Affordable, sustainable and inclusive housing for marginalised communities

D4.1 Effects of Energy Poverty including the analysis of the multinational survey

January 2026

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement no. 101132513. The responsibility for the information and the views set out in this document lies entirely with the authors. The European Commission is not responsible for any use that may be made of the information

Project acronym	**HouseInc**
Programme	Horizon
Topic	HORIZON-CL2-2023-TRANSFORMATIONS-01
Grant Agreement number	101132513
Start day	February 1, 2024
Duration	36 months
Contact	Anne Kantel (anne.kantel@isi.fraunhofer.de)

Document information

Document fact sheet

Full title	D4.1 Effects of Energy Poverty including the analysis of the multinational survey
Work package	WP4
Task(s)	T4.1
Author (organization)	Peter Conradie (IMEC), Sofie Naeyaert (IMEC), Iska Brunzema (FhG ISI), Hannah Janssen (FhG ISI), Sabine Preuss (FhG ISI)
Reviewers	Nives Della Valle (IEECP), Martin Lux (CAS)
Date	01/2026

Document Dissemination level

Dissemination level

x	PU – Public
	PP – Restricted to other programme participants (including the EC)
	RE – Restricted to a group specified by the consortium (including the EC)
	SEN – Sensitive, only for members of the consortium (including the EC)

Document history

Version	Date	Main modification	Institution
V0.9	18.12.2025	Draft distributed for quality review	IMEC
V1	28.01.2026	Final draft distributed for submission	IMEC

Legal Notice

Find the Project

Web: http://houseinc.eu

LinkedIn: https://www.linkedin.com/showcase/houseinc/

About

The objective of HouseInc is to apply innovative methodology to deeply analyse interlinked dimensions of housing inequalities in the context of marginalized communities.

HouseInc will empirically examine economic, social and ecological drivers and assess impacts of various indicators on housing inequality to derive policy recommendations that foster the adoption of effective measures addressing housing inequality across Europe. With a transdisciplinary dialogue, the project develops innovative social, financial and digital solutions that can be up scaled and thus, contribute to a better socio-economic and sustainable integration of vulnerable groups in European societies.

HouseInc takes a systemic view and assesses interlinkages of housing inequalities - emphasizing energy and mobility poverty, digital dimensions, employment opportunities, family and socio-demographic conditions, energy-efficiency, and health - on a micro-, meso- and macro-level. The interdisciplinary HouseInc consortium - consisting of research institutes and universities, policy think tanks, NGOs, and practitioners on the ground - involves case studies to engage directly with members of four marginalized communities in or from Eastern Europe.

Besides a mix-method approach, including modelling and a GIS-based analysis depicting geographical and future housing inequality, we implement a multinational survey to better understand housing inequality in light of recent events such as COVID-19 and Russia's invasion in the Ukraine. The research results will be assessed, mapped, and scaled up using Living Labs and various stakeholder engagement activities to provide innovative solutions addressing housing inequalities and translating them into valid local, regional, national and EU policy recommendations impacting EU and national funding programs and providing a comprehensive overview and guidance for policymakers to mitigate housing inequalities.

Project partners

Fraunhofer ISI

LIBERI TUTTI

MANNHEIM²

Institute of Sociology Czech Academy of Sciences

POLITEIA

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

DIW BERLIN

EPG ENERGY POLICY GROUP

IEECP INSTITUTE FOR EUROPEAN ENERGY AND CLIMATE POLICY

UCL

Università Commerciale Luigi Bocconi

imec

ROMODROM

University of Exeter

Table of contents

Figures

Tables

Abbreviations

Acronym	Meaning
df	Degrees of freedom
AIC	Akaike Information Criterion
EPC	Energy Performance Certificate
EU-SILC	European Union Statistics on Income and Living Conditions
MAE	Mean absolute error
MICE	Multiple Imputation by Chained Equations
OECD	Organisation for Economic Co-operation and Development
OR	Odds ratio
RMSE	Root mean squared error

EXECUTIVE SUMMARY

This deliverable presents the results of Task 4.1 and provides a detailed examination of housing inequalities and energy poverty across selected EU Member States and the United Kingdom, set against the wider context of the COVID-19 pandemic, the war in Ukraine, and rising cost-of-living pressures. As part of WP4, D4.1 contributes to the broader objective of identifying current and emerging forms of non-financial housing inequality. The deliverable also serves as a source document for subsequent work within the project, including later analytical deliverables and country-level snapshots.

D4.1 draws on data from a multinational survey conducted in eight core countries—Belgium, Czech Republic, Germany, Estonia, Finland, Italy, Romania, and the United Kingdom—using a disproportionate stratified random sampling design. The report documents in detail the sampling, stratification, weighting, and imputation procedures applied to the data, alongside descriptive country-level profiles of survey participants.

The analytical focus of the deliverable is on the relationship between housing inequality and outcomes related to energy poverty, health, wellbeing, mobility, and access to space. Rather than constructing a composite housing inequality index, the analysis adopts a clustering approach to preserve the multidimensional nature of housing disadvantage. A k-prototypes clustering algorithm was applied, allowing the combination of numerical and categorical indicators without imposing predefined weights or linear aggregation. Four variables were used to define housing profiles: the number of dwelling problems, perceived household energy efficiency (subjective EPC), living space per OECD-equivalent household member, and subjective space shortage.

This approach resulted in four distinct housing clusters: Moderate, Affluent, Constrained, and Precarious. The Moderate cluster reflects relatively stable and average housing conditions. The Affluent cluster consists largely of financially secure households living in spacious, high-quality homes, often as outright owners. The Constrained cluster is characterised by smaller dwellings, space shortages, and moderate income limitations, frequently affecting families with children and urban residents. The Precarious cluster faces the most severe strain, combining poor housing quality, very low perceived energy efficiency, low incomes, and high exposure to energy poverty.

Multinomial regression analyses show that income, tenure status, and country of residence are the strongest and most consistent predictors of cluster membership. Higher income substantially increases the likelihood of belonging to the Affluent cluster and reduces the likelihood of being Constrained or Precarious. Renting, whether private or social, is strongly associated with higher odds of belonging to the Constrained and Precarious clusters, while outright ownership is closely linked to the Affluent group. Pronounced cross-country differences highlight how national housing systems shape exposure to housing disadvantage.

The results further demonstrate strong overlaps between housing inequality and energy poverty. Both financial and thermal energy poverty are much more prevalent among households in the Constrained and Precarious clusters. Poorer self-reported health and lower current and expected wellbeing are also strongly associated with these clusters, indicating that housing disadvantage is closely linked to broader social and health inequalities. Differences in transport access and use are more limited, but clear inequalities

emerge in access to private and nearby green space, particularly for the Constrained group, which faces consistent shortages both indoors and outdoors.

While the analysis is based on cross-sectional data and does not allow causal conclusions, it provides a structured and comparative description of how housing conditions, energy poverty, wellbeing, and related outcomes cluster together across households. The findings show that housing inequalities are not isolated problems but intersect with tenure, income, health, energy insecurity, and access to space.

From a policy perspective, the cluster-based framework offers a practical tool for EU-level policy design and coordination. By identifying distinct household profiles facing different combinations of constraints, it supports more differentiated targeting of housing, energy, transport, and social policies. The analysis also highlights that housing-related inequalities can shift under changing circumstances, such as during the COVID-19 period, reinforcing the need for recurring monitoring rather than one-off assessments. Applied over time, this approach can help track changes in housing disadvantage, assess whether policy instruments continue to reach those most affected, and identify persistent coordination gaps across policy domains.

1 INTRODUCTION

This deliverable presents the results of Task 4.1, which examines research and data gaps on housing inequalities and energy poverty across the European Union. The task also considers the wider context shaped by the COVID-19 pandemic, the war in Ukraine, and the resulting cost-of-living and housing pressures in many countries.

D4.1. is one of one of four deliverables within WP4, whose broader aim is to examine specific current and future drivers and forms of non-financial housing inequality and to address existing research gaps through empirical data work and modelling. It will also consider expected developments in housing inequalities. D4.1 and D4.3 builds partly on the multinational survey conducted in T4.1, however, we also note some specific differences between these two deliverables.

First, D4.1 extensively describes the sampling, imputation, and stratification procedures used to construct the dataset applied in both deliverables. It also provides a detailed overview of the sample through country-level descriptions of survey participants. D4.3, by contrast, includes more extensive analyses on themes such as general health and housing inequality. D4.1 limits itself to short summaries and statistical outputs, while D4.3 discusses theory and related literature in more depth.

Second, both deliverables include a k-prototypes clustering analysis, but they do so in different ways. D4.3 uses a wider set of clustering variables, while D4.1 relies on four variables: total dwelling problems, occupant density, subjective space shortage, and perceived household energy efficiency (i.e.: subjective EPC).

Third, D4.1 focuses on the representative subset of the data, described in more detail in chapter 0. This results in more attention to absolute measures, such as levels of fuel poverty, and a closer examination of how sociodemographic factors influence cluster membership. Overall, D4.1 also serves as a source document for later work, including the country snapshots.

A disproportionate stratified random sampling approach was applied within the project's core countries: Belgium, Czech Republic, Germany, Estonia, Finland, the United Kingdom, Italy, and Romania. The rationale for selecting these countries is described in D3.1

As explained in section 2, the sampling strategy aimed to include a substantial share of lower-income participants to enable reliable statistical analysis within this subgroup. Stratification ensured that the sample remained nationally representative for age, gender, and income. The research adhered to the ethical guidelines set by the American Psychological Association. The study protocol received approval from the Ethics Committee of the Faculty of Political and Social Sciences at Ghent University.

2 METHOD AND IMPLEMENTATION

2.1 ANALYTICAL FRAMEWORK AND METHODS

The study focuses on several key thematic areas:

(a) the impact of housing inequality on mobility,

(b) the relationship between housing quality and energy poverty, and

(c) the effects of energy poverty on work mobility, health, and leisure outcomes.

In analysing the collected data, the report applies a multinomial regression approach combined with k-prototypes clustering mentioned above to identify and describe patterns among different population groups. The clustering method allows for the integration of both numerical and categorical variables, supporting a comprehensive examination of complex socio-economic and housing-related dimensions, while the clustering overall allows for the identification of patterns within the sample.

This analytical approach makes it possible to identify distinct clusters of vulnerability and resilience within the surveyed populations and to examine their relationships with key outcomes related to energy poverty and housing inequality. To ensure comparability across analyses, the same set of control variables was consistently applied, followed by the variable of interest (for example, use of public transport) in each model.

The combination of k-prototypes clustering and multinomial regression was chosen to align with the study's primary objective of identifying population typologies characterised by different combinations of housing, energy, and socio-economic conditions. Unlike approaches such as structural equation modelling, which are designed to test hypothesised causal pathways between latent constructs and require stronger assumptions about model structure, directionality, and measurement, the present approach is explicitly descriptive and exploratory. Clustering allows heterogeneity across households to be retained and patterns within the data to be identified, while multinomial regression is subsequently used to examine how these empirically derived housing profiles are associated with sociodemographic characteristics and outcomes of interest, such as energy poverty, health, and mobility.

In this framework, cluster membership functions as an explanatory structure capturing broader housing conditions, and outcomes are analysed as associated consequences rather than as part of a strict causal model. While the analysis is grounded in the assumption that housing conditions shape lived experiences and constraints, the results should be interpreted as descriptive associations. Some overlap between clustering variables and outcomes, particularly in relation to energy-related dimensions, reflects underlying housing conditions and supports the use of the regressions primarily for pattern identification and comparison across clusters.

Finally, our analysis was done using R (R Core Team, 2025), with the package mice being used for the imputation of missing values (Buuren & Groothuis-Oudshoorn, 2011).

2.2 SAMPLING STRATEGY

As noted above, a key objective of the study was to ensure sufficient representation of lower-income households, oversampling the group in the bottom 25%. Because recruitment firms typically do not have access to precise income information at the recruitment stage, it was necessary to rely on proxies for lower income status. These proxies varied by country and included characteristics such as occupation status or age. In addition, income in the EUROSTAT framework is equivalised by household size (OECD equivalised income), which means there is no direct one-to-one correspondence between reported income at recruitment and the final income measure used in the analysis. The latter is calculated post hoc, as described in more detail in section 2.9.3.

For these reasons, recruitment quotas were continuously adjusted to prioritise participants who were likely to fall into lower-income groups based on the available proxies and after calculation of OECD equivalised income. Participation was limited to individuals aged 18 and older residing in one of the eight target regions: Belgium (Flanders), Italy, Germany, Romania, Finland, Estonia, the United Kingdom, and the Czech Republic.

While the study also aimed to include respondents from across NUTS2 regions within each country and to achieve a broadly balanced gender distribution, these criteria were secondary to income. As a consequence, our initial sample was unbalanced (see Table 31: Housing characteristics of the whole sample). To support comparability across countries and analyses, and despite the substantive focus on the bottom income quartile, the sample used in this report was subsequently stratified to approximate national representativeness. This process is discussed in more detail in 0.

2.3 SURVEY DESIGN

Data collection was conducted through an online survey between 24 March and 25 May 2025, using an existing research panel from the marketing research firm iVox. Awarding the tender was preceded by a public procurement process during which multiple firms were reviewed. Participants were rewarded for their participation as part of the regular iVox renumeration scheme, which involves the collection of credits that can be exchanged for money. The exact amount however depends on region.

The initial dataset comprised 22,511 respondents. To increase representation from Romania and Czechia, an additional 937 respondents were recruited between 19 August and 1 September 2025. The survey instrument was programmed and administered using Qualtrics.

The questionnaire began with socio-demographic questions, including age, gender, education, area of residence, income, and housing tenure. This was followed by modules on dwelling characteristics, covering aspects such as available amenities and construction period, as well as questions on health and the consequences of COVID-19 for housing mobility among both tenants and homeowners. Additional sections addressed mobility- and place-related factors, including access to public transport and green spaces in respondents' neighbourhoods.

Subsequent modules focused on energy-related aspects of housing. These included questions on home energy efficiency, such as renovation history and self-assessed energy performance, and on energy poverty, covering heating sources, the ability to afford adequate heating and cooling, and the financial burden of electricity costs. The survey also examined the impact of the recent energy crisis, with a focus on changes in spending behaviour, energy bill awareness, and the adoption of energy-saving measures. Respondents' position within the energy efficiency renovation process and their level of environmental consciousness, including concern about climate change and biospheric values, were also assessed.

The questionnaire further included sections on the acceptance of different energy efficiency renovation policies; however, these items are not analysed in this deliverable. Additional modules captured political opinions, including trust in government, satisfaction with government performance, and political orientation. The survey concluded with questions on housing wealth, covering pathways to first home acquisition, self-reported dwelling value, and ownership of additional residential properties. A complete overview of all survey questions can be found on OSF: https://osf.io/sj659/files/eywd5.

Before fieldwork began, the survey underwent multiple rounds of pretesting to assess completion time, clarity of questions, and the correct functioning of data capture. These pretests consisted of a small-scale pilot administration of the questionnaire combined with targeted feedback on question wording, survey length, and technical performance. Participants in the pretests included members of the target population as well as researchers with experience in survey design. Feedback from the pretests led to minor revisions in question phrasing, adjustments to response options, and improvements to the survey flow and data capture.

2.4 TRANSLATION PROCESS

The survey was originally developed in English and subsequently translated into Dutch, Italian, German, Romanian, Finnish, Estonian and Czech. Forward translations were carried out by native speakers familiar with the survey topics. A back-translation procedure was then applied, in which native speakers translated the items back into English (Eremenco et al., 2005). Discrepancies between the original and back-translated versions were reviewed by the research team and resolved in collaboration with the translators to ensure conceptual equivalence, clarity and consistency across languages. Where necessary, minor cultural adaptations were made to ensure that terms, examples and concepts were appropriate for each national context. For instance, the term "housing allowance" was translated in German as "Wohngeld," which is the official term used by the German state for housing-related financial support. In addition, some items required contextual clarification. The question "Which political party did you vote for in the last national general elections?" was adapted for Germany to reflect the two-vote system used in federal elections. An explanatory note was added to ask respondents to indicate their party vote ("Zweitstimme") if they had cast both a candidate vote and a party vote.

2.5 DATA CLEANING

The initial dataset consisted of 25,485 respondents. To protect participant privacy and remove redundant information, several columns were excluded at source, including LocationLatitude, LocationLongitude, RecipientLastName, RecipientFirstName, RecipientEmail, and IPAddress.

Age-related information was calculated by converting reported year of birth into age (2025 minus year of birth). Minors (participants under 18 years) were removed from the dataset, reducing the number of respondents to 25,477. Consent was verified by filtering for participants who agreed to both parts of the informed consent form (ICF1 and ICF2), reducing the sample to 23,445 respondents. Participants who failed attention checks (CONTROL_1 and CONTROL_2) were removed, leaving 14,951 respondents.

We discuss in detail the differences between those who failed the attention checks and those who didn't below. Survey duration was inspected, with a median completion time of approximately 21.8 minutes.

Because only a small number of respondents identified outside the "Male" or "Female" categories, those who selected Gender Queer, Non-Binary, or did not provide an answer were excluded. This produced a dataset of 14,900 respondents. The study focuses on eight countries: Belgium, Finland, Germany, Romania, Czechia, Italy, the United Kingdom, and Estonia. The selection of these countries was the result of a country selection process, detailed in T3.1 and aimed to ensure the inclusion of countries with varying levels of diverse population sizes, geographies, climates, and housing and welfare regimes.

Some participants indicated living in a different location, but when these entries referred to one of the eight countries (for example, "Suomi" for Finland or "Bayern" for Germany), they were recoded accordingly. The remaining 12 respondents who selected "Other" were removed, resulting in a final sample of 14,888.

Step	Action	Respondents Remaining	Respondents removed	Percentage of original sample
1	Initial dataset	25,485		
2	Remove minors (<18)	25,477	8	0.03%
3	Filter for informed consent (ICF1 & ICF2)	23,445	2,032	7.97%
4	Remove attention check failures (CONTROL_1 & CONTROL_2)	14,951	8,494	33.33%
5	Remove participants not identifying as Male or Female	14,900	51	0.20%
6	Removal of "Other" countries	14,888	12	0.05%

Table 1: Data cleaning process and sample reduction

2.6 FAILED ATTENTION CHECKS

As mentioned, our survey included two attention checks, which were embedded in two matrix questions. The first one stated: "Please click on 'yes'. This is a control question" and was placed near the beginning of the survey. The second one stated: "Please click on 'strongly disagree'. This is a control question" and was included near the end of the survey. Both questions and participants who failed these checks were removed from further analysis. We examined how individuals who failed these checks differed from those who passed across several key characteristics. Participants who failed the attention checks had a median completion time of 8.0 minutes (481 seconds), compared to 21.8 minutes (1307 seconds) for those who passed, but this can be attributed by not being able to continue the survey after failure. On average, those who failed spent significantly less time on the survey (M = 4119.1 seconds) than those who passed (M = 6,204.5 seconds), $t(13,511) = -3.82$, p = $p < .001$. Participants who failed the attention checks were younger on average (M = 45.9 years) than those who passed (M = 49.5 years), $t(11,736) = -15.11$, $p < .001$. Those who failed the attention checks also reported lower income (M = 2411.5) compared to those who passed (M = 2,739.4), $t(10,653) = -12.03$, $p < .001$.

