Analysis of the Determining Factors for the Renovation of the Walloon Residential Building Stock

Guirec Ruellan 1,*, Mario Cools 2,3,4 and Shady Attia 1

1 Sustainable Building Design Lab, Department UEE, Faculty of Applied Sciences, Université de Liège, 4000 Liège, Belgium; shady.attia@uliege.be
2 Local Environment Management & Analysis, Department UEE, Faculty of Applied Sciences, Université de Liège, 4000 Liège, Belgium; mario.cools@uliege.be
3 Faculty of Business Economics, Hasselt University, 3590 Diepenbeek, Belgium
4 Centre for Information Management, Modeling and Simulation (CIMS), KU Leuven Campus Brussels, 1000 Brussels, Belgium
* Correspondence: guirec.ruellan@uliege.be

Abstract: The issue of energy retrofitting of existing building stock occupies an increasingly prominent place in energy transition strategies in Europe. Adopting models representing the building stock and accounting for occupancy influence on final housing energy use must be developed to advise new policies. In this respect, this study aims to characterize the Walloon residential building stock and to analyze the existing correlations between the stock’s technical data and its occupants’ socioeconomic data. This study uses existing databases on buildings and inhabitants in Wallonia. Several statistical analyses make it possible to highlight the preponderant criteria and existing correlations between these different criteria. This study affirms the importance of accounting for certain socioeconomic categories, such as low-income groups, in a global strategic reflection on energy renovation. Multiple linear regression shows us that each percent increase in the category of households that declare between 10,000–20,000 EUR of income per year corresponds to an increase of 7.22 kWh/m²·y in the average energy efficiency of the built stock. The results highlight the importance of focusing on renovation strategies for particular types of buildings, such as semi-detached houses, which combine unfavorable technical and socioeconomic factors. Thus, the results confirm the interest of a mixed model approach to adapt to effective renovation policy strategies.

Keywords: sustainable renovation; energy policies; statistical analysis; Walloon region; bottom-up model

1. Introduction

Recent developments in European directives addressing energy issues in buildings (Energy Performance of Buildings Directive (EPBD) 2010/31/EU [1], Energy Efficiency Directive 2012/27/EU [2], and the Directive amending the Energy Performance of Buildings Directive 2018/844 [3]) reflect the importance given in the sector to meeting the objectives of reducing energy consumption and greenhouse gas emissions. Belgium is characterized by an ancient building stock (32% of the housing stock was built before 1945 against 23% for the European average), high proportion of single-family houses (76% in Belgium against 53% for the European average), and low renovation rates that lag behind that of neighboring countries (e.g., 58% of Belgian homes have roof insulation compared to 73% in the UK) [4]. In 2017, the residential sector’s direct energy consumption represented 20% of Belgian energy consumption [5].

For a long time, national regulations focused on new construction to the detriment of renovation [6]. Thanks to the energy renovation campaigns [7], renovation is slowly entering the political field [8,9]. Home energy retrofitting is crucial as it reduces energy
consumption [10] and integrates multiple sustainable development aspects [11]. A European research project from the Eracobuild program thus lists five categories of drivers: (1) durability/building physics, (2) economy, (3) environment, (4) comfort, and (5) others [12].

Despite this clear awareness, all indicators showed little change in the energy renovation rate over the last decade [13], and it is far from the estimated target of approximately 5% [4]. Much research and development on energy retrofitting has focused on the evaluation of technical solutions [14]. In Europe, this has resulted in policies focused on technical aspects [15], and few studies address the development of strategies that encourage retrofitting and improve process productivity [12]. The role of housing energy renovation policies is changing [9], no longer focused exclusively on reducing energy consumption. They now focus more on meeting the needs of the population [16]. Energy renovation policies appear as real social policies [17].

The principal barrier, at all levels, is financial aspects [14]. Consequently, subsidies, incentives, or new business models must be created that highlight the project’s environmental factors. Because of the lock-in effects of unambitious policies, improvements to buildings’ energy efficiency can be stranded below the best standards [14]. Nielsen et al. [18] voiced the need for more research to better target existing buildings and to prioritize actions following the built stock’s characteristics. Walter and Sohn [19] suggested that the most cost-effective renovations should be identified and that an estimated confidence in expected savings should be provided. Calculating the anticipated savings implies the capacity to estimate uncertainties related to climate, behavior, specific building characteristics, and the interactions between these effects. In their literature review, Kavgic et al. [20] considered that the synergies between different policies in favor of energy efficiency, health, and comfort could be improved using adapted models of the built stock.

Two distinct approaches are used to model energy consumption in the residential sector. Top-down models analyze macroeconomic variables to distribute the different uses. Bottom-up models generalize users’ final consumption to extrapolate the total energy consumption on a larger scale [21]. We distinguish between statistical and physical (or engineering) models among the bottom-up models [20]. Top-down models rely on past projections to forecast future trends. They lack detail at the technological level and are not very effective for evaluating technology-centered policies. Bottom-up statistical models require large samples, are inflexible, have difficulty predicting the impact of consumption reduction measures, and are subject to multi-collinearity. Finally, physical bottom-up models require excessive amounts of technical data, do not consider macroeconomic effects, and do not directly assess the impact of human behavior. A few hybrid models were developed to combine physical and statistical models [20].

Cyx et al. [22] distinguished approaches based on the selection of representative and typical buildings. The first case allows for a better approach to the average characteristics of the stock. Still, the second allows for better study of technical solutions adapted to the specificities of each building typology.

We note that the different models developed often have difficulty accounting for energy demand’s sociotechnical factors: how people consume energy in their homes and how they invest in energy savings [20]. There are significant differences in consumption between two dwellings with the same performance, depending on the inhabitants’ behavior [15,23]. Numerous studies have shown how socioeconomic conditions affect energy performance and investments in the field [24]. This is the case for income [25], education [26], or age [27]. The prebound [28] and rebound [29] effects directly translate the impact of behavior on energy efficiency.

