rates may reduce the attractiveness of an area to potential buyers (Kortas, Grigoriev, and Piccillo 2022).
- **`foreignpopprop`**: Share of foreign residents. Immigration may influence housing demand depending on labor market conditions and integration policies (Tai et al. 2017).
- **`studentsprop`**: Share of residents enrolled in basic and secondary education, reflecting family-oriented housing demand and demographic structure.
- **`unemprate`**: Proxy for unemployment rate, defined as registered unemployed divided by the sum of employed and registered unemployed in each municipality. This approximates the active population but excludes unregistered job seekers. Higher unemployment typically reduces housing demand.
- **`schoolprop`**: Number of schools per capita. May serve as a proxy for access to public services and local amenities (Fack and Grenet 2010).

5.3 Descriptive Statistics

The outcome variable, `sqmprice`, exhibits a right-skewed distribution—a common characteristic of housing markets due to the presence of high-value properties in certain municipalities, particularly in urban coastal and metropolitan areas where price variability is more pronounced and demand pressures are elevated. The logarithmic transformation of this variable, `ln_sqmprice`, was also considered in linear models. This conversion facilitates interpretation, as estimated coefficients can be interpreted as elasticities or semi-elasticities. Figure 3 displays histograms of `sqmprice` and `ln_sqmprice`, clearly illustrating the expected improvement in symmetry achieved through the log transformation.

Figure 3: Distribution of Price per Square Meter and Its Logarithmic Transformation



(a) Distribution of 'sqmprice' (Price per Square Meter)

(b) Distribution of 'ln_sqmprice' (Logarithm of Price per Square Meter)

Source: Own elaboration using RStudio, based on data from INE.

Table 2 reports descriptive statistics for the key variables used in the analysis. The median logarithm of the median square meter price (`ln_sqmprice`) is 6.63, corresponding approximately to a median price of EUR 760 per square meter, based on the inverse log transformation. The raw `sqmprice` variable, shown in the table, ranges widely from EUR 156 to EUR 4,167 per square meter, reflecting substantial heterogeneity across municipalities. The average share of green-certified buildings (`greenepcs`) is 34%, indicating a moderate presence of high-efficiency properties in the dataset across the sample period. Socioeconomic control variables such as mean wage (`ln_meanwage`), education ratio (`empeduhighratio`), and unemployment rate (`unemprate`) show expected variation across municipalities, reflecting differences in local development levels and labor market conditions.

Table 2: Descriptive Statistics for Key Variables

Variable	Min	1st Quartile	Median	Mean	3rd Quartile	Max
sqmprice	156	500	737	905.86	1115	4167
ln_sqmprice	5.05	6.30	6.63	6.63	6.95	8.33
greenepcs	0.00	0.22	0.34	0.34	0.46	0.70
ln_meanwage	6.72	6.90	7.00	7.00	7.10	7.82
empeduhighratio	0.22	0.41	0.48	0.48	0.54	0.81
unemprate	0.04	0.10	0.16	0.16	0.21	0.50
ln_popdens	1.44	2.50	4.31	4.31	5.75	8.92
firmsprop	0.06	0.10	0.13	0.13	0.16	0.26
repcrimesprop	0.01	0.02	0.03	0.03	0.04	0.09
foreignpopprop	0.00	0.01	0.05	0.05	0.08	0.44
studentsprop	0.06	0.10	0.13	0.13	0.16	0.28
schoolprop	0.02	0.06	0.09	0.09	0.12	0.27

Source: Own elaboration based on data from ADENE, INE, and other official sources.

The table also highlights substantial variation across municipalities in both housing and socioeconomic characteristics, which is crucial for identification. For example, the logarithm of population density (`ln_popdens`) ranges from 1.44 to 8.92, indicating sharp contrasts between rural and urban areas in terms of settlement patterns. Other variables, such as the proportion of foreign residents (`foreignpopprop`) and the share of firms (`firmsprop`), also exhibit wide dispersion, reflecting structural differences in local demographics, economic activity, and housing demand. This observed heterogeneity supports the inclusion of these covariates in the regression analysis, as they help control for local factors that may influence housing prices over time.