A chi-squared test revealed a significant association between country and attention check outcomes, $\chi^2(7) = 715.77$, $p < .001$. Post hoc analysis of standardized residuals showed that failure rates were lower than expected in Belgium, Germany, and Finland, and higher than expected in Romania, Czechia, and Italy. For example, Germany had fewer failures than expected (standardized residual = -15.46), while Romania had more (residual = 10.14).

Finally, employment status was also associated with attention check failure, $\chi^2(9) = 161.18$, $p < .001$. Standardized residuals indicated that individuals in full-time employment and those engaged in household work were more likely to fail (residuals = 6.45 and 6.35, respectively). Conversely, retired participants and those unable to work were less likely to fail (residuals = -8.70 and -4.32, respectively). Smaller deviations were observed among students, part-time employees, and those on temporary leave. This could be attributed to the financial incentives, with those unable to work, or with more free time, being more willing to fully complete the survey.

2.7 INCOME IMPUTATION

As we will discuss in section 2.9, we apply a stratification procedure to approximate a representative sample within each country along three variables: gender, income and age. To avoid losing participants who did not provide any income data in this sample, we rely on imputing the INCOME_CALC variable. This variable represents a transformed version of the original categorical INCOME variable, which contains 17 options for monthly household income ranging from "Below €500" to "More than €7000" in steps of €500, as well as the options "I don't know" and "I prefer not to say." The midpoint is taken for every option, except for "More than €7000" where €7000 was taken.

However, given that can't force participants to provide their income, or their total assets, both these variables contain missing data. In the case of income, this amounts to 1446 (10%), while for financial assets, it is substantially more at 8508 (57%).

Several approaches can be used for imputing missing data, each with their own drawbacks. A common alternative is mean or median imputation, (Musil et al., 2002) where missing values are replaced with the average or median of the observed data. This method is easy to apply and fast, but it tends to reduce natural variation in the dataset. It can also distort relationships between variables, because all imputed values cluster around a central point. As a result, it often leads to underestimated standard errors and overly narrow confidence intervals.

Another option is regression imputation, in which missing values are predicted using a fitted model based on other variables (Musil et al., 2002). While this preserves some of the associations in the data, it produces values that fall exactly on the regression line and therefore lack random variation. This can make the data look more precise than it really is.

A simpler method is hot-deck imputation (Andridge & Little, 2010), where missing cases are filled by copying observed values from similar respondents (Donders et al., 2006). This can work well when the dataset is large and includes good matching variables. However, it can be unstable if the number of suitable donor cases is limited, and the quality of the imputation depends on how similarity is defined. In addition, hot-deck methods can be hard to document and reproduce, especially when many matching steps are involved.

Compared with these approaches, we prefer Multiple Imputation by Chained Equations (MICE) (Azur et al., 2011). Using multiple rather than single imputations allow us to reflect the statistical uncertainty in the missing values. The chained equations method is also flexible, as it can work with different types of variables (such as continuous or binary). The trade-off is that it requires more computation and more care in specifying the models.

Given the complexity and variety of our dataset, we opted to compared two approaches to data imputation, both applying the MICE procedure, detailed below in section 2.7.1. The aim was to determine which method produced more accurate estimates for intentionally removed income values. In the global approach, all data were pooled while including country as a predictor, whereas in the country-specific approach, imputations were performed separately for each country. We used MICE package for R (Buuren & Groothuis-Oudshoorn, 2011)

2.7.1 METHODOLOGY OVERVIEW

Both approaches use Multiple Imputation by Chained Equations (MICE) to generate 30 completed datasets (m = 30), with a maximum of 15 iterations (maxit = 15). The imputed values are then averaged to produce a single estimate for each missing observation. To enable an objective performance comparison, 20% of the observed INCOME_CALC values were randomly selected and intentionally set to NA. Because the true values of these observations are known, this procedure allows a direct comparison between the imputed and original values.

The first approach, referred to as global imputation, treats the full dataset as a single analytical unit. In this model, COUNTRY_R is included as a categorical predictor alongside other relevant variables such as GENDER, AGE, EMP, and FINBURDEN_ECOST. COUNTRY_R identifies the respondent's country of residence and encodes cross-country differences in economic, social, and institutional contexts that may influence income levels. By including COUNTRY_R as a predictor, the imputation model accounts for systematic differences in income between countries while estimating a single set of relationships between INCOME_CALC and its predictors across the entire dataset. This approach therefore assumes that, although average income levels may differ by country, the underlying associations between income and its predictors are broadly comparable across countries.

The second approach, referred to as country-specific MICE, performs the imputation separately within each country defined by COUNTRY_R. For this approach, the data are first partitioned into country-specific subsets, and an independent MICE model is fitted for each subset. Within these models, COUNTRY_R is excluded as a predictor because it does not vary within a country-specific subset and therefore provides no additional information for imputation. This approach allows the relationships between INCOME_CALC and its predictors to differ across countries, reflecting the possibility that factors such as age, employment status, or financial burden relate to income in country-specific ways rather than through a single global pattern.

2.7.2 VARIABLES USED FOR IMPUTATION

Both imputation approaches used a comprehensive set of variables to predict missing values, chosen for their potential relationship with income and other household characteristics (Azur et al., 2011). These variables can be broadly categorized into demographics and household characteristics, financial indicators, and housing-related attributes.

Variable selection usually means choosing variables that will be included in the analysis and those that help predict missing values (White et al., 2011). In our dataset, several sections were answered only by specific groups, such as questions asked only to homeowners. This makes it impractical to select variables for each sub-analysis, because it would require running a separate imputation process for every analysis model. Each model would need its own imputation setup, which would lead to multiple versions of the imputed variables, each with slight variations. For this reason, we have made the pragmatic consideration to perform our imputation only once and focus on variables that are plausibly related to the two main variables being imputed, income and assets and include variables used in our analysis models. The section below discusses room imputation in more detail, but the same limitations apply.

Demographic and Household Characteristics: These included GENDER, HHTYPE (household type), EMP (employment status), EDU_GROUPED (education level, with smaller education levels grouped), and AGE. COUNTRY_R was also included as a predictor in the global model to capture country-level effects on income.

Financial Indicators: Variables such as FINBURDEN_ECOST (financial burden of energy costs), FIN_BURD_HOUSECOST (financial burden of housing costs), FIN_COMF (financial comfort), CARS_OWN (car ownership), and FIN_DECISION (financial decision-making ability) provided context on economic well-being.

Housing and Lifestyle Attributes: This category encompassed MULT_PROP_FREE (multiple properties owned for free), ROOMS_TOTAL (total rooms), AFF_HEAT (affordability of heating), WELLBE_1 and WELLBE_2 (well-being indicators), ENERGY_APP_TUMBLE_DRYER (ownership of a tumble dryer), GREEN_SPACE_GARDEN_PRIVATE (access to private green space), BILL_CONS (bill consciousness), RENO_HIST (renovation history), PUBLIC_TRANSP_USE (public transport usage), SPACE_SHORT (space shortage), CONSTRUCT_DATE (construction date of dwelling), KITCHEN_ACCESS (kitchen access), DWELLING (dwelling type), and OWNERSHIP (home ownership status). Additionally, health-related variables like CHRON_ILL (chronic illness) and GEN_HEALTH (general health) were included to capture potential impacts on financial capacity.

2.7.3 IMPUTATION ACCURACY ASSESSMENT

The accuracy of each imputation method was evaluated by comparing the imputed values with the original values that were deliberately set to missing. Two commonly used accuracy measures were employed: the root mean squared error (RMSE) and the mean absolute error (MAE)(Li et al., 2024).

- Root Mean Squared Error (RMSE): Measures the average magnitude of the imputation errors. Lower RMSE indicates higher accuracy. RMSE quantifies the square root of the average squared deviation between the imputed and the true values, placing greater weight on larger errors.
- Mean Absolute Error (MAE): Represents the average absolute difference between imputed and true values. Lower MAE indicates higher accuracy and is less sensitive to outliers than RMSE. MAE measures the average absolute difference between the imputed and the true values and provides a more direct interpretation of typical error magnitude.

In addition, accuracy is evaluated by examining the correlation between the imputed values and the corresponding original values, which reflects how well the imputation preserves the relative ordering of observations. A higher correlation (closer to 1) suggests that the imputation effectively preserves the underlying structure and trends of the original data.

2.7.4 RESULTS OF IMPUTATION

The results indicate that the Country-Specific MICE Imputation significantly outperformed the Global MICE Imputation in recovering the removed INCOME_CALC values. The lower RMSE (1,273.35 vs. 1,398.66) and lower MAE (929.87 vs. 1,066.77) for the country-specific approach demonstrate that its imputed values were, on average, closer to the true original income figures. Furthermore, the higher correlation (0.6734 vs. 0.584) suggests that the country-specific method better preserved the underlying patterns and relationships present in the original income data.

Metric	Global MICE	Country-Specific MICE
RMSE	1,398.66	1,273.35
MAE	1,066.77	929.87
Correlation with Original	0.584	0.673

Table 2: Accuracy metrics for the two imputation methods

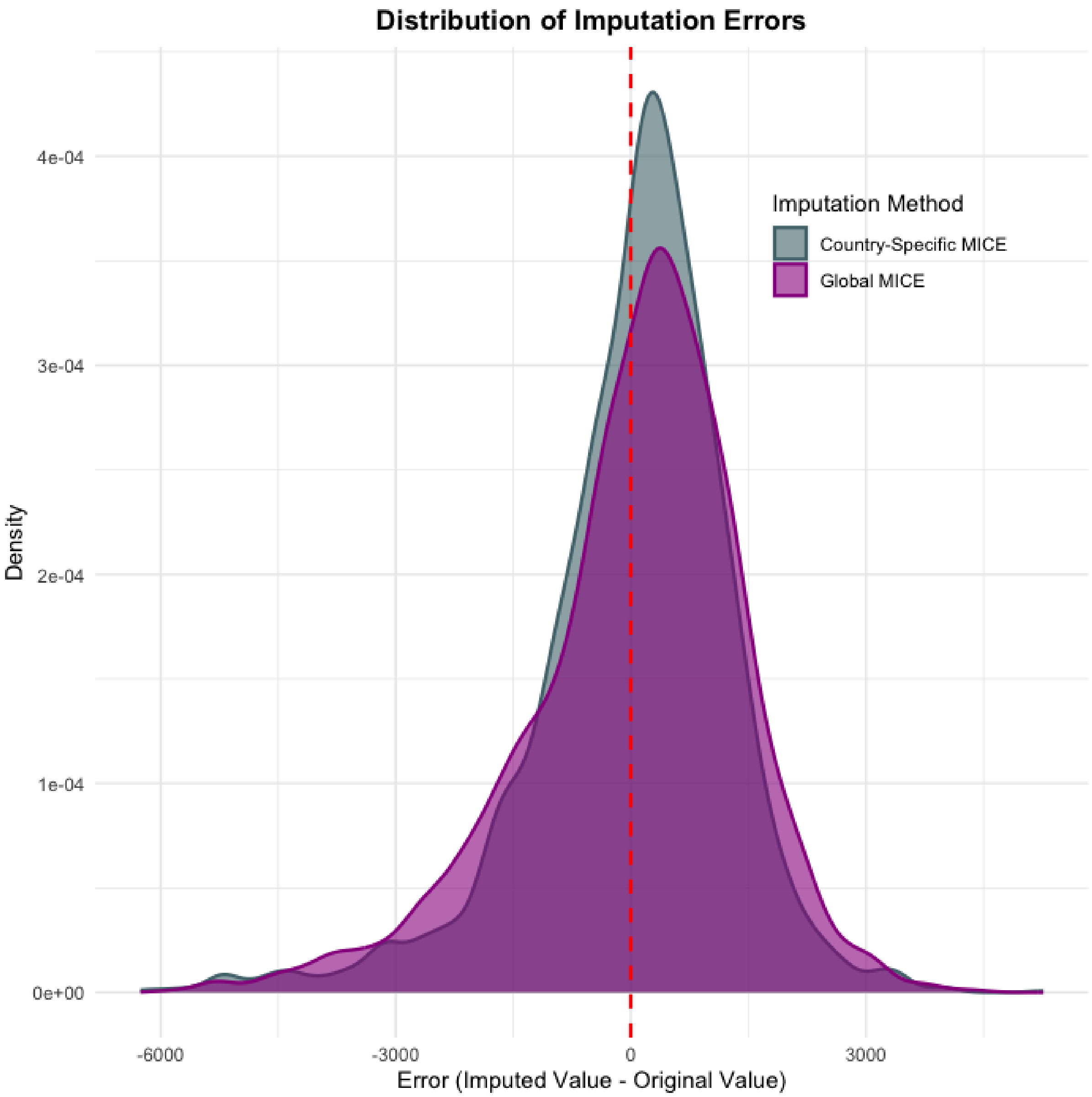

Figure 1: Distribution of imputation errors by method

This finding strongly implies that the factors influencing income (INCOME_CALC) and their respective relationships with income vary substantially across different countries. By performing our imputation independently for each country, the country-specific model was able to capture these unique national contexts, leading to more accurate and reliable imputations.

The density plot in Figure 1 illustrates the accuracy of two distinct multiple imputation approaches for INCOME_CALC. The x-axis represents the imputation error (imputed value minus the original true value), with zero indicating a perfect imputation. The y-axis shows the density of these errors. Optimal accuracy is suggested by density curves that are tall and narrow (indicating low variance in errors) and centred as close to 0 as possible. Country-Specific MICE (grey) demonstrate a tighter distribution around zero compared to the Global MICE method (purple), confirming its superior performance in accurately estimating missing income values.

2.8 SURFACE AREA AND ROOMS IMPUTATION

Both housing floor area and number of rooms contain missing values. In the case of rooms, we are missing 330 (2%), while for surface, we are missing 4450 (29%). Given their importance in the remainder of the project and deliverable (i.e.: to perform the k-prototypes clustering, see section **Fehler! Verweisquelle konnte nicht gefunden werden.**), we applied the same imputation strategy as used for income. For the whole sample, the surface area is moderately correlated with total rooms (r = 0.65), although this varies by country: Finland shows a stronger correlation (0.88), while Belgium is lower (0.37). With only 330 missing values (2%), total rooms thus provide a good basis for imputation. Again, a country-specific approach was used, meaning imputations were performed separately within each country. This allowed the models to capture country-level variation in housing characteristics. As in the income imputations, the procedure generated 30 complete datasets (m = 30) with 15 iterations (maxit = 15). The imputed values were then averaged to produce a single estimate for each missing observation.

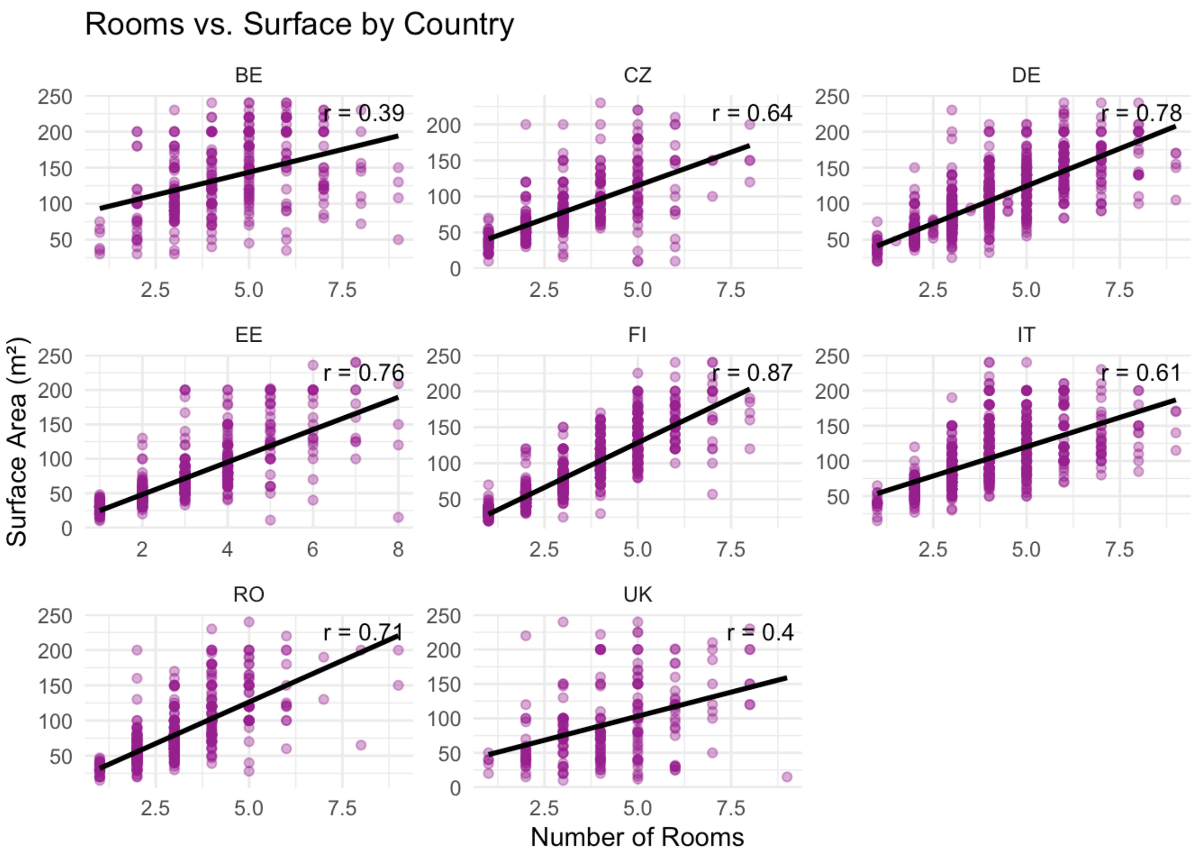

Figure 2: Correlation between dwelling size and amount of rooms

As mentioned earlier, MICE involves selecting variables that are likely to influence our missing variables, as well as variables that will be included in the analysis model (Azur et al., 2011). Because our data will be used in a wide range of analyses, including work on subsets of this dataset, we are, for practical reasons, limited to choosing only predictors that are likely to be related to our missing variables. These included demographic variables (GENDER, AGE, EMP), household and dwelling characteristics (HHTYPE, DWELLING, OWNERSHIP, RENO_HIST, CONSTRUCT_DATE, SPACE_SHORT), and economic indicators (OECD_INCOME, OECD_EQUIV_SIZE_CORRECTED). The outcome variables SURFACE_M2 and ROOMS_TOTAL were also included to inform one another, reflecting their natural correlation.