By definition, sustainable renovation must include a social component [11]. The risk is that high-performance renovation is reserved for the wealthiest [30] and owners [31–33]. Energy and housing policies must be designed to encourage sustainable renovation projects and to limit rent increases [30]. Energy renovation, which has traditionally focused on the efficiency of heating and lighting systems, tends to evolve towards a more global approach that integrates social benefits [12].
In Belgium, several studies were carried out to model different energy aspects of the built stock. Verbeeck and Hens [34] studied the financial viability of different types of energy works on a countrywide scale. In 2008, Kints [35] proposed a typological classification of dwellings in the Walloon region (southern part of Belgium). The Low Energy Housing Retrofit research project [4] analyzed the characteristics of Belgian built stock to define target groups for energy renovation. The SuFiQuaD study [36] proposed a multifactorial evaluation of different types of buildings. The TABULA project [22] defined building types related to the Belgian housing stock, to which different renovation scenarios were applied. In 2015, Gendebien et al. [37] proposed a bottom-up model, mixing hybrid and representative approaches, to analyze the built stock’s energy consumption. The last two jointly analyzed models [38] highlight the importance of reliable and complete data when modeling the built stock to limit significant variations between models. A recent study estimated the heat consumption and heat demand of more than 1,700,000 buildings in the Walloon region using a geographic information system (GIS) [39]. However, all these studies focus on the physical aspects of the built stock. They do not consider the social factors essential to achieving a large-scale energy transition in Belgium [40].

The above review underlines the lack of a model based on the Walloon built stock to guide the elaboration of policies in favor of energy renovation that accounts for a household’s socioeconomic aspects. To make such a model, it is first necessary to select and characterize the appropriate indicators, indicators for which databases are accessible and between which correlations exist to build this model. Therefore, this study aims to analyze different building stock characteristics and to correlate them to their occupants’ socioeconomic conditions. The study aims to allow researcher and policymakers to create long-term renovation strategies informed by these study outcomes. The following questions are answered in this research:

- How to cross analyze building databases that do not represent the same statistical individuals?
- What types of correlations are there between the technical housing characteristics and the socioeconomic characteristics of the occupants? What is the value of these correlations?
- Which trend does the combination of technical and socioeconomic characteristics make the deduction of the built stock’s performance and the identification of representative typologies possible?
- What is the potential impact of these correlations on a building policy in favor of energy renovation?

Finally, this study highlights the essential links between the least affluent households and the least efficient buildings and finds that increased homeownership is accompanied by decreased energy efficiency. Developing energy renovation policies targeted more towards certain groups identified through this analysis could be useful.

2. Methodology

In this section, we present the research methodology, including the study concept. Our research methodology combines a literature review with three main statistical analysis methods. We developed a study conceptual framework that summarizes and visualizes our research methodology, as shown in Figure 1.
Figure 1. Study conceptual framework.

2.1. Boundary Conditions and Data Collection

2.1.1. Boundary Conditions

The boundary conditions are defined in the literature review analysis. As the study’s objective is to analyze the links between technical and socioeconomic characteristics, choosing relevant variables and avoiding excessive variables is essential to simplify the results and to facilitate straightforward interpretation [41].

Two technical characteristics were studied. Our study relies on the data collected within the Energy Performance Certificate (EPC) because the energy efficiency of a dwelling partly conditions its consumption. In this study, we make the same choice as other similar studies in the USA [19], Switzerland [42], Sweden [24], or the Netherlands [15].

The geometry of the dwelling ((1) an apartment, (2) a house with 4 facades, (3) a semi-detached house, or (4) an attached house) is the second characteristic studied to improve EPC quality [24,43,44]; it directly impacts the energy needed to heat the dwelling and indirectly impacts the urban fabric and, therefore, the energy required for household activities [45].

Two socioeconomic characteristics were also studied. Household income not only has a significant impact on the quality and size of housing but also influences energy behavior [27] and investing in energy efficiency [24,25]. The last characteristic studied was ownership. Depending on its status, the occupant does not have the same motivations and possibilities to renovate his dwelling [32,40,46]. The Walloon Region was chosen as a case study because energy regulations vary slightly between Belgium’s three regions (Brussels, Flanders, and Wallonia). Each entity is autonomous to set its own measures (for example, the expected insulation levels for each wall) within the framework of the EPBD [44].

The different datasets were studied on the Walloon municipalities’ scale, with each municipality constituting a statistical individual. This analysis scale makes it possible to
reduce the size of the databases compared to a by-building analysis. Databases are easier to access by the city when they are not the only ones. However, the scale should remain small enough to detect variations between individuals. Finally, the municipal scale avoids any confidentiality issues. The study at the city scale adds the number of households in each city, allowing for identifying possible statistical differences between small villages and large cities.

2.1.2. Data Collection

The chosen study characteristics lead to a search for existing databases that meet the criteria above. The use of existing databases (Table 1) speeds up the study to test the methodology with data already validated elsewhere and to obtain databases representing the entire studied population.

Table 1. Summary table of data.