In addition to the descriptive statistics, Appendix B provides a correlation matrix (Table 6) summarising the pairwise relationships between housing prices and key covariates. As expected, the raw median price per square meter (`sqmprice`) exhibits meaningful positive correlations with `greenepcs` (0.26), `foreignpopprop` (0.58), `ln_popdens` (0.55), and `ln_meanwage` (0.52), illustrating that higher housing prices tend to be associated with areas of greater energy efficiency, higher income levels, denser urban settlements, and a larger share of foreign residents. Conversely, a negative correlation is observed with `unemprate` (-0.27), suggesting that higher unemployment rates are characteristic of municipalities with relatively lower property prices. It is also worth noting that `greenepcs` could be subject to potential endogeneity, as households may prefer newer or higher-quality dwellings that tend to have better EPC ratings and higher prices. Nonetheless, the panel data framework and inclusion of municipal-level controls and fixed effects help mitigate this concern.

6 Empirical Results

This section presents the empirical findings from the econometric analysis of the determinants of housing prices across municipalities, focusing on the impact of energy-efficient certificates and other key covariates. The results are derived from a series of regression models, with particular emphasis on the FE and PPML. The analysis builds on the methodology previously outlined, addressing model selection, robustness checks, and the interpretation of the preferred model's coefficients. All regressions incorporate time fixed effects to account for year-specific shocks and national trends, and robust standard errors clustered at the municipality level are used to address heteroskedasticity and autocorrelation.

6.1 Model Selection and Diagnostics

The choice of the econometric model was based on diagnostic tests to ensure consistent and efficient estimation of the relationship between `greenepcs` and housing prices. First, a POLS model was estimated as a baseline, using `ln_sqmprice` as the dependent variable. However, POLS does not control for unobserved time-invariant differences across municipalities, which may cause biased estimates. To test for such heterogeneity, the Breusch-Pagan Lagrangian multiplier test was conducted, strongly rejecting the null hypothesis that the variance of random effects is zero ($\chi^2(1) = 1102.59$, $p = 0.0000$). This result indicates that a panel data model, such as RE or FE, is more appropriate than POLS.

To choose between RE and FE, the Hausman test checked whether unobserved municipal effects correlate with regressors. It was run twice—excluding and including time fixed effects—to ensure robustness. Both versions strongly rejected the null hypothesis that the RE estimator is consistent ($\chi^2(10) = 72.06$, $p = 0.0000$ without time effects; $\chi^2(13) = 59.96$, $p = 0.0000$ with time effects), favouring the FE model. For further robustness, the Chamberlain-Mundlak test was performed by adding time-averaged regressors in the RE model. This test also rejected the null that the means of regressors are jointly zero ($\chi^2(10) = 52.35$, $p = 0.0000$), reinforcing the preference for the FE model due to correlation between unobserved effects and regressors.

Given the panel data structure, heteroskedasticity and autocorrelation were expected. The modified Wald test confirmed heteroskedasticity ($\chi^2(275) = 8.8 \times 10^{28}$, $p = 0.0000$), and the Wooldridge test detected first-order autocorrelation ($F(1, 269) = 7.577$, $p = 0.0063$). While these issues are often endemic in panel data, their confirmation justified using cluster-robust standard errors. Table 3 compares POLS, RE, FE, and PPML estimates, showing coefficient attenuation in linear models (e.g., `greenepcs` drops from 0.667 in POLS to 0.186 in FE). The PPML model, using `sqmprice` levels as dependent variable instead of logarithm, confirms these patterns.

Table 3: Comparison of POLS, Random Effects, Fixed Effects, and PPML Estimates

Variable	POLS	Random Effects	Fixed Effects	PPML
greenepcs	0.667***	0.270***	0.186**	0.186***
ln_meanwage	0.461***	0.378**	0.208	0.057
empeduhighratio	0.312	0.241	0.226	0.481**
unemprate	-0.227	-0.382**	0.020	-0.216
ln_popdens	0.156***	0.182***	1.626***	1.667***
firmsprop	-0.577	0.431	2.907**	2.200***
repcrimesprop	5.795**	1.800	0.522	-0.777
foreignpopprop	3.550***	3.555***	1.356**	1.015**
studentsprop	2.787***	2.602***	1.205	2.301***
schoolprop	-0.389	-0.059	0.957	1.882
Year Effects				
2021	0.021*	0.027**	0.030**	0.050***
2022	0.035*	0.062***	0.086***	0.113***
2023	-0.024	0.021	0.076*	0.126***
Constant	1.771	2.382**	-2.847	-2.973

Legend: * p<0.1; ** p<0.05; *** p<0.01

Note: The PPML regression uses `sqmprice` (levels) as the dependent variable, whereas the other models use `ln_sqmprice` (logarithm of sqmprice).

Source: Own elaboration based on regression outputs from Stata.