2.9 SAMPLE REPRESENTATIVENESS

2.9.1 BACKGROUND

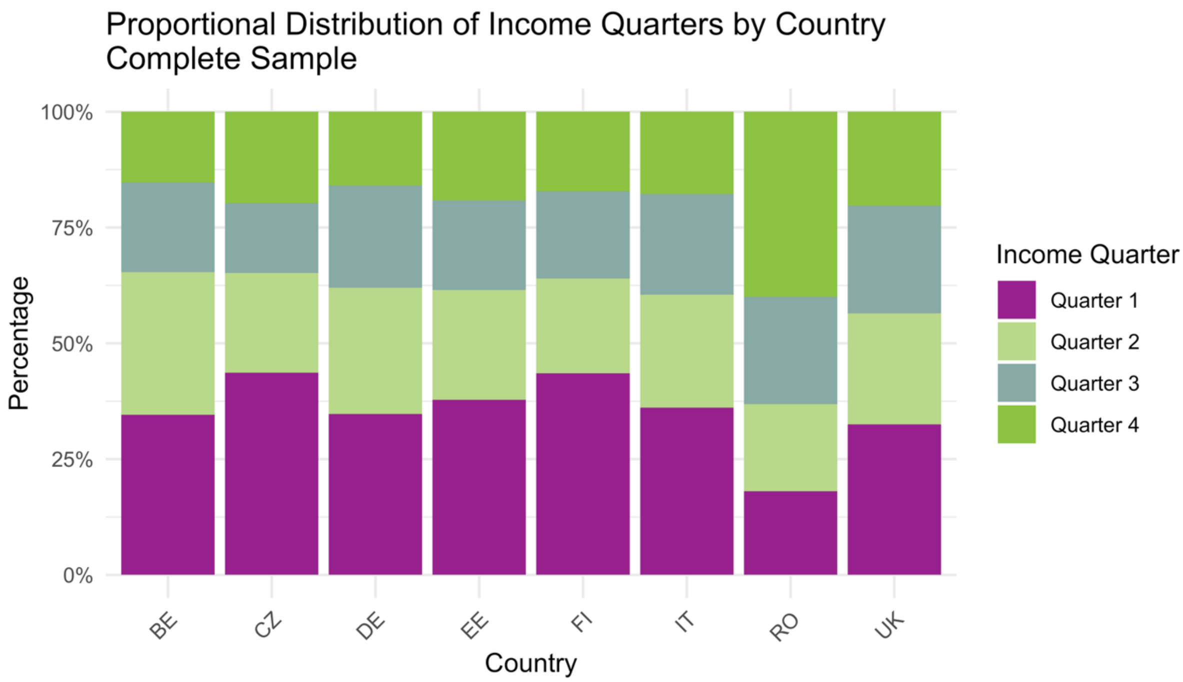

Figure 3: Share of participants in each income quarter

Our sample was initially and purposefully biased toward lower-income groups (see discussion in section 2.2). This limits the generalisability of our results when reporting absolute values such as rates of energy poverty. To address this, we created a sample that approaches representativeness for each selected country. Two broad approaches can be used to correct this type of bias: selecting a new sample through stratification or applying weights to all respondents. In our case, we relied on a controlled stratified sampling procedure that selects cases to match external population margins for income, age, and gender, while respecting feasibility constraints within each country.

Although weighting and stratification aim for the same goal, stratification was suitable here because it shapes the composition of the sample from the start. In a stratified approach, the population is divided into subgroups such as age groups, gender, or income categories, and a fixed number of units is drawn from each subgroup (Groves, 2011, Chapter 4). This ensures that each subgroup appears in the planned proportion and avoids situations in which rare groups are under-represented and require strong corrective weights. Because the structure is fixed before analysis, this approach helps keep the influence of any single case low and avoids extreme weights. These advantages depend on having enough respondents in each subgroup. In our case, with 14,888 participants and over 1,500 respondents per country, using a stratified approach was feasible.

To create the representative dataset, we used EUROSTAT[1] data to determine the target group sizes for income, age, and gender in each country. Details and tables of this data can be found in the appendix from page 76 onwards.

The process had two main parts. First, we identified the best random seed for each country to obtain the largest feasible stratified sample. This was done by testing seeds from 1 to 10,000 and keeping the one that produced the largest valid sample after all stratification steps. Second, we used these seeds to draw the final sample, which is used for all analyses based on representative data.

2.9.2 ADDITIONAL SAMPLING

Most of the data, which includes 14,255 participants, was collected between March 29, 2025, and June 14, 2025. A smaller group of 633 participants was sampled later, between August 27, 2025, and September 1, 2025. This additional sampling concerned 309 participants from Czechia, 1 from Germany, 1 from Italy and 322 from Romania. The exact composition can be viewed in the appendix (Table 33), with these participants selected to increase the representative sample to a minimum of 500 in all eight countries. See section 2.5 for an overview of how our sample was cleaned.

2.9.3 ESTABLISHING INCOME CUTOFFS

Household income was collected by asking participants to indicate the income category they belonged to. Categories started at below €500 and increased in steps of €500, up to €7,000 and above. For participants in CZ, RO, and UK, the amounts were shown in local currency (CZK, RON, GBP) and later converted to euros using the exchange rate of 25 May 2025: 1.195 for GBP, 0.04 for CZK and 0.201 for RON.

To convert categorical values into numerical ones, the midpoint of each category was used. For example, if a respondent reported an income between €1,001 and €1,500, this was converted to €1,250. After establishing the euro amounts, OECD equivalised income was calculated to adjust for household size. Household size was determined using the OECD scale, which assigns a weight of 1.0 to the first adult, 0.5 to each additional adult aged 14 or older, and 0.3 to each child under 14. For instance, a household with two adults and two children has a size of 2.1 (1.0 + 0.5 + 0.3 + 0.3). If this household reports a total annual income of €60,000, the OECD equivalised income is €28,571 (€60,000 divided by 2.1).

The data were also checked for errors in reported household size. In 4% of cases (n=600), the calculated size was below 1, which likely occurred when participants forgot to include themselves when reporting household members. Implausibly large household sizes were also corrected, defined as four times the median, which accounted for 313 cases (2.1%). To address these errors, the reported household size was replaced with the modal value for the corresponding household type. Couples without children were assigned 1.5, couples with children 1.8, those living with parents 2.0, non-family households 1.5, singles 1.0, and single parents 1.5. After these corrections, the mean of INCOME_OECD_EQUIV_CORRECTED was

---

[1] Age: https://ec.europa.eu/eurostat/databrowser/bookmark/b43501a7-4a8b-4aab-946e-bf7f946bbead; Gender: https://ec.europa.eu/eurostat/databrowser/view/demo_pjan/default/table?lang=en; Income: https://ec.europa.eu/eurostat/databrowser/view/ilc_di01$defaultview/default/table

€1,739.81, compared to €1,788.45 for INCOME_OECD_EQUIV before correction, suggesting a minor influence of our correction.

Because national income data are reported yearly rather than monthly, we multiplied the equivalised income provided by our participants by 12. For the UK, Eurostat data were only available up to 2018. To obtain more recent values, we adjusted the 2018 figures for inflation. Eurostat reports in euros, so we applied cumulative eurozone inflation: between 2018 and 2024, prices rose by 21.76% (meaning €100 in 2018 equalled €121.76 in 2024). For reference, the equivalent figure in GBP was 25.38% (£100 in 2018 equalled £125.38 in 2024). This gives a reasonable estimate in the absence of more recent OECD equivalised income data for the UK.

After establishing cutoffs, each participant is then assigned to an income quartile based on their country of residence and the relevant cut-off value. For example, an Italian with a yearly salary of 16,000 would fall into Q2, while a Romanian with the same salary would be in Q4.

Table 3: Income quartile cut-offs

Upper cutoff (EUR)	Belgium	Czechia	Germany	Estonia	Finland	UK	Italy	Romania
Q1	22,159	11,767	19,914	10,747	21,466	14,652	13,808	5,445
Q2	30,392	15,133	27,619	16,140	28,693	21,464	20,605	7,837
Q3	39,403	19,671	38,216	23,577	38,078	31,185	28,821	10,638

2.9.4 PROCESS OF MAKING OUR DATA REPRESENTATIVE

2.9.4.1 Income

The process begins by creating a new sample that is representative of the income distribution. The script compares the current income group distribution in the dataset to EUROSTAT data (i.e. the income cutoffs). It then calculates the largest possible total number of participants that can be drawn from each country while still meeting the target income ratios. Finally, a stratified sampling function is used to select participants, ensuring the sample matches the income-specific targets for each country. Four income groups are chosen (i.e.: quartiles), to ensure sufficiently large group sizes and to retain as many participants as possible. For all countries we aim for a 25% proportion per quarter, the actual final proportions are shown in Table 4.

Table 4: Target income proportions vs actual income proportions

Country	Category	Target proportion	Actual Proportion	Difference
BE	Q1	0.25	0.25	0
	Q2	0.25	0.25	0
	Q3	0.25	0.23	-0.02
	Q4	0.25	0.26	0.01
CZ	Q1	0.25	0.27	0.02
	Q2	0.25	0.26	0.01

	Q3	0.25	0.23	-0.02
	Q4	0.25	0.25	0
DE	Q1	0.25	0.26	0.01
	Q2	0.25	0.25	0
	Q3	0.25	0.24	-0.01
	Q4	0.25	0.25	0
EE	Q1	0.25	0.26	0.01
	Q2	0.25	0.26	0.01
	Q3	0.25	0.23	-0.02
	Q4	0.25	0.25	0
FI	Q1	0.25	0.25	0
	Q2	0.25	0.25	0
	Q3	0.25	0.25	0
	Q4	0.25	0.25	0
UK	Q1	0.25	0.25	0
	Q2	0.25	0.24	-0.01
	Q3	0.25	0.25	0
	Q4	0.25	0.26	0.01
IT	Q1	0.25	0.24	-0.01
	Q2	0.25	0.25	0
	Q3	0.25	0.25	0
	Q4	0.25	0.26	0.01
RO	Q1	0.25	0.23	-0.02
	Q2	0.25	0.25	0
	Q3	0.25	0.25	0
	Q4	0.25	0.27	0.02

2.9.4.2 Age

After establishing income shares, the next step is to ensure that the sample is also representative by age. To do this, custom age bins are created based on target proportions derived from Eurostat. Because the survey does not include respondents below 18, the Eurostat population data are rescaled so that the shares of the remaining age groups increase accordingly. This adjustment means that the proportions in our sample do not match the national statistics exactly, but they accurately reflect the age distribution among adults (i.e. the 18+ proportion).

The Eurostat data are reported in predefined age classes that do not correspond directly to the age bins used in our study. To address this, the population share for each Eurostat age class is first converted into shares per single year of age. For example, in Germany, people aged 0–4 make up 4.5% of the population (3,791,359 out of 83,456,045). Dividing this evenly across the five years in the age class gives approximately 0.9% per year of age. This procedure is repeated across all Eurostat age bins to obtain an approximate share for every single year of age. Once these proportions are calculated, they are rescaled to exclude minors and to reflect only the 18+ population.

The rescaled yearly proportions are then aggregated into larger age bins to preserve as much data as possible while still ensuring comparability across countries. The final bins used are 18–35, 36–49, 50–65, and 66+. For each country, the current age distribution in the sample is compared to these rescaled targets, and the maximum possible sample size per country is determined. We then calculate the exact target counts for each bin and draw a stratified subsample from the income-representative data.

As an illustration in Table 5 for Germany the raw Eurostat share of people aged 18–35 is 21%. After rescaling to exclude minors, this rises to 25%. In our representative sample, the share is 25% (274 respondents). Similarly, for the 50–65 group, the Eurostat share is 22%, which increases to 28% after rescaling, and our sample reflects this with 28% (300 respondents). The complete overview of target proportions for all countries can be viewed in Table 6.

Table 5: German age rescaling example

Age Bin	Eurostat Actual	Rescaled (18+)	Sample Count	Sample
18–35	0.21	0.25	275	0.25
36–49	0.18	0.21	232	0.21
50–65	0.22	0.28	307	0.28
66+	0.21	0.26	283	0.26

Table 6: Target age proportions vs actual age proportions

Country	Category	Target proportion	Actual Proportion	Difference
BE	18–35	0.26	0.25	-0.01
	36–49	0.22	0.22	0
	50–65	0.26	0.26	0
	66+	0.26	0.26	0
CZ	18–35	0.24	0.24	0
	36–49	0.27	0.26	-0.01
	50–65	0.25	0.25	0
	66+	0.24	0.24	0
DE	18–35	0.25	0.25	0
	36–49	0.21	0.21	0
	50–65	0.28	0.28	0
	66+	0.26	0.26	0
EE	18–35	0.26	0.25	-0.01
	36–49	0.25	0.26	0.01
	50–65	0.25	0.25	0
	66+	0.24	0.24	0
FI	18–35	0.27	0.27	0

	36–49	0.22	0.22	0
	50–65	0.24	0.24	0
	66+	0.28	0.28	0
UK	18–35	0.29	0.29	0
	36–49	0.22	0.22	0
	50–65	0.26	0.26	0
	66+	0.23	0.23	0
IT	18–35	0.22	0.22	0
	36–49	0.21	0.21	0
	50–65	0.29	0.29	0
	66+	0.28	0.28	0
RO	18–35	0.26	0.25	-0.01
	36–49	0.25	0.25	0
	50–65	0.26	0.26	0
	66+	0.24	0.24	0

The same procedure is applied across all participating countries. While the exact target proportions differ slightly between countries, depending on their demographic structures, the process ensures that the final stratified sample is both income- and age-representative within the adult population.

2.9.4.3 Gender

Finally, the age- and income-balanced sample is used to achieve gender representativeness. Like the previous steps, the current gender distribution is compared to the exact target proportions for each country (see Table 7 for target and actual proportions per country), drawn from Eurostat[2]. A final "bottleneck" calculation determines the maximum sample size for each country that can fulfil the gender proportion requirements. Target counts for each gender group within each country are then computed and used to draw the final sample.

Table 7: Target gender proportions vs actual gender proportions

Country	Category	Target proportion	Actual Proportion	Difference
BE	Female	0.51	0.51	0
	Male	0.49	0.49	0
CZ	Female	0.51	0.51	0

2 Gender data from Eurostat: https://ec.europa.eu/eurostat/databrowser/view/demo_pjan/default/table?lang=en

	Male	0.49	0.49	0
DE	Female	0.51	0.51	0
	Male	0.49	0.49	0
EE	Female	0.53	0.53	0
	Male	0.47	0.47	0
FI	Female	0.51	0.51	0
	Male	0.49	0.49	0
UK	Female	0.51	0.51	0
	Male	0.49	0.49	0
IT	Female	0.51	0.51	0
	Male	0.49	0.49	0
RO	Female	0.51	0.51	0
	Male	0.49	0.49	0

2.9.5 CUMULATIVE EFFECT ON SAMPLE

The quota-based procedure described above is intended to improve the representativeness of the sample along key sociodemographic dimensions. It is important again to note that the sample was not designed to be fully representative at the data collection stage. Instead, the sampling strategy deliberately oversampled individuals in the lower income groups in order to ensure sufficient observations in the bottom income quintile. The quota adjustments therefore serve to restore representativeness ex post, rather than to preserve it throughout the initial sampling process.

Quotas for income, age, and gender were applied sequentially using simple (marginal) quotas within each country. While this approach substantially improves alignment with population distributions, small deviations from earlier quota targets may occur when subsequent quotas are imposed. For example, achieving the required age or gender proportions may necessitate removing some respondents selected during the income-stratification step, resulting in departures from the original income targets.

Cross-quotas, which enforce joint distributions across income, age, and gender simultaneously, were not applied. Although cross-quotas can provide a closer match to the full multivariate population structure, they would have substantially reduced the achievable sample size, particularly in smaller countries and in narrowly defined subgroups. Given the already constrained sample sizes and the need to retain sufficient observations across income groups, especially for CZ and RO, the use of cross-quotas would have led to an excessive loss of data. The chosen approach therefore represents a pragmatic balance between improving representativeness and preserving adequate sample sizes for analysis.

Overall, the final stratified sample achieves good alignment with national income, age, and gender distributions among the adult population, while maintaining sufficient coverage of lower-income groups, which was a core design objective of the survey.

3 SAMPLE COMPOSITION

3.1 COMPLETE DATASET

3.1.1 SOCIO-DEMOGRAPHIC DETAILS

Given our focus in this deliverable on our representative dataset, we restrict our description here to the key details. The full description and accompanying table can be found in the appendix on page 80.

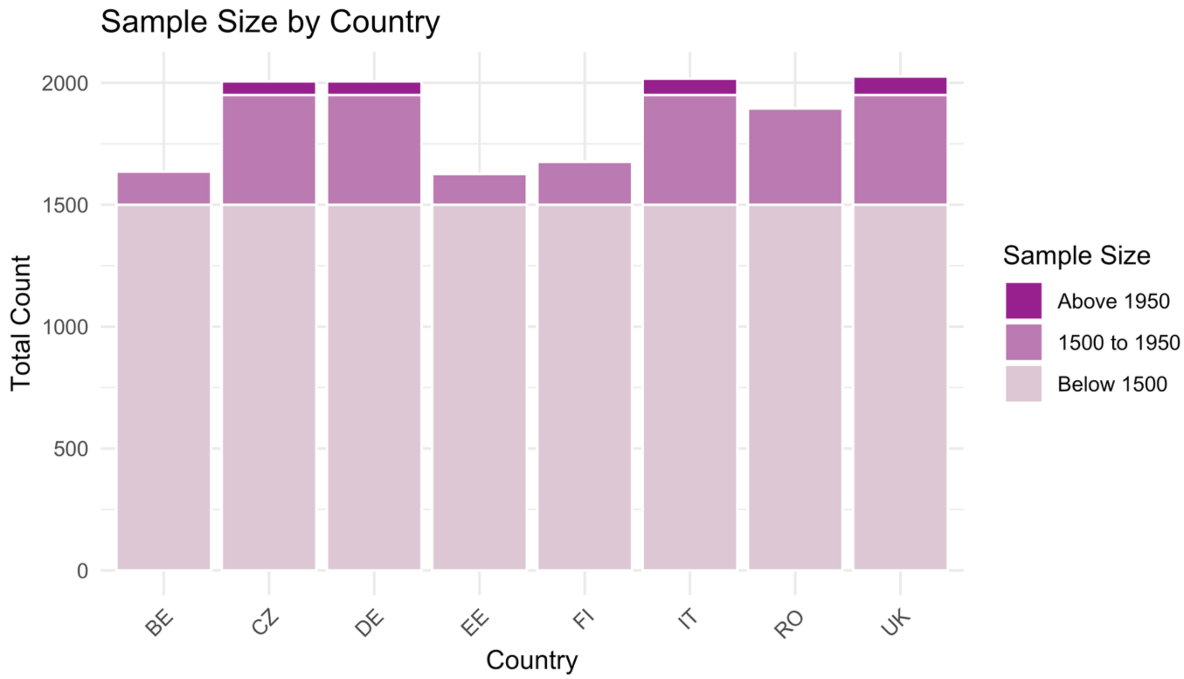

Figure 4: Overview of total amount of participants (complete dataset)

A shown in Figure 4, all countries achieved a final sample size above 1,500 respondents, with the Czech Republic, Germany, the United Kingdom, and Italy exceeding 1,950 respondents, which corresponds to the minimum target sample size for Germany, Italy, and the United Kingdom. The pooled dataset comprises 14,888 respondents across eight countries.
The overall median age of the sample is 50 years, although there is notable variation across countries. Respondents in Belgium, Germany, Finland, the United Kingdom, Romania, and Italy tend to be older, with median ages ranging from 50 to 55 years, while respondents in the Czech Republic and Estonia are younger, with median ages of 44 and 46 years, respectively. This pattern is reflected in the age group distribution. Estonia and the Czech Republic have relatively high shares of respondents aged 18–35 (31% in both countries), whereas Belgium has a lower share in this age group (17%) and a higher proportion of respondents aged 65 and older (25%).
When compared with national median ages from Eurostat (Germany: 45.5; Czech Republic: 44.0; Belgium: 42.0; Finland: 43.4; Italy: 48.7; Romania: 43.8), the sample median ages are generally higher in most countries. This indicates a tendency towards an older respondent profile and subsequently sample bias.
Gender distribution is even across countries, though women make up a slight majority in all samples, with the highest share in Estonia (59%). Marked differences emerge in household income. Median equivalised income is highest in Belgium (€2,250) and the United Kingdom (€2,062), followed by Germany (€2024) and Finland (€1875). In contrast, the lowest medians

are observed in Romania (€833), the Czech Republic and Estonia (€1,167 each). Income quartiles show that Romania is distinct, with 40% of its sample in the top quartile, while Finland and Czech Republic have relatively high proportions in the bottom quartile (43-44%). A more in-depth overview of the characteristics of the complete sample can be found in our appendix on page 80.