<table>
<thead>
<tr>
<th>Category Name</th>
<th>Data Name</th>
<th>Source</th>
<th>Data Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>EPC label and consumption</td>
<td>DG TLPE</td>
<td>2018</td>
</tr>
<tr>
<td>Geometry</td>
<td>Type of building in housing</td>
<td>STATBEL</td>
<td>2018</td>
</tr>
<tr>
<td>Income</td>
<td>Categories of income</td>
<td>IWEPS</td>
<td>2016</td>
</tr>
<tr>
<td>Ownership</td>
<td>Property title</td>
<td>Census</td>
<td>2011</td>
</tr>
<tr>
<td>Households</td>
<td>Number of households</td>
<td>STATBEL</td>
<td>2018</td>
</tr>
</tbody>
</table>

The EPC data is the EPBD application for the Walloon region and comes from the DG TLPE (Direction Général Territoire, Logement, Patrimoine et Environnement), part of the Walloon government. For each EPC certificate, we obtained the city, destination (single-family house or apartment), specific energy (kWh/m²·y), and label (from A++ to G). The database contains information on 495,470 dwellings, which represents almost one third of the Walloon building stock. Since 2010, only buildings that need to be sold had to have an EPC. The other two-thirds of the building stock did not have an EPC label A. A previous study by CEHD [47] conducted in 2017 proved the representativeness of the half million EPCs. Despite the statistical overrepresentation of apartments and underrepresentation of single-family houses, the database represents the seven EPC categories—A to G—in a balanced way. Therefore, the current EPC database provides a useful snapshot on the building stock characteristics and measures the energy efficiency tendencies. The database obtained is sampled to obtain the percentages of the different labels and the average consumption of certificates in each municipality. The sum of these percentages equals 100% and represents the entire housing stock.

The type of housing building is public data published yearly by STATBEL, the Belgian Statistical Office. We have the percentage of terraced houses, semi-detached houses, detached houses, and apartments for each municipality. These percentages equals 100% and represent the entire stock of 1,615,774 dwellings listed by STATBEL.

Income level is a public classification of annually reported income into 6 categories by the IWEPS (Institut Wallon de l’Evaluation, de la Prospective et de la Statistique), the Walloon statistical office. For each municipality, we have the percentage of each category. The sum of these percentages equals 100% and represents all households.

The property title represents the proportion of owners and tenants in each municipality. The public data come from the Census, the population census database conducted by the Belgian government.

The households of the cities represent the sum of income declared by each municipality (statistical individuals). These are public data published annually by STATBEL.

As we can see, the data were not studied at the scale of the building but the municipality scale. This data compression allows for the aggregation of independent databases. Each municipality represented one of the 262 statistical individuals, which is the sample size.
These data were collected by government agencies and met the highest quality standards.

2.2. Data Processing

All the statistical tests were performed and programmed on the R software.

2.2.1. Characterization

Characterization, the first step (Figure 1) of data processing, provides knowledge of the studied statistical variables of the Walloon built stock and details the variability of their distribution.

The collected data were aggregated in a single table to simplify statistical processing. This also allowed for the detection of possible errors, such as changes in the individuals’ names concerned. Because the Walloon region is bilingual (French-German), municipalities may have several names.

This sampling makes it possible to analyze the geographical distribution of each variable independently of the others. Comparing the obtained data with state-of-the-art results and comparisons of the results between different cities, provinces, and neighboring regions and countries is an interesting exercise.

This first part of the analysis was done by combining municipalities across the province to simplify exploiting the results in the first instance. It makes it possible to validate the study’s hypothesis, which is based on a sufficiently large variability of the criteria studied between different geographical entities.

The first results were expounded upon in search of correlations.

2.2.2. Correlation

Studying the correlations between the different characteristics is the second stage of data processing. This step consists of determining the directing coefficient of a simple linear regression between two characteristics to analyze the relationships between characteristics.

To simplify the research of correlations, labels A, A+, and A++ were merged into a single category A. The mean energy consumption estimated by the EPC (epcMean) was calculated for each municipality.

The Spearman correlation coefficient could be calculated to explain the linear correlation between the different statistical variables.

A linear regression could further be performed after obtaining the correlation results.

2.2.3. Multiple Linear Regression

An MLR (multiple linear regression) objective was conducted to obtain a model of the relationship between a dependent variable and several explanatory variables, allowing for identification of each explanatory variable’s impact on the dependent variable. The dependent variable is the average energy efficiency (epcMean) of the certified buildings in each city and better represents the entire housing stock than the proportion of EPC A or G. Thus, we can estimate that the model’s unit is the kWh/m\(^2\)·y.

The independent variables are the different housing types, income categories, ownership rate, and relative number of households in the city. For each category where the sum of the variables is the same for all individuals (e.g., equal to 100%), there was one redundant variable. To avoid redundancy, one variable was eliminated. The choice was made by simulating various combinations of explanatory variables and by keeping the combination where the variance influence factor (VIF), i.e., the collinearity, of the different explanatory variables was the lowest. The proportions of detached houses, income category 6, and renters were redundant information and were not considered explanatory variables.

The first three explanatory variables were the proportion of apartments, semi-detached houses, and attached houses for geometry. The following explanatory variables were the proportions of income categories 1–5 for income. The ownership rate in each city was another explanatory variable. The final explanatory variable was the relative number of house-
holds in the city. It was obtained by dividing the total number of households in the city by the number of households in the region’s largest city (Liège households = 114,115 households). This modification gives a variable of the same size as the other explanatory variables (≤1). We obtained a total of ten explanatory variables.

A calculation of the Mahalanobis distance was performed ahead of the MLR to detect and analyze outliers with an excessive influence on the dataset to assure the analysis’s robustness.

The results were also tested for hypothesis validity. The Breusch–Pagan test verified the homoscedasticity of the errors. The variance influence factor (VIF) allowed for the estimation of multicollinearity. The normal distribution of the residues was verified using the Shapiro–Wilk test.

2.3. Instrument Development

2.3.1. Cross Analysis

The results obtained during the various stages of this methodology provide a better understanding of the composition of the built stock, confirm or correct the results previously found in other studies, and better understand the correlation between different factors that have not yet been compared.