Table 5 (Appendix B) shows the within- and between-municipality variation of key covariates. Some variables, like `ln_popdens`, `firmsprop`, and `schoolprop`, have limited within-municipality variation, which can reduce the precision of FE estimates. This explains the loss of significance in some coefficients, but the PPML model, estimating `sqmprice` in levels, still captures the key relationships robustly.

Although the log transformation reduced skewness in the dependent variable, residual diagnostics—illustrated in Figure 5 (Appendix C)—revealed systematic patterns, suggesting specification concerns. A formal RESET test on the FE model, however, did not indicate significant functional form misspecification. This divergence arises because the RESET test evaluates polynomial expansions of regressors, whereas visual inspection can detect additional patterns, such as local non-linearities. Nevertheless, using a PPML estimator with high-dimensional fixed effects mitigates these issues by modeling the dependent variable in levels, accommodating heteroskedasticity and absorbing unobserved heterogeneity. This specification yields stronger significance for key covariates, particularly `empeduhighratio` and `studentsprop`, outperforming the FE specification.

6.2 Robustness Checks

Several robustness checks were performed to validate the model specifications. First, models excluding `ln_popdens` were estimated to assess its contribution to explaining housing prices and to examine whether other covariates would gain individual significance in its absence. The results, however, showed no improvement, highlighting the critical role of population density as a proxy for urbanisation and demand pressures, which enhances both model fit and robustness (Zhou et al. 2019).

The functional form of `greenepcs` was tested for nonlinearity by adding a quadratic term in the FE model. The squared term was not significant ($p = 0.229$) and the linear term became insignificant ($p = 0.957$). Marginal effects (see Figure 6 in Appendix C) are almost flat, showing no nonlinearity. Additionally, a logarithmic transformation ($\log(greenepcs+1)$) was tested to account for one zero observation in the data. In the model incorporating the log-transformed `greenepcs` term, the direction and significance of other covariates remained largely unchanged. These results provide support for retaining the linear specification of `greenepcs` in the final model.

The inclusion of time fixed effects was further validated and proved critical across all models, as they were jointly significant ($F(3, 274) = 8.33$, $p = 0.0000$ in POLS, with similar results in RE and FE). These effects control for national-level trends, such as macroeconomic developments or policy changes, that affect housing prices. The decline in significance of `ln_meanwage` in the FE model suggests that its effect is largely captured by these time effects, consistent with findings by Liang, Lu, and Zhang (2016) that wages reflect broader economic conditions. The joint significance of the time dummies reinforces their importance in capturing temporal effects.

6.3 Interpreting FE Poisson Estimates

The PPML model with high-dimensional fixed effects is specified as shown in Equation (10). The dependent variable corresponds to the median housing price per square meter (`sqmprice`) in municipality i at time t. The vector of covariates $\mathbf{x}_{it}$ includes energy efficiency (`greenepcs`), educational attainment (`empeduhighratio`), log population density (`ln_popdens`), firms per capita (`firmsprop`), share of foreign residents (`foreignpopprop`), and proportion of residents enrolled in education (`studentsprop`). Municipality fixed effects c_i and time fixed effects γ_t control for unobserved heterogeneity across municipalities and years, respectively.

$$E[sqmprice_{it} \mid \mathbf{x}_{it}, c_i, \gamma_t] = \exp\Big(\beta_0 + \beta_1 \text{greenepcs}_{it} + \beta_2 \text{empeduhighratio}_{it} + \beta_3 \text{ln\_popdens}_{it} + \beta_4 \text{firmsprop}_{it} + \beta_5 \text{foreignpopprop}_{it} + \beta_6 \text{studentsprop}_{it} + c_i + \gamma_t\Big). \tag{10}$$

This model was chosen for its ability to directly model the non-negative and right-skewed `sqmprice` without log transformation, improving interpretability and yielding a higher number of statistically significant covariates compared to the FE model with logged prices. Variables deemed insignificant (`unemprate`, `ln_meanwage`, `repcrimesprop`, and `schoolprop`) were excluded based on a joint significance test ($\chi^2(4) = 3.28$, $p = 0.5122$). Results are summarised in Table 4.

Table 4: Final Poisson Pseudo-Maximum Likelihood with High-Dimensional Fixed Effects model with Clustered Standard Errors

Variable	Coefficient	Std. Error	z	p-value	95% Conf. Interval
greenepcs	0.193	0.054	3.61	0.000	[0.088; 0.299]
empeduhighratio	0.500	0.226	2.21	0.027	[0.057; 0.942]
ln_popdens	1.566	0.299	5.25	0.000	[0.981; 2.151]
firmsprop	2.422	0.653	3.71	0.000	[1.143; 3.702]
foreignpopprop	0.934	0.507	1.84	0.066	[-0.061; 1.928]
studentsprop	2.279	0.810	2.81	0.005	[0.691; 3.867]
Year Effects					
2021	0.050	0.007	6.78	0.000	[0.036; 0.065]
2022	0.117	0.012	9.48	0.000	[0.093; 0.142]
2023	0.132	0.020	6.77	0.000	[0.094; 0.171]
Constant	-1.994	1.536	-1.30	0.194	[-5.005; 1.017]

Clustered robust standard errors at the municipality level.