3.1.2 REGIONAL SPREAD

The regional breakdown highlights uneven representation across territories. In the Czech Republic, Prague accounts for 16% of the sample, with other respondents spread across Moravia and Bohemia. In Germany, the sample is more concentrated in the larger states, with Nordrhein-Westfalen (20%) and Bavaria (17%) accounting for the biggest shares. In Estonia, half the sample comes from Harju county, where the capital Tallinn is located. Finland's sample is concentrated in Uusimaa (32%, including Helsinki) and Pirkanmaa (10%). In the UK, England dominates with 88%, while Scotland, Wales, and Northern Ireland together represent just over 10%.

The Italian sample is geographically broad, though Lombardy (19%) and Lazio (9%) are the largest contributors. In Romania, Bucharest alone accounts for 16% of the sample, while respondents in other counties are more evenly distributed, with no single county exceeding 5%. For Belgium all participants lived in Flanders. Our appendix contains a table with the complete overview on page 80.

3.2 REPRESENTATIVE SAMPLE CHARACTERISTICS

3.2.1 SOCIODEMOGRAPHIC DETAILS

The stratified representative sample includes 7,169 respondents across the eight countries. As discussed earlier, stratification was performed on gender, age groups, and income to ensure comparability across countries. The sample is broadly balanced, with some country-specific differences worth noting. We have a minimum of 500 participants for each country.

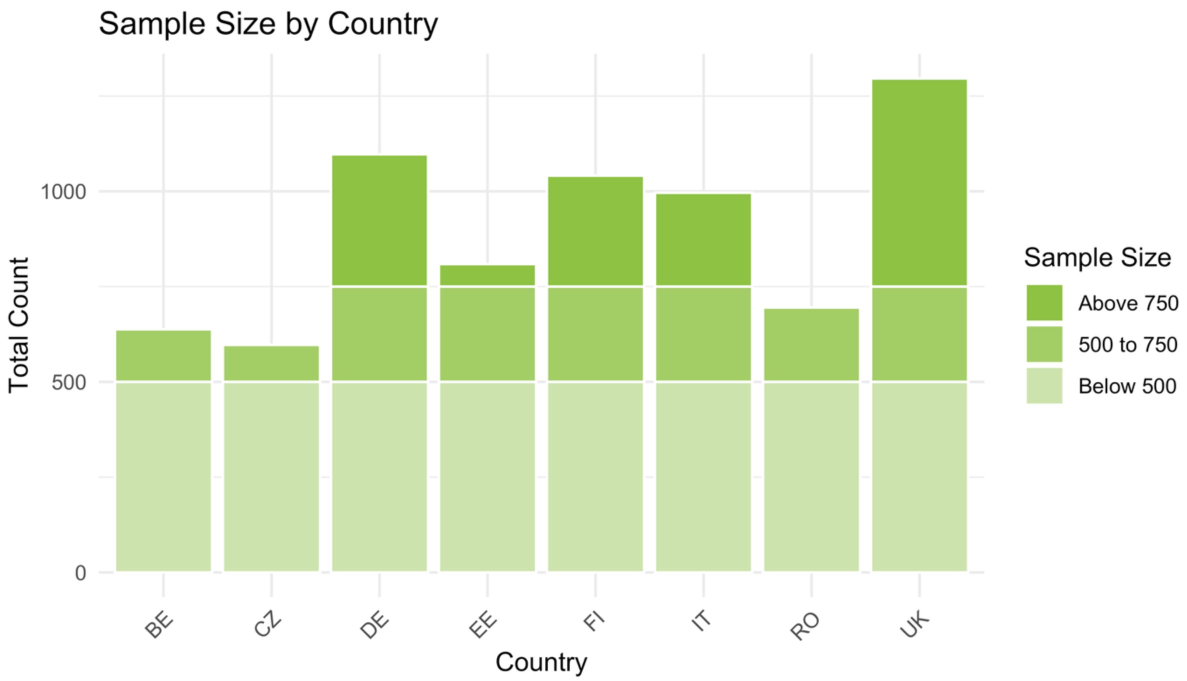

Figure 5: Overview of number of participants for the representative sample in each country

The median age is 51 years, with interquartile ranges suggesting a widespread across younger and older respondents. Germany, Belgium, and Italy show slightly older profiles (median around 52-54 years), while the United Kingdom has the youngest (median 48 years). The age group breakdown is evenly distributed by design, though subtle differences appear: the United Kingdom has more 18-35-year-olds (29%), while Germany and Italy lean more toward the 50-65 bracket (28 – 29%). As is to be expected after the stratification process, gender is nearly balanced overall, with 51% women and 49% men. This holds consistently across all countries, with only minor variation (see Figure 6).

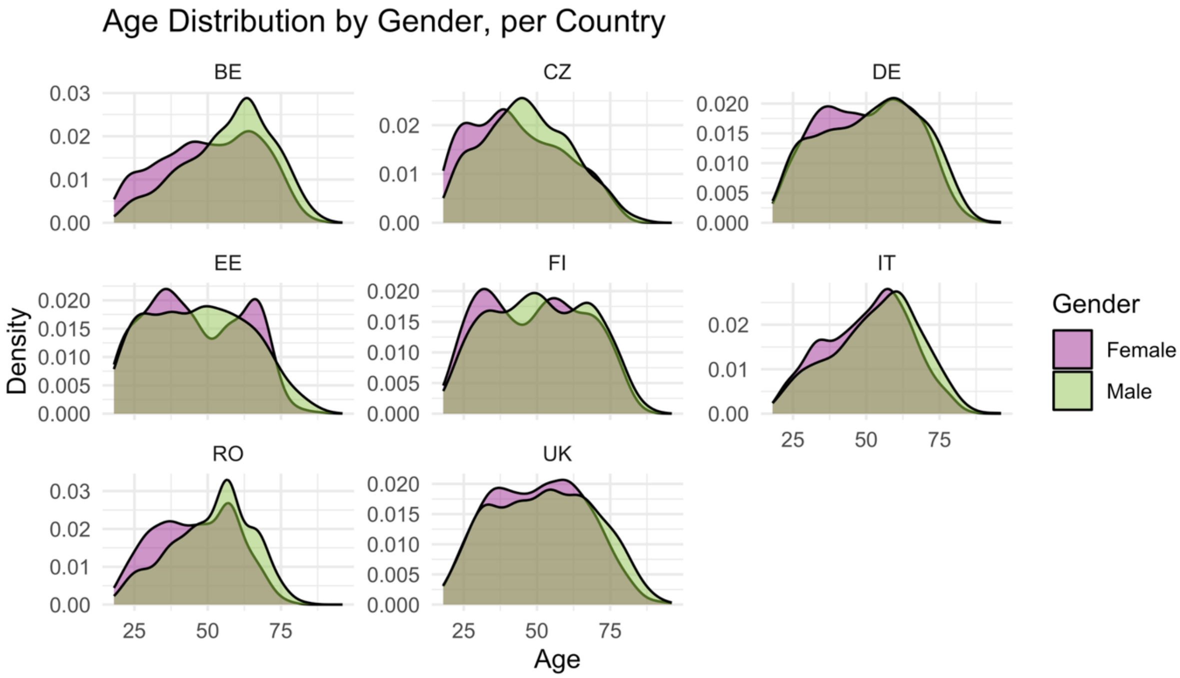

Figure 6: Gender and age distribution across the eight countries for the representative sample

Income patterns show clearer contrasts. Median equivalised income is highest in Belgium, Finland, Germany and the United Kingdom (ranging from €2,250–€2,500), while Romania reports the lowest (median €694, mean €869). The quartile distribution is nearly uniform by design, but Romania's strong concentration in the lowest quartile (23%) paired with a larger share in the highest (27%) signals higher inequality within that subsample.

Household composition also varies. Couples with children are most common in Romania (44%) and Italy (38%), while couples without children dominate in Germany (39%) and Finland (40%). Singles are relatively frequent in Finland (34%), Germany and Estonia (28% each) but rare in Romania and Italy (14% each). Living with parents is strikingly higher in Italy (12%), Romania and Belgium (each around 10%) compared with other countries.

Educational attainment varies across countries. In Finland, 44% of respondents have higher non-university qualifications and 14% hold a master's degree or higher. In Italy, 39% have completed secondary education, with fewer respondents holding university qualifications. Romania reports 36% of respondents with a bachelor's degree, compared to 8% in Italy and

8% in the Czech Republic. Belgium shows a relatively even distribution across educational categories.

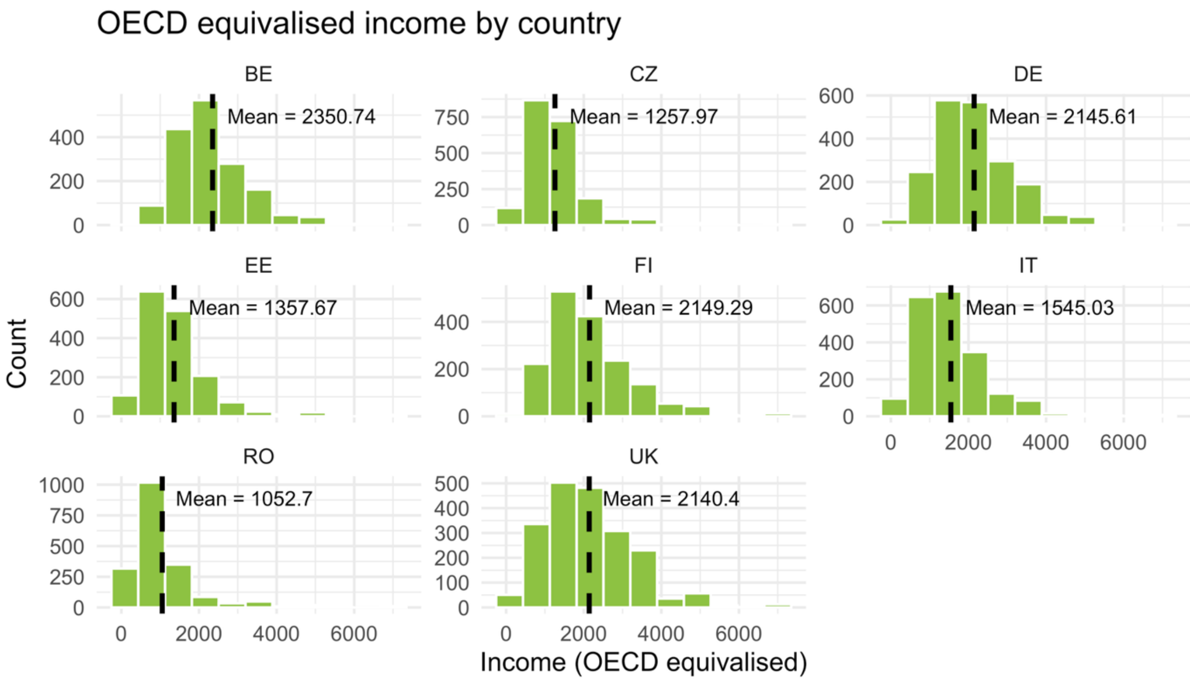

Figure 7: OECD transformed income per country for the representative sample.

Urban–rural differences are pronounced among the participating countries. Romania stands out with 78% living in cities, while the United Kingdom (56%) and Belgium (50%) have the largest share in towns/suburbs. The strong urban population in Romania, however, may point to a sampling bias due to the online recruitment channel.

Furthermore, Employment patterns vary across countries. Full-time employment rates are highest in Estonia (52%) and the Czech Republic (51%), but lower in Italy (38%). Germany (13%) and the United Kingdom (11%) report the highest shares of part-time work.

Respondents whose main activity is household work (homemakers) are most common in Italy (9%), much higher than in other countries. Retired respondents are most represented in Belgium (34%) and Finland (29%). Finland has most students (6%), while Italy and Romania show higher unemployment (approximately 7%).

Finally, migration status varies considerably across countries. The United Kingdom reports the largest share of residents born abroad (11%), followed by Italy and the Czech Republic (around 8%). Finland and Belgium have a very small share of residents born outside the country (2% and 3%), and Romania is almost entirely composed of people born in Romania (99%).

Table 8: Sociodemographic characteristics of the representative sample

Characteristic	Overall	BE	CZ	DE	EE	FI	UK	IT	RO
	N = 7,169[1]	N = 638[1]	N = 597[1]	N = 1,097[1]	N = 809[1]	N = 1,041[1]	N = 1,296[1]	N = 996[1]	N = 695[1]
Age	51 (35, 66)	53 (36, 66)	49 (36, 65)	52 (36, 66)	49 (36, 65)	50 (35, 67)	48 (34, 64)	54 (38, 66)	49 (35, 63)

Age Group									
18–35	1,832 (26%)	159 (25%)	146 (24%)	274 (25%)	202 (25%)	277 (27%)	377 (29%)	220 (22%)	177 (25%)
36–49	1,648 (23%)	141 (22%)	158 (26%)	233 (21%)	207 (26%)	229 (22%)	291 (22%)	214 (21%)	175 (25%)
50–65	1,873 (26%)	169 (26%)	148 (25%)	306 (28%)	206 (25%)	245 (24%)	333 (26%)	287 (29%)	179 (26%)
65+	1,816 (25%)	169 (26%)	145 (24%)	284 (26%)	194 (24%)	290 (28%)	295 (23%)	275 (28%)	164 (24%)
Gender									
Female	3,662 (51%)	323 (51%)	304 (51%)	556 (51%)	426 (53%)	526 (51%)	661 (51%)	509 (51%)	357 (51%)
Male	3,507 (49%)	315 (49%)	293 (49%)	541 (49%)	383 (47%)	515 (49%)	635 (49%)	487 (49%)	338 (49%)
Income (Equivalised, €)	1,806 (1,167, 2,750)	2,500 (1,833, 3,250)	1,250 (900, 1,591)	2,250 (1,563, 3,043)	1,250 (833, 1,833)	2,250 (1,750, 3,167)	2,250 (1,500, 3,125)	1,625 (1,125, 2,250)	694 (500, 972)
Income Quartile									
Q1	1,802 (25%)	160 (25%)	160 (27%)	283 (26%)	209 (26%)	257 (25%)	328 (25%)	242 (24%)	163 (23%)
Q2	1,806 (25%)	162 (25%)	155 (26%)	279 (25%)	211 (26%)	261 (25%)	314 (24%)	251 (25%)	173 (25%)
Q3	1,740 (24%)	149 (23%)	135 (23%)	264 (24%)	187 (23%)	261 (25%)	323 (25%)	249 (25%)	172 (25%)
Q4	1,821 (25%)	167 (26%)	147 (25%)	271 (25%)	202 (25%)	262 (25%)	331 (26%)	254 (26%)	187 (27%)
Household Type									
Couple No Children	2,260 (32%)	204 (32%)	196 (33%)	433 (39%)	192 (24%)	415 (40%)	344 (27%)	295 (30%)	181 (26%)
Couple With Children	2,291 (32%)	186 (29%)	202 (34%)	270 (25%)	265 (33%)	210 (20%)	478 (37%)	377 (38%)	303 (44%)
Living With Parents	475 (6.6%)	63 (9.9%)	47 (7.9%)	43 (3.9%)	42 (5.2%)	23 (2.2%)	72 (5.6%)	117 (12%)	68 (9.8%)
Non Family Household	104 (1.5%)	4 (0.6%)	11 (1.8%)	15 (1.4%)	18 (2.2%)	9 (0.9%)	25 (1.9%)	14 (1.4%)	8 (1.2%)
Single	1,683 (23%)	151 (24%)	102 (17%)	308 (28%)	228 (28%)	351 (34%)	299 (23%)	144 (14%)	100 (14%)
Single Parent	356 (5.0%)	30 (4.7%)	39 (6.5%)	28 (2.6%)	64 (7.9%)	33 (3.2%)	78 (6.0%)	49 (4.9%)	35 (5.0%)
Level of Education									
1 - Limited Education	865 (12%)	48 (7.7%)	30 (5.0%)	194 (18%)	87 (11%)	161 (15%)	100 (7.8%)	225 (23%)	20 (2.9%)
2 - Secondary	1,828 (26%)	228 (36%)	271 (46%)	171 (16%)	161 (20%)	107 (10%)	251 (19%)	393 (39%)	246 (35%)
3 - Higher Non-University	1,630 (23%)	46 (7.3%)	132 (22%)	332 (30%)	179 (22%)	460 (44%)	278 (22%)	114 (11%)	89 (13%)