2.3.2. Simulation

The theoretical model obtained by MLR allows us to express an estimate of the value of the dependent variable (the energy efficiency of the built stock) as a function of the value of the different explanatory variables (of this built stock). The exploitation of the results allows the simulation of one or several renovation scenarios to demonstrate their impact on the Walloon Region. This is a theoretical estimate to be taken with the usual precautions.

2.4. Recommendation

A more open recommendation stage will attempt to use the results obtained and the analysis that will have been made to promote several recommendations to better target the buildings to be renovated, the populations to be assisted, and the type of assistance to be provided.

3. Results

This study’s focus is the Walloon Region in Belgium and includes five provinces: Hainaut, Liège, Luxembourg, Namur, and Walloon Brabant (see Figure 2). The Walloon Region has its energy performance regulations for building and calculation methodologies (EPC) in line with the European EPBD.

The study parameters chosen to characterize the building stock are based on three primary criteria:

• The estimated impact of the parameters on energy use
• The estimated impact of the parameters on the choice of energy-efficient renovation
• The availability of representative data on the study parameters at the Walloon municipality level

The collected data are grouped in a database listing each parameter’s rate in the different Walloon municipalities to provide a view by province and for the whole region. Table 2 describes the variables used afterwards, synthesizing their minimum, mean, and maximum values.
Figure 2. Walloon region map with provincial boundary (d-maps, 2020).

Table 2. Synthesis table of the separate results for each part.

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>epcA</td>
<td>0%</td>
<td>0.88%</td>
<td>6.29%</td>
</tr>
<tr>
<td>epcB</td>
<td>1.4%</td>
<td>8.04%</td>
<td>26%</td>
</tr>
<tr>
<td>epcC</td>
<td>4.88%</td>
<td>12.78%</td>
<td>23.81%</td>
</tr>
<tr>
<td>epcD</td>
<td>7.74%</td>
<td>15.09%</td>
<td>23.83%</td>
</tr>
<tr>
<td>epcE</td>
<td>10.84%</td>
<td>15.93%</td>
<td>23.68%</td>
</tr>
<tr>
<td>epcF</td>
<td>8.83%</td>
<td>14.81%</td>
<td>20.22%</td>
</tr>
<tr>
<td>epcG</td>
<td>10.77%</td>
<td>32.48%</td>
<td>57.74%</td>
</tr>
<tr>
<td>epcMean</td>
<td>297.5</td>
<td>443.5</td>
<td>607.6</td>
</tr>
<tr>
<td>apartments</td>
<td>0.36%</td>
<td>9.56%</td>
<td>61.45%</td>
</tr>
<tr>
<td>attached</td>
<td>1.14%</td>
<td>18.68%</td>
<td>56.93%</td>
</tr>
<tr>
<td>semidetached</td>
<td>9.27%</td>
<td>25.59%</td>
<td>40.70%</td>
</tr>
<tr>
<td>detached</td>
<td>3.4%</td>
<td>46.17%</td>
<td>80.81%</td>
</tr>
<tr>
<td>income1</td>
<td>10.32%</td>
<td>13.70%</td>
<td>20.51%</td>
</tr>
<tr>
<td>income2</td>
<td>14.89%</td>
<td>27.44%</td>
<td>41.17%</td>
</tr>
<tr>
<td>income3</td>
<td>16.42%</td>
<td>21.25%</td>
<td>25.96%</td>
</tr>
<tr>
<td>income4</td>
<td>10.27%</td>
<td>12.84%</td>
<td>15.95%</td>
</tr>
<tr>
<td>income5</td>
<td>4.96%</td>
<td>7.97%</td>
<td>11.11%</td>
</tr>
<tr>
<td>income6</td>
<td>5.6%</td>
<td>16.8%</td>
<td>34.91%</td>
</tr>
<tr>
<td>owner</td>
<td>48.06%</td>
<td>72.15%</td>
<td>88.22%</td>
</tr>
<tr>
<td>households</td>
<td>766</td>
<td>7743</td>
<td>114,115</td>
</tr>
<tr>
<td>relativehouseholds</td>
<td>0.67%</td>
<td>6.78%</td>
<td>100%</td>
</tr>
</tbody>
</table>

3.1. Characterization of the Walloon Built Stock

For this first part, the statistical individuals are grouped by province to facilitate the results’ graphical analysis. Figure 3 shows that Walloon Brabant has a particularly efficient housing stock, with a high number of EPC A, B, and C (37.52%). The provinces of Liège (45.89%) and Hainaut (49.93%) visibly suffer from an overrepresentation of EPC F and G certificates.
The distribution of building types can be categorized into two groups (Figure 4). The provinces of Liège (30.09%) and Hainaut (36.90%) are primarily comprised of attached houses. The provinces of Luxembourg (50.65%), Namur (43.29%), and Walloon Brabant (40.97%) have more detached dwellings. A higher proportion of apartments characterizes the provinces of Liège (19.29%) and Walloon Brabant (23.23%).

Figure 4. Geographical distribution of building type in the Walloon Region.

Figure 5 shows the characteristic homogeneity of the ownership rate distribution (62.25–68.63% owner-occupied houses). This homogeneity reflects the importance of regional and national policies in favor of homeownership. Liège and Hainaut have slightly more rented housing, whereas Namur, Luxembourg, and Walloon Brabant have more owner-occupied housing.
The income distribution in Figure 6 shows more marked variations, particularly in categories 2 (low income) and 6 (high income). Thus, Hainaut (33.27%) and Liège (30.20%) have a large low-income population. Conversely, Brabant-Wallon (21.75%), Luxembourg (17.77%), and Namur (16.25%) have a large high-income population. Category 1 (income of 1000–10,000 EUR) represents special occupants, such as students.

Figure 5. Geographical distribution of property title.