Pseudo R-squared: 0.9615 Wald chi2(9): 1779.74 Prob > chi2: 0.0000

Source: Own elaboration based on PPML regression outputs using `ppmlhdfe` in Stata.

The coefficient for `greenepcs` is positive and highly significant ($\beta = 0.193$, $p < 0.001$), indicating that a one percentage point (pp) increase in the share of highly energy-efficient buildings (rated A+, A, or B) is associated with a 0.193% increase in `sqmprice`, holding other factors constant. This finding aligns with the hypothesis that energy-efficient housing commands a market premium, likely due to lower energy costs, improved thermal comfort, or growing consumer demand for sustainable properties (Fuerst et al. 2015). As the primary variable of interest, `greenepcs` reflects the market's valuation of energy efficiency, a key consideration in the context of rising environmental awareness and energy costs. The robustness of this effect across models, although weaker in FE and PPML with High-Dimensional FE compared to POLS, underscores its importance after controlling for unobserved heterogeneity.

`firmsprop` has a strong effect ($\beta = 2.422$, $p < 0.001$), where a 1pp increase in firms per capita is associated with about a 2.42% increase in prices, capturing economic vibrancy and job opportunities (Gibbons, Overman, and Sarvimäki 2021).

The coefficient for `ln_popdens` is significant ($\beta = 1.566$, $p < 0.001$), suggesting a 1% increase in population density raises housing prices by 1.566%. This reflects supply constraints and heightened demand in urban areas, consistent with Zhou et al. (2019). `studentsprop` is significant ($\beta = 2.279$, $p = 0.005$), showing a 1 percentage point rise in students increases prices by roughly 2.3%, likely due to demand for family-oriented housing near schools.

`empeduhighratio` is significant ($\beta = 0.500$, $p = 0.027$), suggesting that a one-unit increase in the share of highly educated employees increases prices by 50.0%, reflecting higher purchasing power in areas with skilled workers (Fack and Grenet 2010). `foreignpopprop` is marginally significant ($\beta = 0.934$, $p = 0.066$), where a 1 percentage point increase in the share of foreign population corresponds to an approximate 0.93% increase in housing prices, possibly due to immigration-driven demand (Tai et al. 2017).

The time fixed effects for 2021, 2022, and 2023 are all highly significant ($p < 0.001$), with coefficients of 0.050, 0.117, and 0.132, respectively. These positive coefficients indicate that housing prices increased steadily over time relative to the base year (2020). This upward trend likely reflects broader macroeconomic factors, including economic recovery after downturns, rising inflation, and evolving housing market dynamics that influenced demand and pricing during the study period.

7 Conclusion

This dissertation examines the relationship between energy efficiency and housing prices at the municipal level in Portugal, focusing on residential properties from 2020 to 2023. The empirical analysis, employing microeconometric techniques, reveals a robust positive association between the share of high-efficiency EPCs and median housing prices per square meter. The preferred PPML specification with high-dimensional fixed effects indicates that a one percentage point (pp) increase in the share of energy-efficient buildings (rated A+, A, or B) is associated with a 0.193% increase in housing prices, ceteris paribus. This result aligns with studies such as Fuerst et al. (2015) and Brounen and Kok (2011), which report price premiums for energy-efficient homes in other European markets. The findings suggest that the Portuguese housing market values energy efficiency, likely due to reduced energy costs, improved thermal comfort, and increasing consumer demand for sustainable properties.

The analysis also underscores the influence of socioeconomic and infrastructural factors on housing prices. Key covariates such as population density, the share of firms per capita, and the proportion of students significantly drive prices, reflecting demand pressures in urban areas, economic vibrancy, and family-oriented housing needs, respectively. The share of highly educated employees and the foreign population further contribute to higher prices, highlighting the role of skilled labor and immigration-driven demand. These findings are consistent with Zhou et al. (2019) and Gibbons, Overman, and Sarvimäki (2021), which emphasise the impact of urban and economic characteristics on property values. The use of time and municipality fixed effects enhances the robustness of the estimates by controlling for unobserved heterogeneity and national trends.