4 - Higher University (Bachelors)	1,491 (21%)	183 (29%)	52 (8.7%)	156 (14%)	198 (24%)	171 (16%)	399 (31%)	82 (8.2%)	250 (36%)
5 - Higher University (Masters+)	1,331 (19%)	121 (19%)	110 (18%)	244 (22%)	184 (23%)	142 (14%)	260 (20%)	182 (18%)	88 (13%)
Missing	24	12	2	0	0	0	8	0	2
Living Area									
City	3,342 (47%)	153 (24%)	336 (56%)	532 (48%)	521 (64%)	526 (51%)	339 (26%)	396 (40%)	539 (78%)
Rural	1,322 (18%)	163 (26%)	125 (21%)	239 (22%)	125 (15%)	123 (12%)	231 (18%)	194 (19%)	122 (18%)
Town / Suburb	2,505 (35%)	322 (50%)	136 (23%)	326 (30%)	163 (20%)	392 (38%)	726 (56%)	406 (41%)	34 (4.9%)
Employment Status									
Employed Full Time	3,238 (45%)	273 (43%)	307 (51%)	523 (48%)	417 (52%)	426 (41%)	601 (46%)	375 (38%)	316 (45%)
Employed Part Time	624 (8.7%)	49 (7.7%)	37 (6.2%)	139 (13%)	76 (9.4%)	72 (6.9%)	139 (11%)	89 (8.9%)	23 (3.3%)
Household Work	244 (3.4%)	12 (1.9%)	14 (2.3%)	11 (1.0%)	14 (1.7%)	14 (1.3%)	40 (3.1%)	91 (9.1%)	48 (6.9%)
Other	37 (0.5%)	0 (0%)	2 (0.3%)	2 (0.2%)	8 (1.0%)	8 (0.8%)	6 (0.5%)	3 (0.3%)	8 (1.2%)
Retired	1,816 (25%)	217 (34%)	146 (24%)	290 (26%)	148 (18%)	306 (29%)	278 (21%)	251 (25%)	180 (26%)
Self Employed	355 (5.0%)	14 (2.2%)	38 (6.4%)	33 (3.0%)	50 (6.2%)	37 (3.6%)	74 (5.7%)	72 (7.2%)	37 (5.3%)
Student	303 (4.2%)	34 (5.3%)	24 (4.0%)	48 (4.4%)	31 (3.8%)	58 (5.6%)	31 (2.4%)	44 (4.4%)	33 (4.7%)
Temporary Leave	83 (1.2%)	4 (0.6%)	11 (1.8%)	16 (1.5%)	14 (1.7%)	11 (1.1%)	13 (1.0%)	4 (0.4%)	10 (1.4%)
Unable to Work	177 (2.5%)	26 (4.1%)	5 (0.8%)	14 (1.3%)	17 (2.1%)	35 (3.4%)	75 (5.8%)	2 (0.2%)	3 (0.4%)
Unemployed	292 (4.1%)	9 (1.4%)	13 (2.2%)	21 (1.9%)	34 (4.2%)	74 (7.1%)	39 (3.0%)	65 (6.5%)	37 (5.3%)
Born in Current Country									
No	415 (5.8%)	19 (3.0%)	46 (7.7%)	71 (6.5%)	40 (4.9%)	19 (1.8%)	138 (11%)	75 (7.5%)	7 (1.0%)
No Answer	3 (<0.1%)	0 (0%)	0 (0%)	1 (<0.1%)	0 (0%)	1 (<0.1%)	0 (0%)	1 (0.1%)	0 (0%)
Yes	6,751 (94%)	619 (97%)	551 (92%)	1,025 (93%)	769 (95%)	1,021 (98%)	1,158 (89%)	920 (92%)	688 (99%)

[1] Median (Q1, Q3); n (%)

3.2.2 HOUSING CHARACTERISTICS OF THE REPRESENTATIVE SAMPLE

Detached houses are most common overall (28%), with the highest shares in Romania (34%), Belgium (33%), and Germany (30%). Semi-detached and terraced houses are more prevalent in the United Kingdom (48%) and Belgium (45%), while large flats (10+ dwellings) dominate in Finland (40%) and the Czech Republic (38%). Smaller flats with one or two dwellings are more common in Estonia (17%) and Romania (13%), reflecting different housing stock compositions across countries.

Housing tenure structure also shows significant variation. Most respondents own their homes without a mortgage (44%), with particularly high rates in Romania (75%), Estonia and Italy (each 54%). Mortgaged ownership is frequent in Belgium (36%) and Finland (33%), while private renting is most common in Germany (41%) and the Czech Republic (28%). Social renting is generally low but slightly higher in Finland (14%) and the United Kingdom (12%).

Construction periods reflect historical housing patterns. Homes built between 1971 and 2006 dominate the sample (42%), especially in Estonia (48%) and Finland (52%). Older housing (pre-1940) is more common in the Czech Republic (18%) and the United Kingdom (19%), while newer construction (post-2007) accounts for 18% overall, with Belgium (23%), Finland (21%) and het United Kingdom (22%) showing higher shares. A small proportion of respondents are unsure of the construction year, ranging from 2.4% in Finland and 10% in the UK.

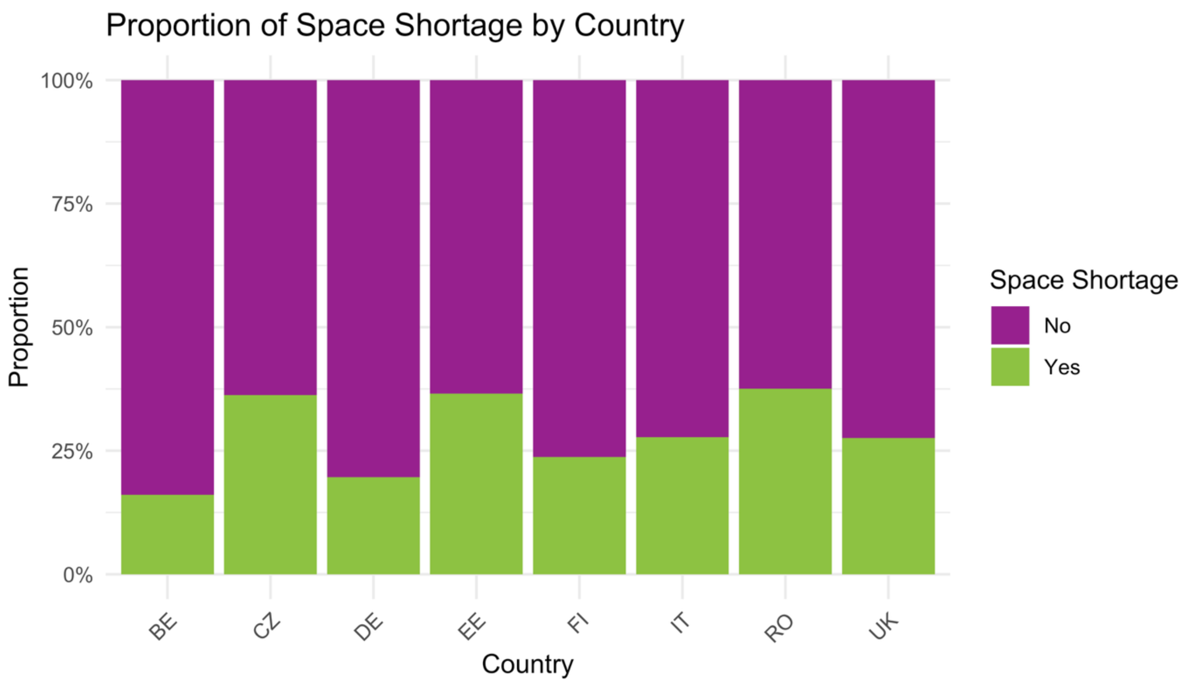

Figure 8: Space shortage per country, representative sample

Perceived space shortage affects 26% of respondents, most prominently in Romania (40%) and the Czech Republic (34%), indicating that housing density and size are key concerns in these countries (see Figure 8). Kitchen availability is nearly universal (99 - 100%). Housing size varies across countries. The mean surface area is largest in Belgium (143 m²) and Italy (103 m²) and smallest in Romania (81 m²) and Estonia (75 m²). Most dwellings in the sample have three (26%) or four rooms (20%). Very small units (one room) are rare (5%), as are very large dwellings with eight rooms or more (under 4%). Country differences are evident: two-room dwellings are especially common in Romania (36%) and Estonia (28%), while three-room dwellings dominate in the Czech Republic (32%) and Italy (31%). Larger dwellings with five or more rooms are most frequent in Belgium and the United Kingdom.

The reported market value of dwellings also differs substantially, with the highest average in Belgium (€364,877) and the lowest in Romania (€78,536), reflecting differences in property markets. Major energy renovations are unevenly distributed. While 35% of the sample report no renovations, recent updates (from 2020 onward) are more common in Belgium (22%), the United Kingdom and Estonia (17%). Renovations from earlier periods show variation, with Germany reporting the largest share of energy renovations before 2005 (9%).

Most households report being able to afford adequate heating, with 69% indicating no difficulty. However, challenges are evident in Romania (34%), Italy and the United Kingdom (28%), where many households can afford to heat only part of their home.

In addition, there are clear gender differences in heating affordability: women are more likely than men to report having no or only some ability to afford adequate heating, while men are more likely to report no difficulty. The gender effect should be interpreted with caution. Gender refers to the respondent, not all household members, and therefore does not indicate gendered differences in housing conditions but rather who completed the survey.

Cooling comfort in summer shows more variation: only 51% overall manage to keep their dwelling comfortably cool. Comfort levels are highest in Belgium and the United Kingdom (61–64%), while Estonia (53%) and Finland (27%) report the lowest levels.

Table 9: Dwelling characteristics of the representative sample

Characteristic	Overall	BE	CZ	DE	EE	FI	IT	RO	UK
	N = 7,169[1]	N = 638[1]	N = 597[1]	N = 1,097[1]	N = 809[1]	N = 1,041[1]	N = 996[1]	N = 695[1]	N = 1,296[1]
Type of Dwelling									
Detached House	1,990 (28%)	210 (33%)	157 (26%)	324 (30%)	223 (28%)	204 (20%)	241 (24%)	236 (34%)	395 (30%)
Flat: 10+ Dwellings	1,687 (24%)	52 (8.2%)	225 (38%)	179 (16%)	262 (32%)	416 (40%)	252 (25%)	243 (35%)	58 (4.5%)
Flat:1-2 Dwellings	579 (8.1%)	25 (3.9%)	32 (5.4%)	81 (7.4%)	137 (17%)	40 (3.8%)	76 (7.6%)	89 (13%)	99 (7.6%)
Flat:3-9 Dwellings	1,143 (16%)	63 (9.9%)	113 (19%)	342 (31%)	128 (16%)	75 (7.2%)	245 (25%)	73 (11%)	104 (8.0%)
Other	54 (0.8%)	1 (0.2%)	4 (0.7%)	6 (0.5%)	6 (0.7%)	8 (0.8%)	6 (0.6%)	6 (0.9%)	17 (1.3%)
Semi-detached/Terraced House	1,716 (24%)	287 (45%)	66 (11%)	165 (15%)	53 (6.6%)	298 (29%)	176 (18%)	48 (6.9%)	623 (48%)
Tenure status									
Free Residence	203 (2.8%)	22 (3.4%)	13 (2.2%)	12 (1.1%)	38 (4.7%)	9 (0.9%)	49 (4.9%)	37 (5.3%)	23 (1.8%)

Owned, No Mortgage	3,190 (44%)	273 (43%)	274 (46%)	292 (27%)	439 (54%)	275 (26%)	535 (54%)	522 (75%)	580 (45%)
Owned, With Mortgage	1,660 (23%)	230 (36%)	95 (16%)	226 (21%)	172 (21%)	341 (33%)	214 (21%)	58 (8.3%)	324 (25%)
Rented: Private	1,560 (22%)	88 (14%)	170 (28%)	450 (41%)	142 (18%)	271 (26%)	166 (17%)	61 (8.8%)	212 (16%)
Rented: Social	556 (7.8%)	25 (3.9%)	45 (7.5%)	117 (11%)	18 (2.2%)	145 (14%)	32 (3.2%)	17 (2.4%)	157 (12%)
Construction Period									
Before 1940	826 (12%)	72 (11%)	105 (18%)	142 (13%)	89 (11%)	51 (4.9%)	75 (7.5%)	48 (6.9%)	244 (19%)
1941-1970	1,490 (21%)	128 (20%)	141 (24%)	244 (22%)	150 (19%)	199 (19%)	219 (22%)	124 (18%)	285 (22%)
1971-2006	3,036 (42%)	248 (39%)	214 (36%)	467 (43%)	388 (48%)	543 (52%)	478 (48%)	350 (50%)	348 (27%)
2007-Later	1,315 (18%)	144 (23%)	79 (13%)	185 (17%)	134 (17%)	223 (21%)	145 (15%)	119 (17%)	286 (22%)
Don't Know	502 (7.0%)	46 (7.2%)	58 (9.7%)	59 (5.4%)	48 (5.9%)	25 (2.4%)	79 (7.9%)	54 (7.8%)	133 (10%)
Perceived Space Shortage	1,867 (26%)	102 (16%)	203 (34%)	193 (18%)	269 (33%)	213 (20%)	259 (26%)	277 (40%)	351 (27%)
Kitchen Availability	7,115 (99%)	635 (100%)	593 (99%)	1,091 (99%)	798 (99%)	1,039 (100%)	992 (100%)	682 (98%)	1,285 (99%)
Surface Area (m²)	95 (50)	143 (50)	85 (44)	100 (43)	75 (45)	86 (44)	103 (40)	81 (45)	93 (57)
Total Rooms									
1	353 (4.9%)	8 (1.3%)	47 (7.9%)	32 (2.9%)	75 (9.3%)	106 (10%)	19 (1.9%)	43 (6.2%)	23 (1.8%)
2	1,353 (19%)	42 (6.6%)	143 (24%)	205 (19%)	224 (28%)	267 (26%)	105 (11%)	251 (36%)	116 (9.0%)
3	1,865 (26%)	116 (18%)	193 (32%)	308 (28%)	244 (30%)	248 (24%)	309 (31%)	205 (29%)	242 (19%)
4	1,415 (20%)	134 (21%)	89 (15%)	211 (19%)	144 (18%)	188 (18%)	269 (27%)	101 (15%)	279 (22%)
5	1,056 (15%)	137 (21%)	73 (12%)	160 (15%)	53 (6.6%)	146 (14%)	153 (15%)	62 (8.9%)	272 (21%)
6	579 (8.1%)	81 (13%)	31 (5.2%)	108 (9.8%)	36 (4.4%)	52 (5.0%)	77 (7.7%)	19 (2.7%)	175 (14%)
7	303 (4.2%)	59 (9.2%)	13 (2.2%)	42 (3.8%)	23 (2.8%)	21 (2.0%)	40 (4.0%)	9 (1.3%)	96 (7.4%)
8	170 (2.4%)	42 (6.6%)	7 (1.2%)	21 (1.9%)	10 (1.2%)	8 (0.8%)	17 (1.7%)	3 (0.4%)	62 (4.8%)
9	75 (1.0%)	19 (3.0%)	1 (0.2%)	10 (0.9%)	0 (0%)	5 (0.5%)	7 (0.7%)	2 (0.3%)	31 (2.4%)
Dwelling Value (EUR)	219,355	364,877 (126,502)	199,894 (118,590)	301,460 (163,496)	141,363 (116,650)	199,444 (123,562)	183,863 (127,082)	78,536 (82,855)	276,085 (177,067)

	(157,59 8)								
Missing	4.357	355	357	820	469	598	574	418	766
Major Energy Renovations									
1 - No	2,526 (35%)	197 (31%)	172 (29%)	427 (39%)	258 (32%)	379 (36%)	475 (48%)	231 (33%)	387 (30%)
2 - Before 2005	463 (6.5%)	28 (4.4%)	40 (6.7%)	94 (8.6%)	51 (6.3%)	39 (3.7%)	53 (5.3%)	54 (7.8%)	104 (8.0%)
3 - 2005-2009	435 (6.1%)	36 (5.6%)	59 (9.9%)	57 (5.2%)	67 (8.3%)	41 (3.9%)	40 (4.0%)	64 (9.2%)	71 (5.5%)
4 - 2010-2014	531 (7.4%)	57 (8.9%)	59 (9.9%)	75 (6.8%)	62 (7.7%)	63 (6.1%)	60 (6.0%)	66 (9.5%)	89 (6.9%)
5 - 2015-2019	706 (9.8%)	83 (13%)	74 (12%)	77 (7.0%)	97 (12%)	100 (9.6%)	69 (6.9%)	82 (12%)	124 (9.6%)
6 - From-2020	1,103 (15%)	143 (22%)	68 (11%)	147 (13%)	138 (17%)	142 (14%)	143 (14%)	107 (15%)	215 (17%)
7 - Yes: Unknown Year	518 (7.2%)	44 (6.9%)	67 (11%)	45 (4.1%)	68 (8.4%)	85 (8.2%)	50 (5.0%)	60 (8.6%)	99 (7.6%)
8 - Don't Know	887 (12%)	50 (7.8%)	58 (9.7%)	175 (16%)	68 (8.4%)	192 (18%)	106 (11%)	31 (4.5%)	207 (16%)
Affordability of Heating									
No	514 (7.2%)	20 (3.1%)	35 (5.9%)	46 (4.2%)	73 (9.0%)	49 (4.7%)	127 (13%)	52 (7.5%)	112 (8.6%)
Some	1,682 (23%)	139 (22%)	152 (25%)	299 (27%)	128 (16%)	82 (7.9%)	280 (28%)	233 (34%)	369 (28%)
Yes	4,973 (69%)	479 (75%)	410 (69%)	752 (69%)	608 (75%)	910 (87%)	589 (59%)	410 (59%)	815 (63%)
Ability to keep cool									
No	1,643 (23%)	89 (14%)	129 (22%)	234 (21%)	429 (53%)	279 (27%)	170 (17%)	175 (25%)	138 (11%)
Some	1,853 (26%)	158 (25%)	216 (36%)	313 (29%)	95 (12%)	224 (22%)	316 (32%)	202 (29%)	329 (25%)
Yes	3,673 (51%)	391 (61%)	252 (42%)	550 (50%)	285 (35%)	538 (52%)	510 (51%)	318 (46%)	829 (64%)
[1] Mean (SD); n (%)									

4 RESULTS

4.1 CLUSTERING APPROACH

As noted earlier, the aim of this deliverable is to explore the relationship between housing inequality and outcomes in areas such as health, mobility, leisure, and energy. While the survey includes several variables related to housing inequality, such as self-reported housing quality and perceived space shortage, it does not provide a single composite indicator. One possible approach in similar research areas, for example energy poverty, is the construction of a composite or multidimensional index. Such indices require explicit choices regarding variable selection, weighting schemes, and aggregation rules, which can impose normative assumptions about the relative importance of different housing dimensions and reduce transparency in how heterogeneity across households is represented.

To avoid these limitations, a clustering approach was adopted. Specifically, a cluster analysis was conducted using the k-prototypes algorithm, which is designed for datasets containing both numerical and categorical variables. Unlike an index, which collapses multiple dimensions into a single scale, clustering groups households with similar housing conditions across multiple dimensions without imposing predefined weights or linear aggregation. This allows the multidimensional structure of housing inequality to be preserved and facilitates the identification of distinct housing profiles that may relate differently to outcomes in health, mobility, leisure, and energy.