Finally, the history of the Walloon region is well-reflected in the different characteristics studied. Liège and Hainaut have significantly suffered from deindustrialization. They have dense built-up areas (attached houses and apartments); low-quality (EPC F and G), less well-off populations; and more tenants. Luxembourg and Walloon Brabant were urbanized more recently. Walloon Brabant benefits from its proximity to Brussels, and the province of Luxembourg benefits from its proximity to the Grand Duchy of Luxembourg. These poles of activity lead to the arrival of new occupants with higher than average incomes. New dwellings that meet the latest energy standards are then built to accommodate them, allowing for greater penetration of four-sided houses and built stock efficiency.

3.2. Correlation between Walloon Built Stock Characteristics

Figure 7 shows the Spearman correlation results in matrix form. Larger circles denote higher correlations, and lighter yellow circles indicate more positive correlations.
The proportions of EPC A, B, C, and D are positively correlated with each other and negatively correlated with EPC F and G’s ratios and with average energy efficiency values (epcMean). These correlations confirm that the geographical distribution of EPC certificates is not random. The low correlation between the proportion of EPC E and other parameters indicates its independence. These certificates can be found in all the configurations.

The proportion of apartments is positively correlated to EPC A, B, and C and negatively correlated to EPC F and G. In contrast, the proportion of semi-detached houses is positively correlated to EPC F and G and negatively correlated to EPC B, C, and D. The proportion of attached houses was also positively correlated with the proportion of label E. No significant correlation was found between the proportion of detached houses and the EPC labels. The greater compactness of the apartments partly explains their better energy performance. A higher proportion of apartments in newly built dwellings meet existing energy standards, and attached and semidetached houses (although more compact) are significantly less efficient than detached houses. Attached and semi-detached houses often encounter difficulties in insulating specific walls, such as street fronts or shared walls. Some attached and semi-detached houses are older dwellings of lower quality—called worker houses in Belgium.

Low-income categories 2 and 3 are positively correlated with EPBD certificates F and G. High-income categories 5 and 6 are positively correlated with EPBD certificates A, B, C, and D. Income category 1 appears to be a special category (students), insofar as it presents correlations opposite to that of income 2/3. One hypothesis is that the highest
income earners live in recently built or renovated housing. Another hypothesis is that higher-income earners live in larger areas, increasing energy efficiency per square meter. A third hypothesis is that low-income earners are more likely to perform energy retrofits themselves and would not provide proof of insulation during the certification process.

The rate of ownership is positively correlated with high income (categories 5 and 6) and detached houses and negatively correlated with the rate of apartments, attached houses, and low income (category 2). Numerous studies have found a negative correlation between the rate of ownership and the city population and a weak correlation between the rate of ownership and the rate of different energy labels. A higher ownership rate can explain the negative correlation for geometries with lower compactness (detached houses) and more energy efficiency work performed by the owners (without a contractor).

In summary, the correlations reveal three typologies that are widely represented in the Walloon region:

- Old worker houses in the industrial valley are more present in big cities. These attached or semi-attached houses are of lower quality and energy efficiency. Lower-income households occupy these houses; more than the average number of such houses are rented.
- Newer and more efficient apartments are primarily built in large cities and rented more than average.
- Four-sided houses in small towns are mostly owner-occupied by higher income.

3.3. Multiple Linear Regression between Walloon Built Stock Characteristics

An analysis of the MLR results highlights the simultaneous impact of the selected explanatory variables on each city’s built stock’s average energy efficiency. The model studied for the MLR can be summarized as follows in Equation (1).

\[ \text{epcMean} \sim \text{apartments} + \text{attached} + \text{semidetached} + \text{income1} + \text{income2} + \text{income3} + \text{income4} + \text{income5} + \text{owner} + \text{households} \]  

The dependent and independent variables used for this MLR are described in more detail in Table 2, with their minimum value, mean value, and maximum value.

The results of the MLR are presented in Table 3. The first results are the model quality and have a multiple R-squared = 0.6194. The independent variables explain almost 60% of the dependent variable’s variance, which is higher than that of other similar studies [15,19,26,27]. This shows the interest of the chosen parameters for explaining the average energy efficiency of the built stock. Each estimator can be interpreted as the share of energy efficiency (kWh/m²·y) of the built stock induced by a 100% increase in this variable.

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>90.31</td>
<td>116.78</td>
</tr>
<tr>
<td>apartments</td>
<td>−274.71</td>
<td>53.64</td>
</tr>
<tr>
<td>attached</td>
<td>−26.59</td>
<td>23.65</td>
</tr>
<tr>
<td>semidetached</td>
<td>141.27</td>
<td>39.12</td>
</tr>
<tr>
<td>income1</td>
<td>−492.34</td>
<td>244.31</td>
</tr>
<tr>
<td>income2</td>
<td>721.83</td>
<td>93.64</td>
</tr>
<tr>
<td>income3</td>
<td>19.82</td>
<td>171.57</td>
</tr>
<tr>
<td>income4</td>
<td>232.70</td>
<td>306.46</td>
</tr>
<tr>
<td>income5</td>
<td>117.75</td>
<td>372.37</td>
</tr>
<tr>
<td>owner</td>
<td>235.08</td>
<td>61.17</td>
</tr>
<tr>
<td>relativehouseholds</td>
<td>59.81</td>
<td>27.03</td>
</tr>
</tbody>
</table>

Residual standard error: 34.07 on 250 degrees of freedom

Multiple R² = 0.6194
Adjusted R² = 0.6041
Interestingly, the most precise estimates concern the proportion of apartments, semi-detached houses, income category 2 (income of 10,000–20,000 EUR), and owners. The proportion of income category 1 and the number of households in the city also have a p-value of less than 0.05, indicating that further analysis of the coefficients relating to these explanatory variables can be performed.