The positive association between EPCs and housing prices suggests that Portuguese cities pursuing sustainable urban development may experience rising property values. As municipalities invest in energy-efficient infrastructure, the increased prevalence of high-efficiency EPCs could drive housing prices upward, particularly in urban areas with high demand. However, this trend may be controversial, as rising prices could exacerbate affordability challenges for lower-income households, potentially deepening housing inequality in sustainably developing cities. Policymakers must balance these economic benefits with measures to ensure equitable access to energy-efficient housing.

The results carry important implications for policymakers in Portugal and the EU. The positive association between energy efficiency and housing prices, as captured by `greenepcs`, supports the effectiveness of the EPBD and Portugal's EPC system in incentivising sustainable construction and renovation. The significant effect of `empeduhighratio` suggests that municipalities with a higher share of highly educated employees, a proxy for higher-income groups, exhibit greater demand for energy-efficient properties, likely due to greater financial capacity and awareness of sustainability benefits. Similarly, the strong effect of `ln_popdens` indicates that urban areas, often associated with higher-income populations, drive this market sensitivity. However, for lower-income households, improving energy efficiency remains challenging due to the high costs of retrofitting, exacerbating issues of energy poverty and affordable housing, particularly in less affluent municipalities.

Public policies are essential to mitigate these disparities. Initiatives like "Eco-Bairros"[8] and regional energy renovation programmes exemplify how local interventions can promote energy-efficient urban regeneration while addressing social inclusion. Additionally, targeted financial instruments, such as subsidies or low-interest loans, are necessary to support lower-income households in upgrading their homes, aligning with the EU Taxonomy Regulation's goals of decarbonising buildings and promoting social equity in the transition to sustainable housing.

[8]"Eco-Bairros" is a public policy initiative in Portugal aimed at promoting sustainable neighbourhoods through integrated urban regeneration. It focuses on improving energy performance, enhancing mobility and public space, and fostering social inclusion in existing urban areas. Launched as part of national climate adaptation and urban development strategies, the programme supports local governments in implementing energy-efficient and low-carbon measures at the neighbourhood level.

The findings in this study strengthen the case for integrating energy efficiency into urban planning and housing policies, aligning with the EU's climate goals under the EPBD and EU Taxonomy frameworks. This research highlights Portugal's proactive approach to energy certification, providing a model for other Southern European countries with similar socioeconomic challenges. By demonstrating the market's valuation of energy-efficient buildings, the study underscores the potential for economic incentives to drive sustainable development, particularly in regions with diverse municipal characteristics.

Future research could extend this work in several directions. First, incorporating transaction-level data, if available, would allow for a hedonic pricing approach to estimate the marginal effect of individual property characteristics. Second, expanding the dataset to include more recent years or Portugal's autonomous regions could enhance generalisability. Third, a particularly interesting avenue would be to investigate whether the investment costs of renovating low-efficiency buildings to achieve high EPC ratings yield sufficient returns upon sale. Such analysis could guide homeowners and developers in prioritising energy-efficient upgrades. Fourth, exploring the impact of energy efficiency on rental markets could provide a broader view of housing dynamics. Finally, employing dynamic panel models or spatial econometric techniques could address time dynamics and spatial dependencies, further refining the understanding of energy efficiency's role in housing markets.

Another promising line of research would be to examine potential heterogeneity in the effect of EPC ratings across Portugal. For instance, energy efficiency premiums may vary not only by region, but also according to municipalities' demographic and socio-economic conditions, such as median schooling levels or income distribution. These differences could be addressed through interaction variables, for example combining `greenepcs` with NUTS regional dummies, or interacting with education and income indicators. Such an approach would identify whether the value of energy efficiency is amplified in developed urban areas or more relevant in weaker socio-economic municipalities, informing targeted policy interventions.

In conclusion, this study provides compelling evidence that energy efficiency, as measured by high-efficiency EPCs, is associated with higher housing prices in Portugal's residential market. The findings highlight the economic value of sustainable buildings and support the continued implementation of energy certification programs. By addressing data limitations through a municipal-level approach, this dissertation offers a nuanced contribution to the literature and informs policy efforts toward sustainable development. The results also underscore the need for targeted policies to support lower-income households in achieving energy-efficient homes, addressing Portugal's broader challenges of housing affordability. These insights pave the way for future research to further explore the intersection of environmental sustainability and economic incentives in housing markets.

References

Arellano, M. (1987). "Computing Robust Standard Errors for Within-Groups Estimators". *Oxford Bulletin of Economics and Statistics* 49 (4), pp. 431–434.