An alternative is structural or latent variable modelling, for example through factor analysis, where underlying dimensions of housing inequality are inferred from correlations among observed variables. While such methods can reduce dimensionality in a data-driven way, they typically assume that housing inequality can be represented along one or more continuous latent constructs and rely on linear relationships between variables. In this deliverable, the analytical objective is instead to identify distinct household profiles characterised by different combinations of housing conditions, including mixed numerical and categorical indicators. Moreover, the presence of categorical variables and the interest in non-linear patterns of co-occurrence limit the applicability and interpretability of factor-analytic approaches. For these reasons, a clustering strategy was considered more appropriate for capturing heterogeneity in housing situations without imposing latent structural assumptions.

To identify relevant clusters, a set of variables was selected to represent key housing characteristics. We opted for a combination of subjective measures that takes into consideration how spacious or energy efficient people feel their homes are (subjective feelings of space shortage and subjective Energy Performance Certificate (EPC) rating, combine with more objective variables such as the presence of problems and how crowded the home is. These were taken in preference to, for example, other dwelling characteristics such as the location of the dwelling (rural, vs urban), dwelling type (flat vs detached house) or construction date, or tenure status, because the primary aim of the clustering was to capture lived housing conditions and constraints that are most directly experienced by households, rather than structural or contextual attributes of the dwelling. Subjective assessments of space shortage and energy performance reflect how residents perceive and cope with their housing situation, while objective indicators such as overcrowding and the presence of physical problems provide complementary information on material conditions. Variables such as dwelling location, dwelling type, construction date, or tenure status were therefore excluded from the clustering, as they describe broader structural characteristics that may shape housing outcomes but do not in themselves indicate housing inequality. These variables are instead used in subsequent analyses to contextualise and interpret the identified clusters, ensuring a clear distinction between variables used to define housing profiles and those examined as correlates or outcomes. Below we discuss the variables in more detail.

The first variable is a sum of the total number of dwelling problems. This measure reflects the number of physical defects reported for each household and is created by summing six binary indicators, each representing a specific type of problem:

- Leaking roof (yes / no)
- Damp walls or floors (yes / no)
- Rot in windows or floors (yes / no)
- Mould around windows (yes / no)
- Mould on walls (yes / no)
- Mould on furnishings (yes / no)

Each variable takes the value 1 if the problem is present and 0 if not. The resulting variable, DP_TOTAL, represents the count of housing problems in a dwelling, ranging from 0 (no problems) to 6 (all six problems reported).

The second variable is a subjective EPC rating. Participants were asked to rate the energy efficiency of their home on a scale that was converted to a range between 1 and 10, where 10 represents the highest perceived energy efficiency and 1 the lowest. We emphasised the link between energy efficiency and costs: *A dwelling with a green EPC label is very energy-efficient with low energy costs, and on the other side a dwelling with a red EPC label is not at all energy-efficient with high energy costs.* Given that EPC rating across the EU is not uniform and many people are typically unaware of their EPC, we rely on perceived household energy efficiency, or subjective EPC.

The third variable is the imputed surface area of the dwelling divided by the OECD-equivalent household size. This measure provides an estimate of how much living space is available per adjusted household member, reflecting the degree of crowdedness within the home.

The final variable is a binary indicator capturing whether respondents feel that their home suffers from a shortage of space. All numerical variables were standardised before running the clustering.

A concern was to ensure that none of the variables were highly correlated. To check this, Pearson correlation coefficients were calculated. As shown in Table 10, no problematic correlations were found; the strongest correlation was between the perceived house energy efficiency and the total number of dwelling problems (–0.28). Space shortage was converted to for this purpose to a 1 or 0. The four selected variables were therefore retained for the k-prototypes clustering. We perform our cluster analysis with the complete sample.

Table 10: Clustering data and data type; Pearson correlation table for the clustering variables

Data type and description			Pearson Correlation			
Variable	Description	Data Type	1	2	3	4
DP_TOTAL (1)	Sum of total dwelling problems	Numerical (Standardised)	1	-0.28	-0.11	0.22
EPC_10 (2)	Perceived house energy efficiency (scale 1–10)	Numerical (Standardised)	-0.28	1	0.09	-0.13
SURFACE_OECD_PP (3)	Imputed surface area divided by OECD-equivalent household size	Numerical (Standardised)	-0.11	0.09	1	-0.27
SPACE_SHORT_num (4)	Subjective experience of space shortage	Categorical (Factor, 0 or 1)	0.22	-0.13	-0.27	1

A second concern is determining the appropriate number of clusters (k). To address this, we ran eight successive models, varying k from 2 to 8. This range was chosen given that two is the minimum and 8 was a reasonable maximum number of clusters.

For each model, we examined the total cost function, which measures how well the clustering fits the data by summing the dissimilarities within all clusters. Lower values indicate that households within each cluster are more similar to each other. An elbow plot of the cost function showed clear diminishing returns beyond four clusters (see Figure 9). The largest decrease in within-cluster dissimilarity occurred when increasing from k = 2 to k = 3 (a reduction of 7,762.51), followed by a smaller but still notable improvement from k = 3 to k = 4 (6,703.22). Beyond this point, the rate of improvement slowed considerably, with reductions falling to 4,893.39 for k = 5 and 2,043.41 for k = 6. This flattening pattern suggests that four clusters offer a suitable balance between capturing variation within the data and keeping the clusters simple and interpretable. As a result, we proceed with the four-cluster solution.

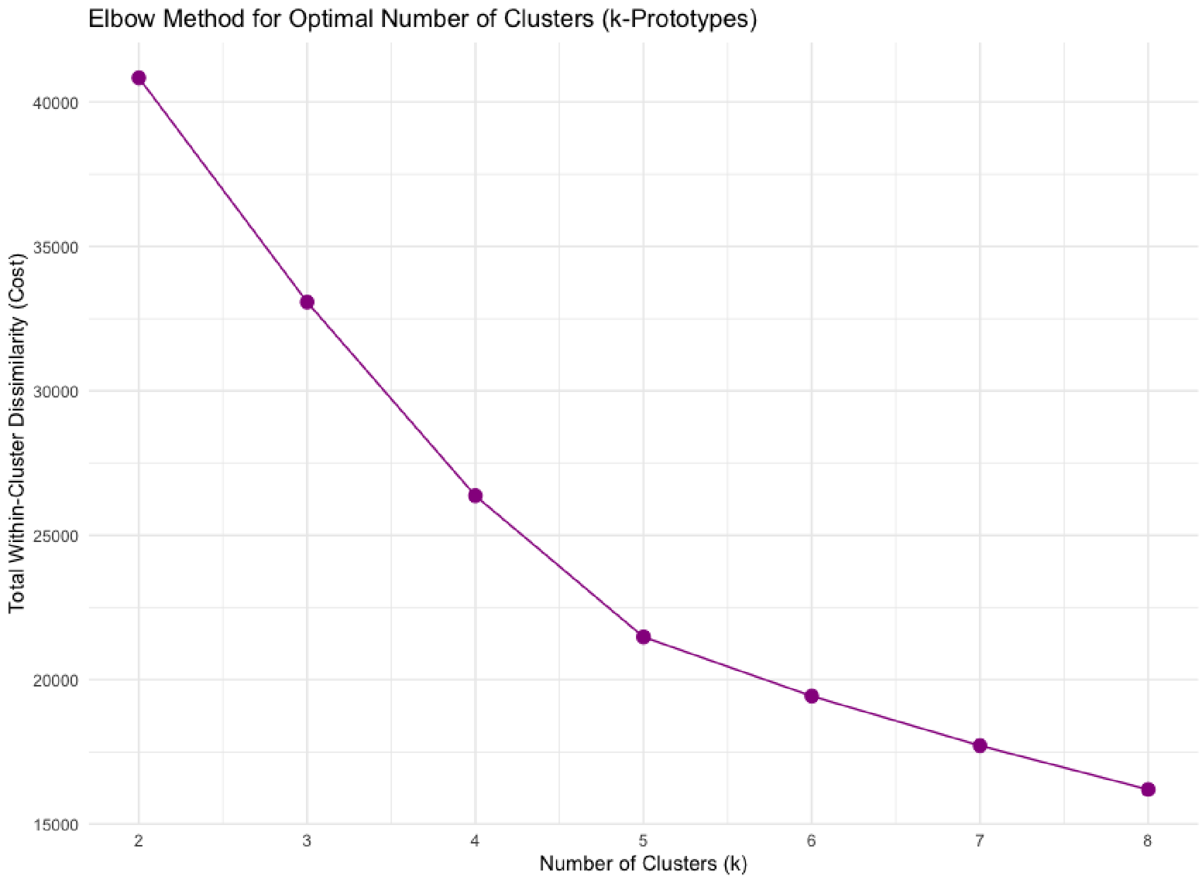

Figure 9: Total within-cluster distance for eight different k-prototypes results

4.2 DESCRIPTION OF THE CLUSTERS

Each cluster captures a particular combination of housing and socio-economic characteristics, allowing for a clearer understanding of how housing inequality manifests across the sample. In the following section, we describe each cluster in detail, highlighting their main housing conditions, demographic profiles, and economic situations, starting first with the variables used in the clustering and then accompanied by a selection of additional sociodemographic and economic variables.

Moderate cluster (N = 3,561): The first and biggest cluster is the Moderate cluster. Households show relatively balanced housing and socio-economic conditions. They report few dwelling problems on average (0.15), and subjective EPC scores are moderate (5.71). Their homes have a mean surface area of 83.6 m², with 3.7 rooms on average, and they experience no reported space shortages. The average density (51 m² per OECD persons) and family size (1.69) suggest typical occupancy levels. Their income is near the sample average (€1,935), and the distribution across income quartiles is fairly even, though slightly fewer are in the lowest quartile. Most respondents are middle-aged (mean age 51), with 44% in full-time employment and 27% retired. Nearly half (45%) own their homes outright, while a fifth have mortgages, and a quarter rent privately. Half of these households live in cities, with others spread across towns and rural areas. The group includes a mix of household types, with couples with children (32%) and multi-person households without children (40%) being the most common.

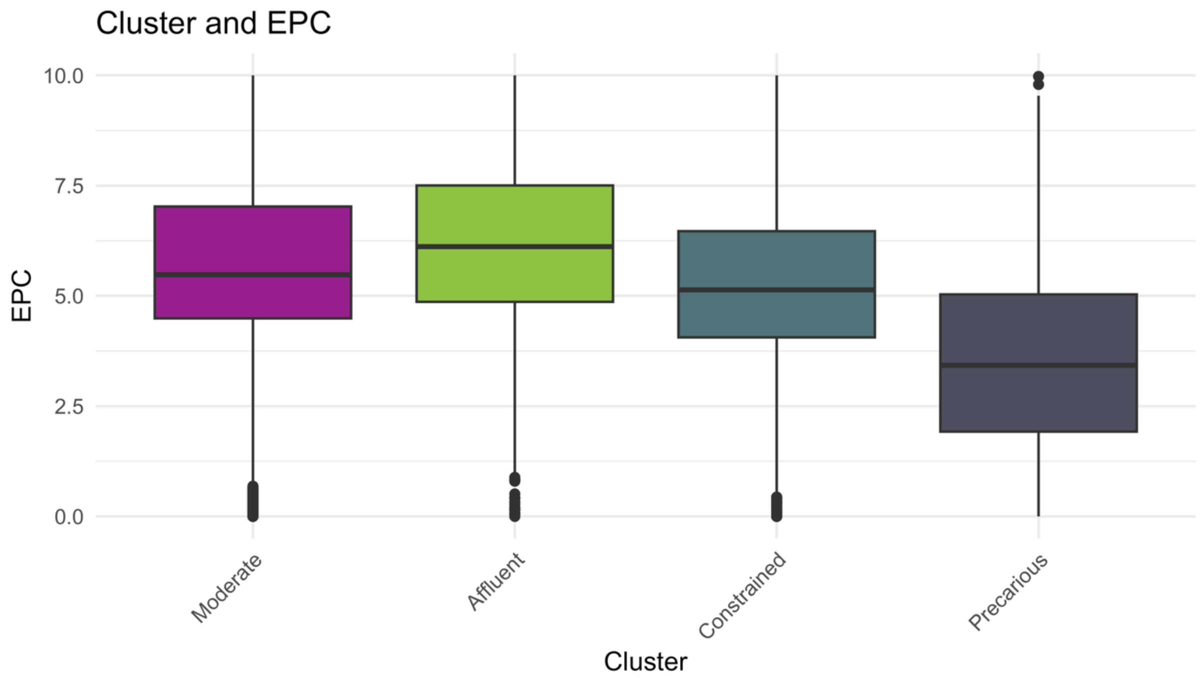

Figure 10: Cluster and subjective EPC

Affluent cluster (N = 1,402): The Affluent cluster is characterised by comfortable living conditions and high material well-being. They have few dwelling problems (0.20), high EPC scores (6.10), and large homes averaging 157.7 m² with about five rooms. Despite their large surface area, they still report some space shortage, although it is quite rare (7%). The high m² per OECD equivalent person (119) reflects spacious conditions relative to family size rather than crowding. These households have the highest mean income (€2,715) and are concentrated in the top income quartiles in their respective countries, with 42% in Q4. Respondents are slightly older (mean age 55), and the group has a relatively high share of retired individuals (37%). Over half own their homes outright (57%), and 29% still have mortgages, while renting is uncommon. They tend to live in towns and suburban areas (42%) and are often couples without children or in "other" household types (47%). This cluster represents stable, well-established households with strong economic resources and comfortable housing.

Constrained cluster (N = 1,502): The Constrained cluster faces notable housing limitations. Every household in this group reports a space shortage (100%), and their dwellings are the smallest on average (67.7 m²) with only three rooms. They have relatively high levels of dwelling problems (0.45) and somewhat lower subjective EPC scores (5.24). Their average income (€1,672) is below the sample mean, with nearly one-third in the lowest income quartile in their respective country. Their family size (1.71) and age profile (mean 46) are comparable to the Moderate group, but a higher share is employed full-time (51%), indicating working households under spatial constraints. Ownership is less common: only 35% own their homes outright, while 27% rent privately and 10% occupy social housing. The majority live in cities (57%), suggesting limited space and higher density urban environments. This cluster mainly includes families with children (37%) and multi-person households (34%), reflecting a combination of working households struggling with small, constrained living spaces.

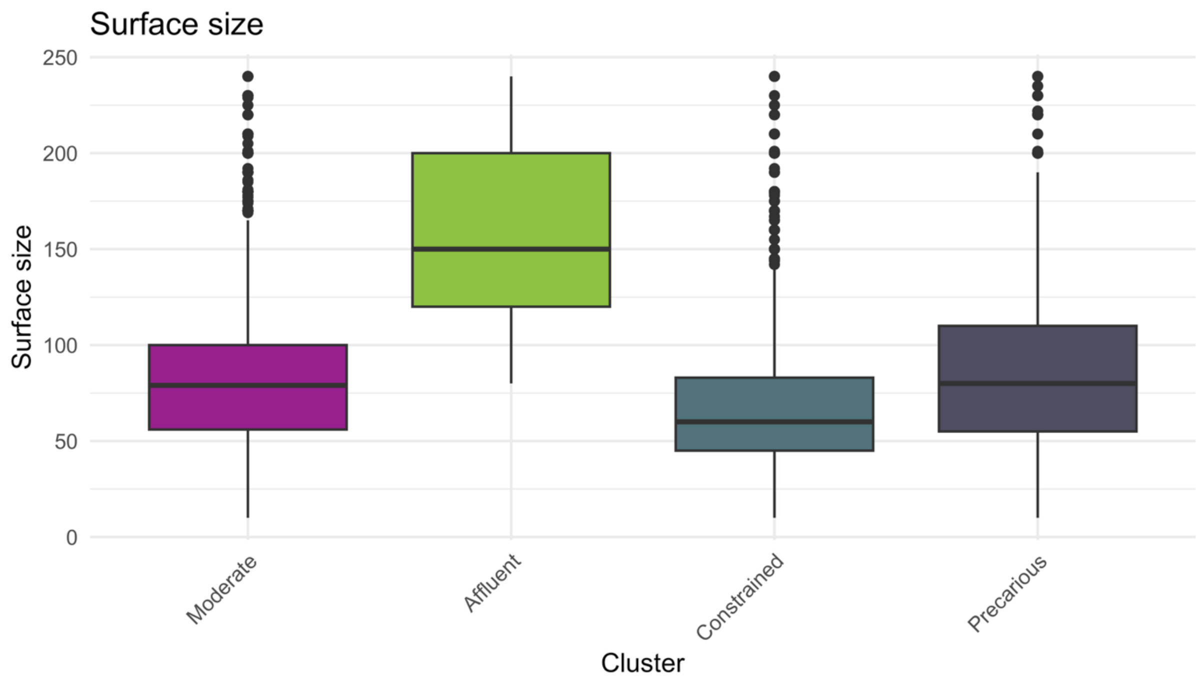

Figure 11: Cluster and surface area of dwelling

Table 11: Key characteristics of four cluster solution (representative sample)

	Moderate	Affluent	Constrained	Precarious	p-value [1]
	N = 3,561	N = 1,402	N = 1,502	N = 704	
Total amount of dwelling problems	0.15 (0.36)	0.20 (0.45)	0.45 (0.65)	2.62 (1.05)	<0.001
Energy Performance Certificate Score	5.71 (1.95)	6.10 (1.94)	5.24 (1.97)	3.64 (2.00)	<0.001
OECD persons per M2	51.36 (17.97)	118.78 (32.47)	41.37 (18.14)	51.81 (29.02)	<0.001
Space Shortage (Yes/No)	0 (0%)	100 (7.1%)	1,502 (100%)	265 (38%)	<0.001
Surface Area (m^2)	83.59 (37.22)	157.73 (42.14)	67.72 (34.81)	86.38 (44.61)	<0.001
Total Number of Rooms	3.67 (1.59)	5.06 (1.59)	2.97 (1.43)	3.67 (1.64)	<0.001
OECD family size	1.69 (0.64)	1.37 (0.37)	1.71 (0.66)	1.82 (0.71)	<0.001
OECD Income	1,935.26 (1,064.67)	2,714.96 (1,307.93)	1,672.06 (1,056.66)	1,598.95 (1,116.25)	<0.001
Age	51.30 (17.45)	55.35 (17.08)	45.72 (15.36)	45.30 (15.12)	<0.001
Employment Status					<0.001
Working	2,054 (58%)	756 (54%)	977 (65%)	430 (61%)	
Retired	970 (27%)	516 (37%)	226 (15%)	104 (15%)	
Rest	537 (15%)	130 (9.3%)	299 (20%)	170 (24%)	
Living Area					<0.001
City	1,778 (50%)	414 (30%)	859 (57%)	291 (41%)	

Rural	596 (17%)	397 (28%)	194 (13%)	135 (19%)	
Town Suburb	1,187 (33%)	591 (42%)	449 (30%)	278 (39%)	
Household Type					<0.001
Couple with Children	1,156 (32%)	308 (22%)	558 (37%)	269 (38%)	
Single Parent	150 (4.2%)	39 (2.8%)	107 (7.1%)	%) 60 (8.5	
Single	843 (24%)	396 (28%)	330 (22%)	114 (16%)	
Other	1,412 (40%)	659 (47%)	507 (34%)	261 (37%)	

1 Kruskal-Wallis rank sum test; Pearson's Chi-squared test

Precarious cluster (N = 704): Our final and smallest group, the Precarious cluster experiences the most difficult housing conditions. They have by far the highest number of dwelling problems (2.62 on average) and the lowest EPC scores (3.64), indicating poor housing quality and low energy efficiency. About 38% report a space shortage, and while their average home size (86.4 m²) is comparable to the Moderate group, housing quality is clearly lower. Their income is the lowest (€1,599), with over a third in the bottom income quartile in their respective countries. They tend to be younger (mean age 45) and show slightly higher unemployment (7%) and inability to work (3.6%) than other groups. Home ownership is limited, with only 37% owning their homes outright and 19% with mortgages, while private and social renting are common (29 and 11%, respectively). These households are split between cities (41%) and towns or suburbs (39%). Families with children and multi-person households each make up about a third of this group. Overall, this cluster reflects economic and housing insecurity, with poor dwelling conditions.