The variable relating to the proportion of apartments has a largely negative estimator (−274.71 kWh/m²·y). The proportion of semi-detached houses has a largely positive estimator (141.27 kWh/m²·y), confirming that more semi-detached houses decrease energy-efficiency. Conversely, more apartments seem to produce more energy-efficient housing stock.

The variable with the most important estimator is the proportion of income category 2 (from 10,000–20,000 EUR), strongly correlated (estimate = 721.83 kWh/m²·y) with poor energy efficiency. Each percent increase in this category corresponds to an increase of 7.22 kWh/m²·y in the built stock’s average energy efficiency. Additionally, this correlation is not linear, as household income increases. From income categories 3–5, the estimator remains very insignificant (p-value > 0.4). The variable corresponding to the proportion of income category 1 confirms that it is a special category (estimate = −492.34 kWh/m²·y). The negative estimator can be explained, among other things, by students living in apartments or by assisting spouses. The cases that correspond to this income category must be identified in future studies.

The ownership rate variable provides more information than the rest of the analysis. There was a weak correlation between the share of ownership and the share of different energy labels; the MLR indicates that the ownership rate and average energy efficiency vary in a similar way (estimate = 235.08 kWh/m²·y). The lack of direct correlation in the previous analyses can be explained by considering that an increase in the ownership rate in a city is correlated with a decrease in the proportion of income category 2.

Finally, the number of households also has a positive impact (estimate = 59.81 kWh/m²·y) on estimating the energy efficiency of buildings. Large cities have denser buildings (apartments and semi-detached houses). However, they also have older buildings and households with lower incomes. The city’s density favors building compactness and complicates insulation from the outside (external boundary wall, street alignment, etc.).

Figure 8 summarizes the mean and confidence interval of the regression model coefficient estimates obtained on R more visually.

Figure 8. Mean and 95% confidence interval of regression model coefficient estimates.
The MLR is preceded by a test step, allowing the exclusion of outliers. The calculation of the Mahalanobis distance highlights Ottignies-Louvain-la-Neuve. We calculated a Mahalanobis distance of 94.67 with a chi-square of approximately 18.31 with 95% probability for this city. The city has undergone profound changes following the relocation of the French-speaking part of the University of Louvain, leading to thousands of new apartments between 1973 and 1980. These outliers were removed for MLR. The MLR study was then performed on 262 Walloon cities.

Tests applied to the model verify that it meets the statistical assumptions. The null hypothesis of normality of the residuals was accepted according to the Shapiro–Wilk normality test, with a \( p \)-value of 0.056. The null hypothesis of homoscedasticity was accepted according to the studentized Breusch Pagan test, with a \( p \)-value of 0.222. The magnitude of multicollinearity was low, considering that the variance influence factor was less than 5.5 for all explanatory variables.

3.4. Results Extrapolation

3.4.1. Cross-Analysis

The concomitant analysis of the results of the three methods makes it possible to refine their learning. The main results of these analyses are summarized in Table 4. In this way, correlations that are confirmed throughout the entire study can be distinguished. Table 4 also shows some different results between the methods, which deserve a separate explanation.

**Table 4.** Synthesis table of the separate results for each part.

<table>
<thead>
<tr>
<th>Characterization</th>
<th>Energy</th>
<th>Geometry</th>
<th>Income</th>
<th>Ownership</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Extensive inefficient habitat. Very few efficient dwellings.</strong></td>
<td>Extensive inefficient habitat. Very few efficient dwellings.</td>
<td>All categories are well-represented.</td>
<td>High rate of low income (between 10,000–20,000 EUR)</td>
<td>High rate of owner Geographical homogeneity</td>
</tr>
<tr>
<td><strong>Geographical heterogeneity</strong></td>
<td></td>
<td>Geographical heterogeneity</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Correlation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>There is little correlation between ownership and energy efficiency. Rented properties seem to be more likely to be flats or semi-detached houses, occupied by low-income occupants in large cities.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Linear regression</strong></td>
<td>Linear regression highlights geometry’s importance in determining energy consumption, with more efficient apartments and very inefficient semi-detached houses. Household income is only important for income group 1, who live in more efficient housing, and income group 2, who live in very inefficient housing. Paradoxically, homeownership has a negative effect on energy efficiency.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Concerning the analysis of the distribution of energy labels, it is possible to distinguish an important stratification of labels according to the city. The most efficient dwellings are not spread evenly over the entire region. According to the MLR, cities with more apartments have more energy-efficient stock. Conversely, the number of semi-middle-sized houses and households and the number of homeowners are correlated with a decreased...
average efficiency of built stock. Therefore, one way to increase the built stock’s efficiency would be to make targeted proposals aimed at these categories of housing or households.

By studying the distribution of geometries, we see great variation from one region and city to another. The correlation between the number of households in the city and the proportion of different geometries responds to densification logic. The different correlations reflect the impact of the history of urbanization in the Walloon region. Some more compact geometries, such as apartments, are also more efficient. However, this is not true when comparing attached, semi-detached, and detached houses. The latter are less compact but more energy efficient than the first two. The MLR confirms that apartments are more efficient than average buildings and that targeting semi-detached houses is essential. This may lead to questions about why semi-detached houses are so different from attached and detached houses.

Income category 2 includes all households with a net taxable income from 10,000–20,000 EUR. All analyses show the particular importance that must be given to this part of the population to improve the built stock’s efficiency. Characterization shows that the most represented segment of the population lives in less efficient housing than average and, according to the MLR, is the primary indicator of the low energy efficiency of the built stock. This also implies a higher exposure to fuel poverty risk and fewer means of improving the situation.