— (2003). "Panel Data Econometrics". *Oxford University Press.* Working Paper, CEMFI.

Ayala, A. de, I. Galarraga, and A. Spadaro (2016). "The Price of Energy Efficiency in the Spanish Housing Market". *Energy Policy* 94, pp. 16–24.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). "Inference on Treatment Effects after Selection among High-Dimensional Controls". *Review of Economic Studies* 81 (2), pp. 608–650.

Brounen, D. and N. Kok (2011). "On the Economics of Energy Labels in the Housing Market". *Journal of Environmental Economics and Management* 62 (2), pp. 166–179. DOI: 10.1016/j.jeem.2010.11.006.

Buildings Performance Institute Europe (2014). *Energy Performance Certificates across the EU: A Mapping of National Approaches.* Tech. rep. BPIE.

Cajias, M. and S. Bienert (2014). "Greenwashing in the Real Estate Sector? The Role of Certification Credibility". *Journal of Sustainable Real Estate* 6 (1), pp. 162–179.

European Central Bank (2020). *Guide on Climate-Related and Environmental Risks: Supervisory Expectations Relating to Risk Management and Disclosure.* https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.202011final_guide_on_climate-related_and_environmental_risks~4afe3f1fb5.en.pdf. Accessed: 2025-07-22.

European Commission (2024). *Energy performance of buildings.* URL: https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings_en.

European Parliament and Council (2024). *Directive (EU) 2024/1275 of 24 April 2024 on the energy performance of buildings (recast).* Official Journal of the European Union, L 202, 8.5.2024. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024L1275.

Evangelista, R., E. A. Ramalho, and J. Andrade e Silva (2020). "On the Use of Hedonic Regression Models to Measure the Effect of Energy Efficiency on Residential Property Transaction Prices: Evidence for Portugal and Selected Data Issues". *Energy Policy* 144, p. 111672. DOI: 10.1016/j.eneco.2020.104699.

Fack, G. and J. Grenet (2010). "When Do Better Schools Raise Housing Prices? Evidence from Paris Public and Private Schools". *Journal of Public Economics* 94 (1–2), pp. 59–77. DOI: 10.1016/j.jpubeco.2009.10.009.

Fuerst, F. and P. McAllister (2011). "Green Noise or Green Value? Measuring the Effects of Environmental Certification on Office Values". *Real Estate Economics* 39 (1), pp. 45–69. DOI: 10.1111/j.1540-6229.2010.00286.x.

Fuerst, F. et al. (2015). "Does Energy Efficiency Matter to Home-Buyers? An Investigation of EPC Ratings and Transaction Prices in England". *Energy Economics* 48, pp. 145–156. DOI: 10.1016/j.eneco.2014.12.012.

— (2016). "Energy Performance Ratings and House Prices in Wales: An Empirical Study". *Energy Policy* 92, pp. 20–33. DOI: 10.1016/j.enpol.2016.01.024.

Gibbons, S., H. Overman, and M. Sarvimäki (2021). "The Local Economic Impacts of Regeneration Projects: Evidence from UK's Single Regeneration Budget". *Journal of Urban Economics* 125, p. 103315. DOI: 10.1016/j.jue.2020.103315.

Hausman, J. A., B. H. Hall, and Z. Griliches (1984). "Econometric Models for Count Data with an Application to the Patents–R&D Relationship". *Econometrica* 52 (4), pp. 909–938.

Huang, Z. and X. Du (2021). "Does Air Pollution Affect Investor Cognition and Land Valuation? Evidence from the Chinese Land Market". *Real Estate Economics* 49 (4), pp. 1042–1074. DOI: 10.1111/1540-6229.12344.

Koengkan, M. and J. A. Fuinhas (2022). "Heterogeneous Effect of 'Eco-Friendly' Dwellings on Transaction Prices in Real Estate Market in Portugal". *Energies* 15 (18), p. 6784. DOI: 10.3390/en15186784.

Kok, N. and M. Jennen (2012). "The Impact of Energy Labels and Accessibility on Office Rents". *Energy Policy* 46, pp. 489–497. DOI: 10.1016/j.enpol.2012.04.015.

Kortas, F., A. Grigoriev, and G. Piccillo (2022). "Exploring Multi-Scale Variability in Hotspot Mapping: A Case Study on Housing Prices and Crime Occurrences in Heerlen". *Cities* 129, p. 103814. DOI: 10.1016/j.cities.2022.103814.