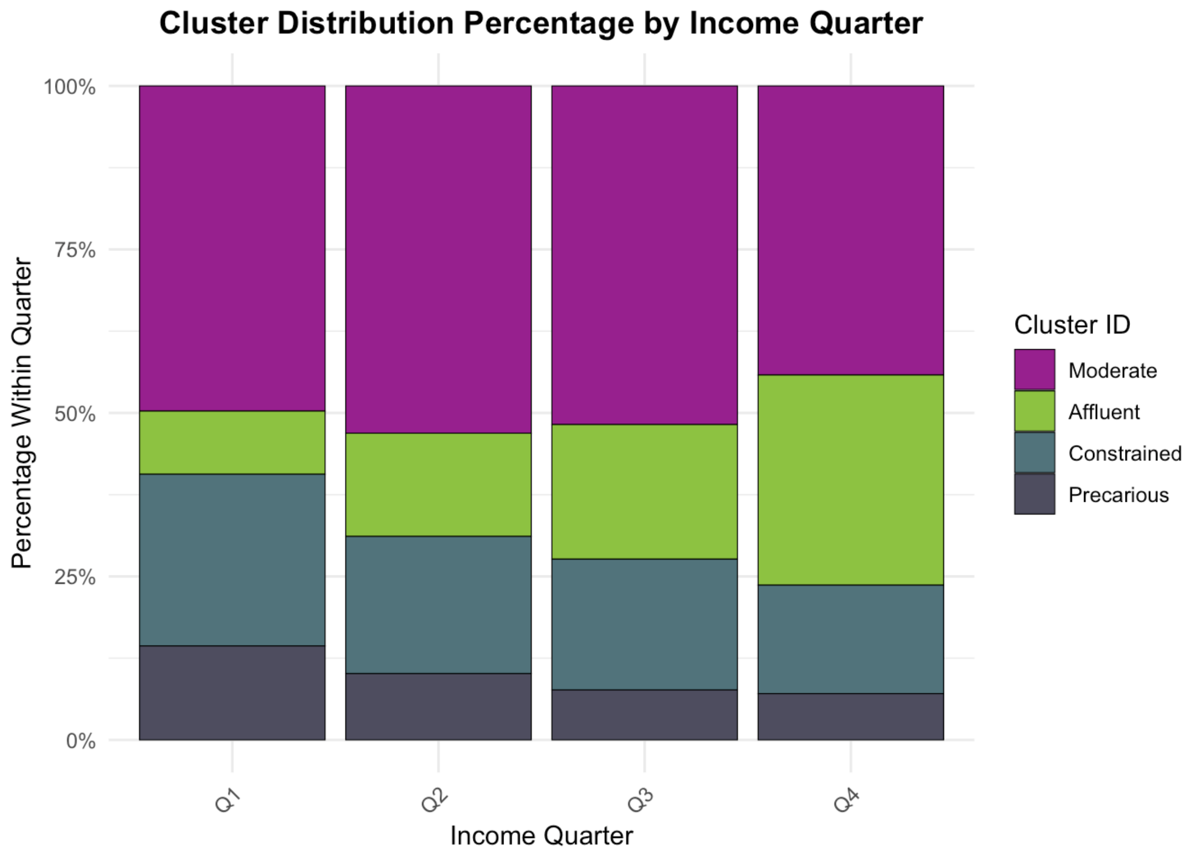

Figure 12: Distribution of cluster membership by income quarter (representative sample)

4.3 ANALYTIC APPROACH

As noted earlier, the purpose of this D4.1 is to examine how housing differences affect various aspects of people's lives, including health, mobility, and energy use. For this analysis, we use a series of multinomial logistic regression models. This modelling approach is appropriate because the outcome variable consists of multiple, unordered categories, corresponding to the housing clusters identified earlier. A linear regression framework would be unsuitable in this context, as it assumes a continuous outcome and can generate predictions outside the valid range of probabilities.

To help us understand how housing inequality relates to other aspects of life, we focus on one variable of interest at a time, assessing how it is associated with membership of the four clusters. This helps us identify whether certain outcomes are more likely among specific housing groups.

To ensure that these relationships are not distorted by other factors, each model includes a set of control variables. These are sociodemographic and contextual characteristics that may also influence both housing conditions and the outcomes we study, such as age, education, employment status, ownership type, and living area. For example, when examining how thermal energy poverty relates to cluster membership, these control variables help account for the fact that older participants, or people living in different geographic regions might also experience different thermal comfort levels, independently of their housing situation. By including such controls, we adjust for background differences between households, allowing the model to more accurately isolate the specific association between the inability to keep the home adequately warm or cool and the housing clusters. The base model therefore provides a consistent framework for comparison across all regressions.

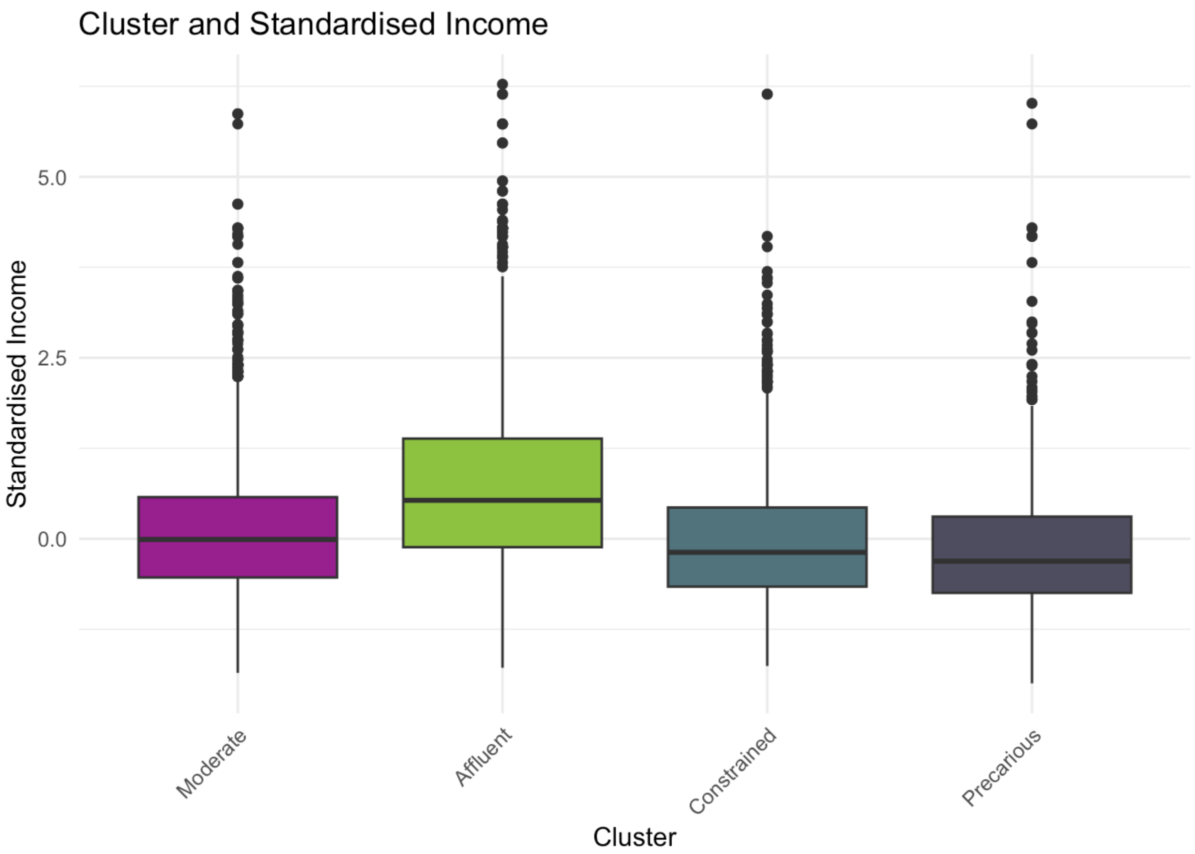

Figure 13: Distribution of cluster and standardised income (representative sample)

Moreover, our base model identifies the main sociodemographic and contextual factors linked to membership in the different housing clusters. This helps to understand the broader patterns that shape housing inequality. By looking at which demographic characteristics, housing tenure types, and geographic settings are related to being in the Affluent, Constrained, or Precarious clusters compared to the Moderate group, we can describe the social and contextual predictors behind these differences.

Cluster membership is treated as the explanatory framework that captures broader housing conditions, while outcomes such as energy poverty, health, and mobility are analysed as associated consequences. The models therefore examine whether these outcomes are more or less likely among households belonging to different housing clusters, rather than testing a strict cause–effect relationship. The underlying assumption is that housing conditions shape lived experiences and constraints, including difficulties in meeting energy needs, but the results should be read as descriptive associations rather than causal estimates. Nonetheless, partial endogeneity cannot be fully ruled out. In particular, perceived household energy efficiency (subjective EPC) is included as one of the clustering variables, while energy-related outcomes are examined in the regression models. This overlap reflects the conceptual link between housing quality and energy-related hardship, but it also implies that the estimated associations may partly capture shared underlying dimensions of housing conditions rather than independent effects. For this reason, the regressions are intended to support pattern identification and comparison across clusters.

Table 12: Grouping variables for education, employment status and household type

Original Category	Grouped Category
Education	
Lower Secondary	Limited Education
None	Limited Education
Primary	Limited Education
Upper Secondary	Secondary
Vocational	Higher Non-University
Bachelors	Higher University
Doctorate	Higher University
Masters	Higher University
Employment status	
Employed Full Time	Working
Employed Part Time	Working
Self Employed	Working
Retired	Retired
Household Work	Rest
Other	Rest
Student	Rest
Temporary Leave	Rest
Unable to Work	Rest
Unemployed	Rest
Household type	
Couple With Children	Couple with Children
Single Parent	Single Parent
Single	Single
Couple No Children	Other
Living With Parents	Other
Non-Family Household	Other

Table 11 summarises our categorical variables such as employment, some categories were grouped together (for Education, Household Type, and Employment Status) to ensure adequate sample sizes within each cluster. The final groupings are shown in Table 12.

While the cluster analysis was performed on the full dataset (n = 14,888), the analyses presented in this section use the representative subsample (n = 7,169). In some cases, only a subset of participants could respond to specific questions—for example, tenants and homeowners were asked separately about reasons for moving during the COVID-19 pandemic. These instances are reflected in the corresponding sample sizes. A small number of participants also declined to answer certain questions. For instance, 11 participants (0.2%) did not provide information on their general health. These cases were excluded from the health analysis but retained for all others. As a result, the total sample size is slightly lower than 7,169 in a few analyses.

For each model, we report the pseudo-R^2, which indicates how well the model explains variation in cluster membership. Higher values suggest a better fit, though the measure is not directly comparable to the R^2 used in linear regression. It should therefore be interpreted as a relative measure for comparing models estimated on the same data, rather than as an absolute measure of explained variance.

We also present the residual deviance and the Akaike Information Criterion (AIC), which provide additional measures of model fit and allow for comparison between models. Lower values of these statistics generally indicate a better-fitting model. Given the large sample size (n = 7,169), the threshold for statistical significance is set at $p \leq 0.01$. Beyond the multinomial regressions, certain results are accompanied by additional statistical tests where appropriate.

A common concern in regression analysis is multicollinearity, where predictors are highly correlated with one another. To assess this, the generalised variance inflation factor (GVIF) was calculated for all predictors in the base model and adjusted for degrees of freedom (adjusted GVIF)(Cohen et al., 2003). In the first model, which used the continuous age variable, the results indicated strong associations between age, education, and employment status (GVIF = 26.47 for Age; 12.98 for Education; and 4.41 for Employment). This suggested that age overlapped with other sociodemographic variables, potentially introducing bias into the estimates.

Table 13: VIF for before and after grouping age

	Initial model		**Final model**	
Predictor	GVIF	Adjusted GVIF	GVIF	Adjusted GVIF
Gender	2.14	1.46	2.14	1.46
Age / Age Bins	26.47	5.14	20.75	1.66
Employment status	4.41	1.45	6.09	1.57
Education level	12.98	1.53	13.10	1.54
Living area	4.11	1.42	4.11	1.42
Household Type	5.59	1.33	5.76	1.34
Ownership status	3.66	1.18	3.66	1.18

OECD Equivalised income (1000s)	6.11	2.47	6.11	2.47
Country of residence	14.90	1.21	15.05	1.21

To address this, age was recoded into categorical groups consistent with those used in constructing the representative sample. After this adjustment, multicollinearity decreased notably. The recalculated VIF values were all well below the common threshold of 5, with most below 2. The highest adjusted values were for education (1.54), employment (1.57), and income (2.47), while all other predictors remained closer to 1.3 or lower. These results, shown in Table 13, indicate that multicollinearity is not a concern in the final specification of the model.

4.4 IMPACT OF SOCIODEMOGRAPHIC VARIABLES

As a first step, the analysis examines how key sociodemographic and contextual characteristics relate to membership of the housing clusters (Table 14). The multinomial logistic regression results (Nagelkerke R^2 = 0.30, N = 7,169) show distinct differences across the four clusters, with the Moderate group serving as the reference category.

Table 14: Multinomial regression for our sociodemographic base model

Predictor	Affluent	Constrained	Precarious
(Intercept)	0.17**	0.49*	0.22**
Male[1]	1.03	0.86	0.71**
36–49[2]	0.99	1.02	1.2
50–65[2]	1.24	0.79*	0.70*
65+[2]	1.35	0.50**	0.41**
Employment: Retired[3]	1.14	0.93	1.12
Employment: Other/Non-Active/Temp. Leave[3]	0.96	1.01	1.13
Education: Secondary[4]	1	1.06	1.04
Education: Higher Non-University[4]	0.98	0.93	0.9
Education: Higher University[4]	1.14	0.95	0.88
Living Area: Rural[5]	2.36**	0.76*	1.36
Living Area: Town/Suburb[5]	1.63**	0.91	1.45**
Household Type: Single Parent[6]	1.31	1.3	1.44
Household Type: Single[6]	2.71**	0.92	0.77
Household Type: Other[6]	1.67**	0.9	0.97
Ownership: Owned, With Mortgage[7]	1.05	1.47**	1.12
Ownership: Rented: Social[7]	0.26**	1.96**	2.08**

Ownership: Rented: Private[7]	0.31**	1.72**	2.07**
Ownership: Free Residence[7]	0.55	1.58	1.69
Income (1000s €)	1.68**	0.88**	0.87*
Country: Czechia[8]	0.24**	1.54	1.36
Country: Germany[8]	0.28**	0.71	0.57
Country: Estonia[8]	0.17**	1.64*	0.72
Country: Finland[8]	0.14**	0.88	0.15**
Country: United Kingdom[8]	0.26**	1.2	1.70*
Country: Italy[8]	0.47**	1.38	1.90*
Country: Romania[8]	0.27**	2.13**	2.30**

Reference cluster: Moderate

Reference Categories: [1]Gender: Female; [2]Age: 18-35; [3]Employment: Working; [4]Education: Primary/None; [5]Living Area: City Centre; [6]Household Type: Couple with Children; [7]Ownership: Owned, No Mortgage; [8]Country: Belgium

Notes: Results displayed as odds ratios. Nagelkerke Pseudo R2 = 0.300; Residual Deviance = 15,218.05; AIC = 15,380.05; * $p \leq 0.01$; ** $p \leq 0.001$. N = 7,169

Gender and age had notable but uneven effects. Men were less likely than women to belong to the Precarious cluster (OR = 0.71, $p \leq 0.001$), while no significant gender difference appeared for the Affluent or Constrained clusters. Older individuals were less likely to be Precarious or Constrained compared with younger respondents. For example, people aged 65 and over had lower odds of being Constrained (OR = 0.50, $p \leq 0.001$) or Precarious (OR = 0.41, $p \leq 0.001$). The same age group was somewhat more likely to be Affluent (OR = 1.35).

Employment status showed limited influence. Being retired or temporarily inactive did not strongly predict cluster membership once other factors were controlled. Education was also not a strong differentiator: neither secondary nor higher education levels were significantly linked to any specific cluster when compared with primary or no education.

Location played a more consistent role. Living in a rural area was associated with higher odds of being in the Affluent cluster (OR = 2.36, $p \leq 0.001$) but lower odds of being in the Constrained cluster (OR = 0.76, $p \leq 0.01$). Those living in towns or suburbs were also more likely to be Affluent (OR = 1.63, $p \leq 0.001$) and Precarious (OR = 1.45, $p \leq 0.001$) compared with city-centre residents.

Household type was another factor distinguishing clusters. Single-person households had much higher odds of being Affluent (OR = 2.71, $p \leq 0.001$) compared with couples with children, while single parents showed no clear difference. Other household types (for instance, shared or multigenerational) were also more likely to be Affluent (OR = 1.67, $p \leq 0.001$). No strong effects were observed for Constrained or Precarious households in this regard.

Tenure type showed some of the strongest associations. Relative to outright homeowners, those with a mortgage were more likely to belong to the Constrained cluster (OR = 1.47, $p \leq 0.001$). Renting was linked to disadvantage: social renters had much lower odds of being in the Affluent cluster (OR = 0.26, $p \leq 0.001$) but much higher odds of belonging to the Constrained (OR = 1.96, $p \leq 0.001$) or Precarious (OR = 2.08, $p \leq 0.001$) clusters. The same pattern appeared for private renters. These results underline the role of tenure insecurity in explaining differences between housing clusters.