The analysis of the results on the ownership rate is more ambivalent. The characterization highlights the homogeneity of the ownership rate, and the correlation confirms little direct relationship between the efficiency of the built stock and ownership rate. However, the MLR shows that increased ownership rates correspond to a significant decrease in the built stock’s energy efficiency. Other characteristics correlated with ownership rate tend to weaken the direct correlation. An increase in the number of high-income households is accompanied by a parallel improvement in energy efficiency and increased ownership rate. This correlation is significant as it is the opposite of the expected results. The tenants were assumed to live in less efficient dwellings because of split incentives, but this is not the case.

3.4.2. Prospective Simulation

Obtaining a model from MLR allows for analysis of the estimators. It can be directly applied to other statistical individuals (e.g., other sets of buildings) or on a regional scale to calculate the different explanatory variables’ total impact instead of the relative impact. Therefore, the average number of households (rather than the total) in the Walloon municipalities is considered.

For example, different model variables can be applied to the values corresponding to the Walloon Region’s different provinces. We obtained a prediction range that can be compared with the observations in Table 5.

| Table 5. Predicted value of the model with a 95% confidence interval and observed value. |
|-----------------------------------------------|---------------|---------------|---------------|---------------|
|                                 | Lower Value  | Mean Value    | Upper Value   | Observed Value |
|                                 | (kWh/m²·y)   | (kWh/m²·y)   | (kWh/m²·y)   | (kWh/m²·y)    |
| Walloon Brabant province        | 330.16       | 343.14        | 356.12        | 349.53        |
| Hainaut province                | 441.73       | 450.03        | 458.33        | 454.64        |
| Liège province                  | 390.64       | 400.32        | 410.01        | 434.69        |
| Luxembourg province             | 400.94       | 408.09        | 415.25        | 405.68        |
| Namur province                  | 409.63       | 416.59        | 423.56        | 420.35        |
| Walloon region                  | 407.32       | 415.02        | 422.72        | 427.50        |
This comparison logically allows us to obtain very close results because they come from the same databases. Nevertheless, Table 5 shows that the model could explain the significant variations in the built stock’s average energy efficiency between provinces.

4. Discussion

4.1. Findings and Recommendations

This research aimed to characterize Walloon’s energy efficiency and to quantify the existing correlations with the socioeconomic factors that influence the renovation decision-making of house owners. The overall aim is to help policymakers propose long-term-renovation strategies adapted to both the buildings and their inhabitants.

The building stock characterization revealed the heterogeneity of building energy efficiency among the Walloon region’s municipalities, with more than 45% of energy-inefficient houses (labels F and G) in Liège and Hainaut and more than 35% of energy-efficient homes (label A, B, and C) in Walloon Brabant. The correlation analysis shows that apartments, overall, are more energy-efficient. Semi-detached houses are correlated with lower energy efficiency compared to detached and terraced houses.

The MLR provides more novel results. The literature on split incentives suggested that owned residences were on average 6% more efficient than rented homes [33] due to the split incentive [31]. On the contrary, the MLR highlights the correlation between increased ownership rate and decreased energy efficiency. Each percent increase in the ownership category corresponds to an increase of 2.35 kWh/m²·y in the built stock’s average energy efficiency. Even if the owner-tenant dilemma exists [31], it is not currently a determining factor in the energy efficiency of the Walloon built stock. The landlord/tenant dilemma occurs when tenants are not allowed to own the building services [32].

Furthermore, the study confirms the positive correlation between some household income categories and housing energy efficiency. Very low-income groups (10,000–20,000 EUR of net annual taxable income) occupy much less efficient housing than the rest of the population. Each percent increase in this category corresponds to a rise of 7.22 kWh/m²·y in the built stock’s average energy efficiency. Nevertheless, this correlation between income and energy efficiency is not linear at all, contrary to the results obtained in other studies [25]. The differences observed between the middle and high-income groups are not very significant. Thus, there is a real threshold effect.

Therefore, a targeted renovation policy needs to be established with a particular focus on low-income households (from 10,000–20,000 EUR of net taxable income). These households must be the subject of complementary studies to characterize them deeper and to identify the effective actions to be implemented. Heavy renovation scenarios should be studied specifically for less energy-efficient building archetypes. Demolition/reconstruction scenarios could be proposed for these lower-quality archetypes. Energy regulations for buildings should be established on a broad geographical scale to subject everyone to the same rules and to close the disparity gap. Therefore, it is essential to adapt local policies in favor of renovation according to building owners and tenants’ built and socioeconomic contexts.

4.2. Strength and Limitations

This study’s principal contribution is cross-referencing buildings’ technical data (energy performance and archetypes) with socioeconomic data of the inhabitants (income and property title). An increasing number of studies found in literature analyzed the social aspects of the inhabitants and the technical aspects of buildings separately—the two determining factors in improving the built stock. However, we could not find a building stock model that integrates the inhabitants’ social aspects and to couple them such as our work on the national level [22,36,37]. Simultaneously, statistical models and studies that study these correlations at the international level rely on aggregated data sets that are usually based on questionnaires, which require much time to collect [15,27]. More and more databases can be used to improve our understanding of the characteristics of the built stock. However, these databases are complicated to link because they do not
always relate to the same statistical individual, whether it is a building, a person, a city, or a province. Therefore, this study succeeded in performing a unique statistical analysis that crosses different sets of data. The use of municipal data allowed us to correlate data on the individual level, resulting in an acceptable level of data quality.

Moreover, the MLR formula succeeded in explaining with 62% certainty the variation of the dependent variable (energy) concerning the independent variables (geometry, income, and ownership). Reducing the uncertainty to less than 50% is a significant contribution of this study because it allows us to explain the correlation between energy efficiency and the socioeconomic factors with high confidence. For example, similar studies in the Netherlands [15], USA [26], and China [27] investigated the correlation between the technical aspects of the building and the socioeconomic aspects of the inhabitant, but with lower certainty. They obtain an $R^2$ below 0.4 for their regression model.