Kranzl, L. et al. (2020). *X-Tendo: Extending the Energy Performance Assessment and Certification Schemes via a Modular Approach*. Tech. rep. Grant Agreement No. 845958. Buildings Performance Institute Europe (BPIE). DOI: 10.3030/845958.

Kuethe, T. H. and S. Coggeshall (2010). "Green Certification and House Prices: Evidence from the Chicago Housing Market". *Journal of Real Estate Research* 32 (4), pp. 435–452.

Liang, W., M. Lu, and H. Zhang (2016). "Housing Prices Raise Wages: Estimating the Unexpected Effects of Land Supply Regulation in China". *Journal of Housing Economics* 34, pp. 21–38. DOI: 10.1016/j.jhe.2016.07.002.

Ou, Y. et al. (2025). "The Price Premium of Residential Energy Performance Certificates: A Scoping Review of the European Literature". *Energy and Buildings* 301, p. 113700. DOI: 10.1016/j.enbuild.2025.115377.

PORDATA (2023). *Population living in dwellings in poor condition, Portugal and European Union*. https://www.pordata.pt/pt/estatisticas/habitacao-e-familias/habitacao/populacao-que-vive-em-alojamentos-em-mas-condicoes. Accessed: 2025-07-18.

Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition". *Journal of Political Economy* 82 (1), pp. 34–55.

Tai, M. Y. et al. (2017). "Foreign Buyers and Housing Price Dynamics". *International Review of Economics Finance* 52, pp. 65–74. DOI: 10.1016/j.iref.2017.03.013.

White, H. (1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity". *Econometrica* 48 (4), pp. 817–838.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. 2nd. Cambridge, MA: MIT Press.

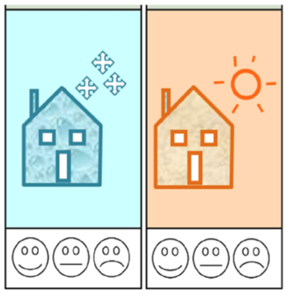
Yoshida, J. and T. Sugiura (2016). "Quantile Treatment Effects of Energy Performance Certificates in the Real Estate Market". *Energy Economics* 55, pp. 278–288.

Zhou, Q. et al. (2019). "Do Housing Prices Promote Total Factor Productivity? Evidence from Spatial Panel Data Models in Explaining the Mediating Role of Population Density". *Land Use Policy* 91, p. 104410. DOI: 10.1016/j.landusepol.2019.104410.

A How EPCs Differ Across Europe

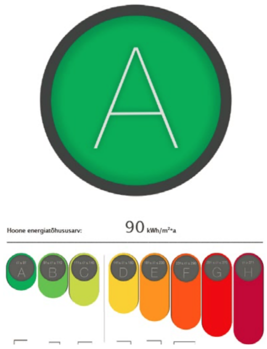
(a) Extract from Portugal's EPC

Source: Agência para a Energia (ADENE), as shown in Kranzl et al. (2020).



(b) User-friendly Italian EPC design

Source: Ministero della Transizione Ecologica (MiTE), as shown in Kranzl et al. (2020).

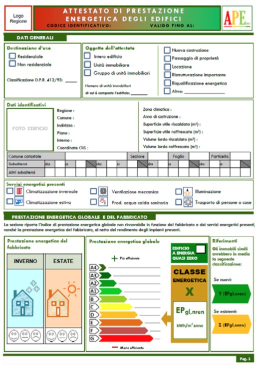


(c) Extract from Estonia's EPC

Source: Enefit, as shown in Kranzl et al. (2020).

(d) EPC Label in Germany

Source: German Energy Agency (dena).



(e) EPC Label in Italy

Source: MiTE, as shown in Kranzl et al. (2020).

(f) EPC Label in Portugal

Source: ADENE, as shown in Kranzl et al. (2020).

Figure 4: Comparison of EPC Label Designs in Estonia, Germany, Italy, and Portugal

B Panel Variation and Correlations

This appendix presents descriptive statistics and correlation information for the key variables in the panel dataset.