In our case, the income2 category (10,000–20,000 EUR of net annual taxable income) negatively correlates with energy efficiency, while the apartment category has a positive correlation with energy efficiency. Thus, these evidence-based results can help decision-makers set their priorities when designing short-term renovation strategies.

However, the study considers neither the building’s age, which is an essential criterion for estimating buildings’ energy efficiency in Belgium, nor the house size. Unfortunately, these data were not available in a reference frame similar to that of the study.

Finally, the study provides a better understanding of Wallonia’s building stock, and the results help propose a data-driven adaptation policy in favor of energy renovation. Moreover, this study’s statistical methods can be transferred or applied to other data sets or case studies.

4.3. Implication on Practice and Future Research

This study can enable policymakers and local authorities to understand better the interactions between the building’s energy efficiency characteristics and socioeconomic factors. Those factors must be considered when developing and implementing renovation strategies for the built stock. Local governments will have better knowledge of buildings’ stock in their city and may encourage targeted initiatives for the least efficient buildings.

However, this study needs to be completed and improved. The integration of new study parameters, such as the dwelling’s age and size, will result in a better understanding of the interactions’ between different factors. The age of the buildings provides plentiful information on the type of construction and the standards in force. Investigating the building vintage would deepen and clarify the study of correlations. Additionally, a clustering study should follow this work to categorize existing buildings better.

5. Conclusions

The presented paper characterizes the Walloon building stock and analyses the existing correlations between the building stock energy efficiency and its occupant socioeconomic factors. The overall aim is to guide policymakers, professionals, and scientists towards an effective renovation strategy that can better cater to occupants’ socioeconomic conditions. In this context, the study succeeded in quantifying the correlation between household income and building energy efficiency states. The study combines four different databases on different scales and spatial levels, analyzing them using multiple linear regression. The building stock characterization allowed us to reveal the heterogeneity of building energy efficiency among the Walloon region’s municipalities, with more than 45% of energy-inefficient houses (labels F and G) in Liège and Hainaut and more than 35% of energy-efficient homes (label A, B, and C) in Walloon Brabant. Conversely, the MLR highlights the positive correlation between increased ownership rate and decreased energy efficiency. Belgium underwent a massive suburbanization in the last 70 years, and homeownership has always been stimulated by tax incentives [48]. It is time to re-engineer the tax incentive system to encourage renovation strategies and accelerate its rate [9].
Furthermore, the study confirms the positive correlation between some household income categories and housing energy efficiency. Very low-income groups (10,000–20,000 EUR of net annual taxable income) occupy much less efficient housing than the rest of the population. Each percent increase in this category corresponds to a rise of 7.22 kWh/m²·y in the built stock’s average energy efficiency. This correlation between income and energy efficiency is not linear at all. Indeed, the differences observed between middle and high-income groups are not very significant. There is a real threshold effect. Therefore, a targeted renovation policy needs to be established, with a particular focus on low-income households (from 10,000–20,000 EUR of net taxable income). These households must be the subject of complementary studies to characterize them deeper and to identify the effective renovation measures to be implemented.

Author Contributions: Conceptualization, G.R. and S.A.; methodology, G.R and M.C.; software, G.R.; validation, G.R., M.C. and S.A.; formal analysis, G.R. and S.A.; investigation, G.R.; resources, G.R.; data curation, G.R.; writing—original draft preparation, G.R.; writing—review and editing, G.R., M.C. and S.A.; visualization, G.R.; supervision, S.A.; project administration, S.A.; funding acquisition, S.A. All authors have read and agreed to the published version of the manuscript.

Funding: Liege University in Belgium funded this research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. Income data can be found here: https://walstat.iweps.be. Geometry, ownership and households data can be found here: https://bestat.statbel.fgov.beRestrictions apply to the availability of the energy data. Energy data was obtained from DG TLPE and are available from the authors with the permission of DG TLPE.

Acknowledgments: The author would like to acknowledge the DG TLPE that gave access to the EPC data. The authors express their thanks to Gilles Tihon, Griet Verbeeck, and Sigrid Reiter and appreciate their valuable comments and feedback.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG TLPE</td>
<td>Direction Général Territoire, Logement, Patrimoine et Environnement</td>
</tr>
<tr>
<td>EPC</td>
<td>Energy Performance Certificate (Performance Energétique des Bâtiments)</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information System</td>
</tr>
<tr>
<td>IWEPS</td>
<td>Institut Wallon de l’Evaluation, de la Prospective et de la Statistique</td>
</tr>
<tr>
<td>MLR</td>
<td>Multiple Linear Regression</td>
</tr>
</tbody>
</table>

References

29. Hens, H.; Parisi, W.; Deuring, M. Energy consumption for heating and rebound effects. Energy Build. 2010, 42, 105–110. [CrossRef]
30. Mjörnell, K.; Femenias, P.; Annadotter, K. Renovation Strategies for Multi-Residential Buildings from the Record Years in Sweden—Profit-Driven or Socioeconomically Responsible? Sustainability 2019, 11, 6988. [CrossRef]
33. Kavousian, A.; Rajagopal, R.; Fischer, M. Ranking appliance energy efficiency in households: Utilizing smart meter data and energy efficiency frontiers to estimate and identify the determinants of appliance energy efficiency in residential buildings. Energy Build. 2015, 99, 220–230. [CrossRef]


47. Cassilde, S. *Analyse de la Base de Données des Certificats PEB en Wallonie* [Report]; Centre d’Etudes en Habitat Durable: Charleroi, Belgium, 2017.