Table 5: Summary Statistics of Panel Data Variation

Variable	Overall Mean	Between Std. Dev.	Within Std. Dev.	Obs.
sqmprice	905.861	574.681	140.609	1,077
ln_sqmprice	6.631	0.578	0.128	1,077
greenepcs	0.339	0.129	0.057	1,112
ln_meanwage	6.999	0.138	0.063	1,112
empeduhighratio	0.479	0.083	0.026	1,112
unemprate	0.164	0.074	0.022	1,112
ln_popdens	4.309	1.544	0.012	1,112
firmsprop	0.133	0.034	0.006	1,112
repcrimesprop	0.028	0.009	0.004	1,112
foreignpopprop	0.052	0.062	0.011	1,112
studentsprop	0.130	0.030	0.003	1,112
schoolprop	0.093	0.033	0.005	1,112

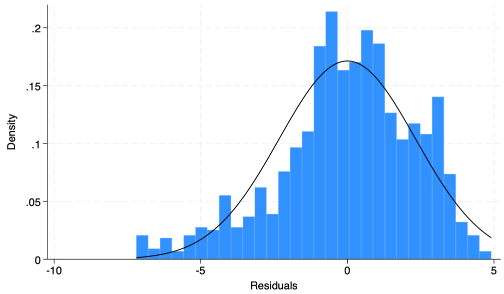
Note: "Between" is the cross-sectional variation, "Within" is the variation over time within municipalities.
Source: Own elaboration based on outputs from Stata (`xtsum`).

Table 6: Pairwise Correlation Matrix for Key Variables

	sqm	l_sqm	green	l_wage	edu	un-emp	l_pop	firms	crimes	for-eign	stu	sch
sqm	1.00											
l_sqm	0.93*	1.00										
green	0.26*	0.32*	1.00									
l_wage	0.52*	0.52*	0.04	1.00								
edu	0.51*	0.50*	0.09*	0.53*	1.00							
unemp	-0.27*	-0.30*	-0.19*	-0.48*	-0.28*	1.00						
l_pop	0.55*	0.61*	0.32*	0.46*	0.43*	-0.22*	1.00					
firms	0.18*	0.02	0.09*	-0.11*	0.11*	0.20*	-0.19*	1.00				
crimes	0.41*	0.33*	-0.16*	0.27*	0.22*	-0.04	-0.04	0.21*	1.00			
foreign	0.58*	0.57*	0.01	0.23*	0.24*	-0.16*	0.19*	0.26*	0.53*	1.00		
stu	0.54*	0.51*	0.18*	0.47*	0.40*	-0.21*	0.52*	-0.15*	0.21*	0.24*	1.00	
sch	0.31*	0.32*	-0.23*	-0.18*	-0.15*	0.07*	-0.42*	-0.01	0.04	-0.12*	-0.17*	1.00

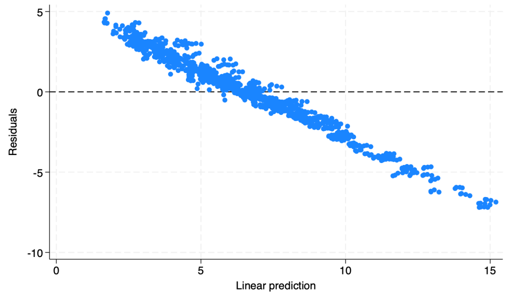
Note: * indicates significance at the 5% level. Abbreviations: sqm = sqmprice, l_sqm = ln_sqmprice, green = greenepcs, l_wage = ln_meanwage, edu = empeduhighratio, unemp = unemprate, l_pop = ln_popdens, firms = firmsprop, crimes = repcrimesprop, foreign = foreignpopprop, stu = studentsprop, sch = schoolprop.
Source: Own elaboration based on outputs from Stata.

C Model Specification and Diagnostic Checks



(a) Histogram of residuals from the Fixed Effects model

Source: Own elaboration using Stata, based on model estimation results.



(b) Residuals vs fitted values (Fixed Effects model)

Source: Own elaboration using Stata, based on model estimation results.

Figure 5: Residual diagnostics of the Fixed Effects model

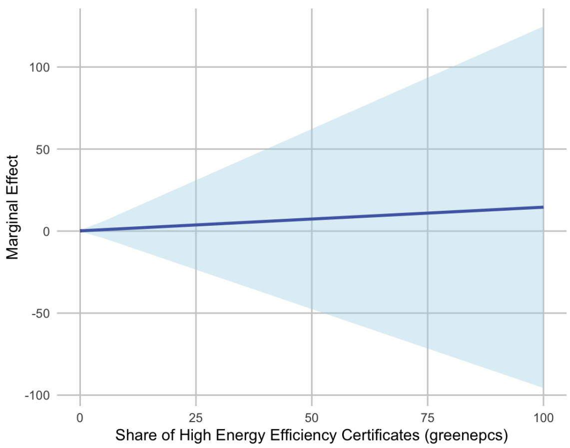


Figure 6: Marginal effects of the quadratic term ($greenepcs^2$) in the Fixed Effects model

Source: Own elaboration using RStudio, based on marginal effects analysis.