Income was one of the most consistent predictors. Each additional thousand euros of household income increased the odds of being Affluent (OR = 1.68, $p \leq 0.001$) and reduced the odds of being Constrained (OR = 0.88, $p \leq 0.001$) or Precarious (OR = 0.87, $p \leq 0.01$) when compared with the Moderate cluster. This aligns with the descriptive findings showing that Affluent households have larger, higher-quality dwellings, while Precarious households face greater space and energy problems.

Finally, country of residence also influenced cluster membership. Compared with Belgium, our reference country, respondents from most other countries were less likely to be Affluent, including those from Czechia, Germany, Estonia, Finland, United Kingdom, Italy, and Romania (all $p \leq 0.001$). However, residents of Romania (OR = 2.13, $p \leq 0.001$) and Estonia (OR = 1.64, $p \leq 0.01$) were more likely to be Constrained, while those from Romania, Italy, and United Kingdom were more likely to be Precarious. Finland showed an opposite pattern, with respondents being less likely to belong to the Precarious group (OR = 0.15, $p \leq 0.001$).

Overall, income and tenure were the strongest predictors of cluster membership, followed by living area and age. Higher-income homeowners, particularly those in rural or suburban settings, were more often Affluent. Middle-income households with mortgages or rental contracts tended to be Constrained, while younger, lower-income renters—especially in certain countries—were more likely to be Precarious. These associations reflect the structural links between income, housing security, and residential conditions across Europe.

Consistent with our clustering, the Precarious cluster represents households facing the most vulnerable housing situations. Members of this group were generally younger, more often female, and had lower incomes than those in the Moderate cluster. Renting was the defining feature of this group: both social and private tenants were more than twice as likely to belong to the Precarious cluster compared with outright homeowners. Income had a clear protective effect, with even small increases reducing the likelihood of being Precarious. Unlike the Affluent group, which was more common in rural areas, Precarious households were concentrated in towns and suburban areas.

It is important to interpret the gender effect with caution. Since the questionnaire was completed at the household level, the reported gender identifies the respondent rather than all household members. The observed associations between gender and cluster membership therefore cannot be taken as evidence of gendered differences in housing conditions. Instead, they likely capture which household member completed the survey.

However, cross-tabulation between gender and household type indicates that the gender distribution is broadly balanced for most household types but differs markedly among single-parent households. Among couples with children, 1,101 women and 1,190 men completed the questionnaire; among single-person households, 868 women and 815 men responded. For other household types, such as shared or multigenerational arrangements, the distribution was also nearly even (1,424 women and 1,415 men). However, single-parent households were disproportionately represented by women, with 269 female and 87 male respondents. As might be expected, the association between gender and household type was statistically significant ($\chi^2 = 94.893$, df = 3, $p < 0.001$), suggesting that part of the observed gender pattern may stem from this imbalance.

Finally, while we the focus on this work is on the representative sample, we additionally performed the same analysis as shown in Table 14 on our complete sample (seen on page 95). By and large, results concur, yet we briefly discuss some differences.

The clearest shifts appeared for the Precarious cluster. In the representative sample, people aged 50–65 had lower odds of being Precarious (OR = 0.70, $p \leq 0.01$), while this effect was not significant in the complete sample, where the estimate was closer to 1 (OR = 0.87). A similar change was observed for rural residents. In the complete sample, those living in rural areas were more likely to be Precarious (OR = 1.57, $p \leq 0.001$), but this association was no longer significant in the representative sample (OR = 1.36). Income remained a protective factor in both datasets, though the strength of the association was somewhat weaker after stratification: each additional thousand euros reduced the odds of being Precarious to OR = 0.87 ($p \leq 0.01$) in the representative sample, compared with OR = 0.81 in the complete sample.

For the Affluent cluster, most results were stable, but a few adjustments appeared. In the complete sample, single-parent households were more likely to be Affluent (OR = 1.53, $p \leq 0.001$), whereas this association was not significant in the representative sample (OR = 1.31). Other predictors, such as the higher likelihood of single-person households being Affluent (OR = 2.71, $p \leq 0.001$), remained very similar across both models.

The Constrained cluster presented a clearer case where weighting strengthened an existing pattern. Households with a mortgage were more likely to be Constrained in both datasets, but the effect was stronger in the representative sample (OR = 1.47, $p \leq 0.001$) than in the complete sample (OR = 1.28). This indicates that mortgage holders, who are more common in population-representative samples, form a more central part of the Constrained group than suggested by the unweighted data. Some country differences also shifted in magnitude. For example, residents of Romania had higher odds of being Constrained in both datasets, but the association increased from OR = 1.88 ($p \leq 0.001$) in the complete sample to OR = 2.13 ($p \leq 0.001$) in the representative sample. In contrast, the corresponding estimate for Czechia weakened and lost significance (OR = 1.47, $p \leq 0.001$ in the complete sample; OR = 1.54, not significant in the representative sample). However, the broadly the results are in line, with only slight changes and robust predictors like income only weakening or strengthening slightly.

4.5 IMPACT ON ENERGY POVERTY

4.5.1 CONCEPTUAL BACKGROUND

One key question the deliverable aims to address is how energy poverty and housing inequalities are related. Energy poverty is closely linked to housing conditions and inequalities in the housing sector. It generally concerns the affordability of and access to energy and energy services (see, e.g., Bouzarovski & Petrova, 2015; Martiskainen et al., 2021; Simcock, Frankowski, et al., 2021; Simcock, Jenkins, et al., 2021). According to (Bouzarovski et al., 2021), it is "a condition in which a household is unable to secure a socially and materially needed level of energy services in the home".

This condition is often conceptualised through the so-called energy poverty triangle, which highlights the interaction between household resources, energy prices and systems, and housing characteristics (Boardman, 2013). Within this framework, affordability and accessibility emerge as key dimensions, shaped by limited income and high energy costs on the one hand, and by inadequate infrastructure, building quality, and access to modern energy services on the other. Affordability relates to limited income, e.g., in the face of high energy costs. It can be measured based on a household's energy expenditure related to disposable income, but also through being in arrears with energy bills or more subjectively based on the perception of energy costs being a burden for households. Accessibility depends on adequate infrastructure and reliable supply and can refer to the access to renewable energy sources as opposed to fossil fuels or to relevant energy services. Energy services encompass functions like heating, cooling, lighting, cooking or mobility, which enable basic comfort and dignity (Fell, 2017) . Meeting these needs without excessive spending or forced reductions in consumption is essential for maintaining a decent living standard (Brown et al., 2020; Jiglau et al., 2024; Krugmann & Goldemberg, 1983; Shyu, 2024) . While these definitions capture households that currently experience energy poverty, they do not fully reflect vulnerability to energy poverty, which refers to the risk of experiencing inadequate energy services in the future due to the interaction between building characteristics and households' socioeconomic and adaptive capacities, even when current energy needs are still met (Della Valle et al., 2025).

Poor insulation, structural defects, or low energy efficiency and limited access to modern energy systems can exacerbate energy poverty, making it not only an economic issue but also a structural one. In this context, we examine two dimensions of energy poverty, operationalized as (1) financial energy poverty, where energy costs become a burden or lead to arrears, and (2) thermal energy poverty, where households cannot maintain adequate indoor temperatures. To explore these dimensions in detail, the following sections present the regression results for the two indicators.

To quantify these dimensions, empirical research predominantly draws on several harmonized pan-European statistical instruments. The primary source for consensual indicators—such as the self-reported inability to keep a home warm or the presence of utility arrears—is the EU Statistics on Income and Living Conditions (EU-SILC), which allows for longitudinal monitoring across Member States (Thomson et al., 2017). These are often supplemented by Household Budget Surveys (HBS) to capture actual energy expenditure data, and the European Quality of Life Survey (EQLS) or Eurobarometer modules to assess broader social impacts and subjective well-being. By utilizing these micro-data sources, it is possible to link specific household socioeconomic capacities with structural housing characteristics.

4.5.2 FINANCIAL ENERGY POVERTY

A variety measures to define energy poverty. These might include the Ten Percent Rule (TPR)(Boardman, 1991). The TPR considers a household to be fuel poor when it must allocate 10% or more of its disposable income to housing-related energy services, including electricity, space heating, and water heating. Other alternatives include the Low Income/High Cost (LIHC) indicator classifies a household as energy poor when its spending on energy services is higher than the median across households and, after covering these energy costs, its remaining income falls below the official poverty threshold (Hills, 2012). Moore (2012), by contrast proposes a budget-standard approach based on Minimum Income Standards (MIS), in which a household is considered energy poor if, after housing costs and other minimum living expenses are deducted, its remaining income is insufficient to cover required energy costs. See also Moore (2012) for an expansive discussion fuel poverty definitions.

In this deliverable, we understand a person to be financially energy poor when energy costs represent a burden or when they experience arrears on energy bills. For analytical purposes, we apply a less strict operationalisation by classifying respondents as highly financially energy poor if they agree or strongly agree that their energy costs (e.g., for heating, cooling, hot water, and electricity) represent a major financial burden for them.

When controlling for different socio-demographic variables and country of residence, financial energy poverty is related to the housing inequality clusters. When financial energy poverty is high, the odds of being in the Affluent cluster increase by 19% (OR = 1.19, +19%, p ≤ 0.001) compared to the Moderate cluster. The associations are similar but larger for the Constrained and Precarious group. Compared to respondents who face low financial energy poverty, for respondents who are confronted with energy poverty the odds are 76% higher to be in the Constrained group (OR = 1.76, +76%, p ≤ 0.001) and 177% higher to be in the Precarious group (OR = 2.77, +177%, p ≤ 0.001) compared to the Moderate group. The results are unexpected for the Affluent group but expected for the Constrained and Precarious groups.

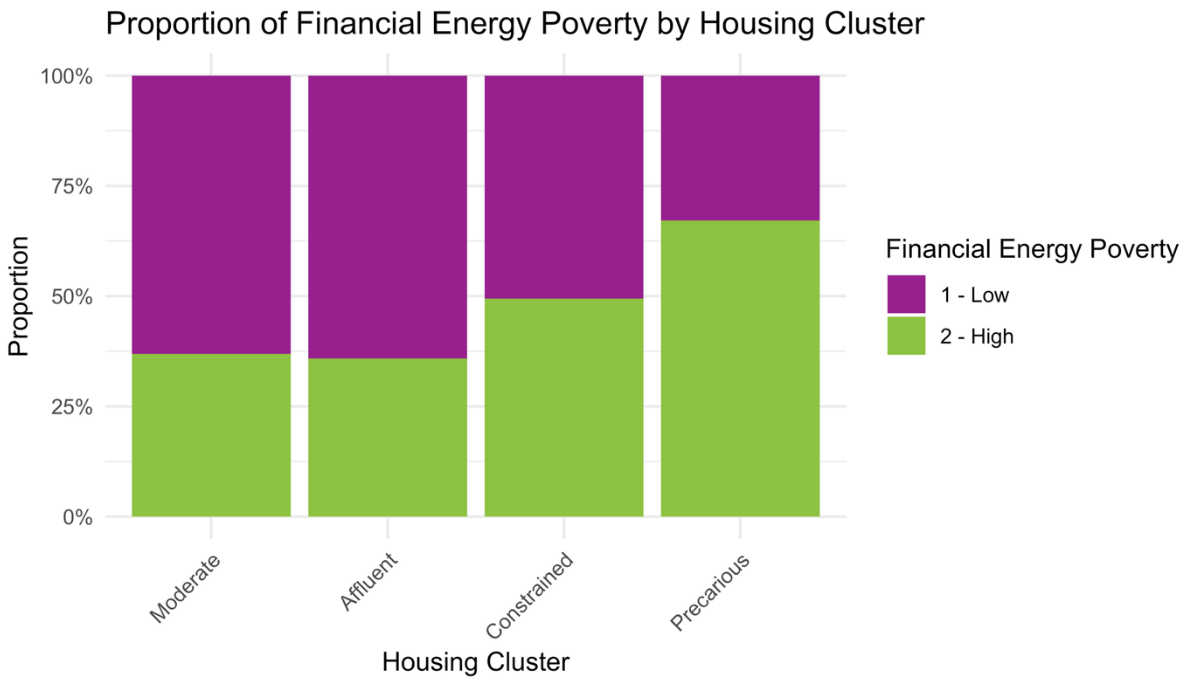

Figure 14: Proportion of respondents per cluster suffering from financial energy poverty

Predictor	Affluent	Constrained	Precarious
(Intercept)	0.16**	0.43**	0.14**
Male[1]	1.06	0.88	0.73**
36–49[2]	1.07	0.98	1.14
50–65[2]	1.37*	0.75*	0.67*
65+[2]	1.38	0.47**	0.37**
Employment: Retired[3]	1.17	0.93	1.15

Employment: Other/Non-Active/Temp. Leave[3]	0.98	1.01	1.11
Education: Secondary[4]	0.99	1.05	1.01
Education: Higher Non-University[4]	1.00	0.90	0.86
Education: Higher University[4]	1.13	0.95	0.86
Living Area: Rural[5]	2.24**	0.74*	1.31
Living Area: Town/Suburb[5]	1.50**	0.89	1.40*
Household Type: Single Parent[6]	1.42	1.25	1.32
Household Type: Single[6]	2.66**	0.96	0.76
Household Type: Other[6]	1.66**	0.91	0.97
Ownership: Owned, With Mortgage[7]	1.11	1.42**	1.06
Ownership: Rented: Social[7]	0.27**	1.89**	1.94**
Ownership: Rented: Private[7]	0.30**	1.64**	1.87**
Ownership: Free Residence[7]	0.48*	1.49	1.59
Income (1000s €)	1.69**	0.92	0.92
Country: Czechia[8]	0.22**	1.46	1.39
Country: Germany[8]	0.27**	0.64*	0.54*
Country: Estonia[8]	0.15**	1.41	0.61
Country: Finland[8]	0.14**	0.88	0.18**
Country: United Kingdom[8]	0.25**	1.05	1.56
Country: Italy[8]	0.45**	1.12	1.53
Country: Romania[8]	0.23**	1.99**	2.26**
Financial Energy Poverty: High[9]	1.19	1.76**	2.77**

Reference cluster: Moderate
Reference Categories:
[1]Gender: Female; [2]Age: 18-35; [3]Employment: Working; [4]Education: Primary/None;
[5]Living Area: City Centre; [6]Household Type: Couple with Children; [7]Ownership: Owned, No Mortgage; [8]Country: Belgium; [9]Financial Energy Poverty: Low.
Notes: Results displayed as odds ratios. Nagelkerke Pseudo R^2 = 0.320; Residual Deviance = 15,103.42; AIC = 15,271.42; * $p \leq 0.01$; ** $p \leq 0.001$. N = 7,169

4.5.3 THERMAL ENERGY POVERTY

We use thermal energy poverty to describe whether a person is able to keep their home warm in winter or cool in summer.

Table 15: Multinomial regression evaluating the impact of thermal energy poverty

Predictor	Affluent	Constrained	Precarious
(Intercept)	0.16**	0.33**	0.10**
Male[1]	1.04	0.88	0.76*
36–49[2]	0.99	1.00	1.18

50–65[2]	1.24	0.75*	0.67*
65+[2]	1.35	0.49**	0.42**
Employment: Retired[3]	1.13	0.91	1.07
Employment: Other/Non-Active/Temp. Leave[3]	0.96	0.99	1.13
Education: Secondary[4]	1.00	1.06	1.06
Education: Higher Non-University[4]	0.97	0.94	0.93
Education: Higher University[4]	1.15	0.97	0.95
Living Area: Rural[5]	2.34**	0.76*	1.36
Living Area: Town/Suburb[5]	1.63**	0.89	1.42*
Household Type: Single Parent[6]	1.29	1.29	1.33
Household Type: Single[6]	2.69**	0.89	0.69*
Household Type: Other[6]	1.68**	0.86	0.90
Ownership: Owned, With Mortgage[7]	1.04	1.44**	1.09
Ownership: Rented: Social[7]	0.26**	1.86**	1.82**
Ownership: Rented: Private[7]	0.31**	1.64**	1.82**
Ownership: Free Residence[7]	0.55	1.47	1.47
Income (1000s €)	1.71**	0.93	0.97
Country: Czechia[8]	0.25**	1.52	1.41
Country: Germany[8]	0.28**	0.68	0.53*
Country: Estonia[8]	0.18**	1.55	0.75
Country: Finland[8]	0.14**	0.89	0.18**
Country: United Kingdom[8]	0.26**	1.22	1.58
Country: Italy[8]	0.47**	1.38	1.74
Country: Romania[8]	0.27**	2.14**	2.24**
Thermal Energy Poverty[9]: Only warm	0.87	1.73**	1.60**
Thermal Energy Poverty[9]: Only cool	1.29	1.63**	2.87**
Thermal Energy Poverty[9]: Neither warm nor cool	1.17	2.07**	4.21**

Reference cluster: Moderate
Reference Categories: [1]Gender: Female; [2]Age: 18-35; [3]Employment: Working; [4]Education: Primary/None; [5]Living Area: City Centre; [6]Household Type: Couple with Children; [7]Ownership: Owned, No Mortgage; [8]Country: Belgium; [9]Thermal Energy Poverty: Warm and cool.
Notes: Results displayed as odds ratios. Nagelkerke Pseudo R^2 = 0.330; Residual Deviance = 14,988.41; AIC = 15,168.41; * $p \leq 0.01$; ** $p \leq 0.001$. N = 7,169

Controlling for socio-demographic variables and the country of residence, thermal energy poverty is statistically significantly related to the housing clusters for most clusters. Compared to respondents who can keep their home warm and cool, the odds for respondents who can only keep their home warm in winter, but cannot keep cool in summer, to be in the Constrained housing cluster increase by 73% (OR = 1.73, +73%, p ≤

0.001) and to be in the Precarious cluster increase by 60% (OR = 1.60, +60%, p ≤ 0.001) compared to the Moderate cluster.

Similarly, those who can only keep cool in winter but cannot keep warm in summer are more likely to be in the Constrained (OR = 1.63, +63%, p ≤ 0.001) or Precarious cluster (OR = 2.87, +187%, p ≤ 0.001) compared to the Moderate cluster. The relationships are even stronger for respondents who can neither keep warm nor cool - representing the strongest form of thermal energy poverty we investigate. If a respondent is neither able to keep warm nor cool the odds to be in the Constrained cluster increase by 107% (OR = 2.07, +107%, p ≤ 0.001) and the odds to be in the Precarious housing cluster increase by 321% (OR = 4.21, +321%, p ≤ 0.001) compared to the Moderate cluster.

For the Affluent cluster, in contrast, we do not find a statistically significant association. Hence, the odds to be in the Affluent housing cluster compared to the Moderate cluster do not change statistically significantly for different levels of thermal energy poverty.

The results highlight that thermal energy poverty is related to other forms of housing inequality, measured through the different housing clusters.

4.6 IMPACT ON MOBILITY

A second key concern relates to housing inequality and mobility. Here we assess three key aspects: the ownership of cars, along with the availability of various public transport options and finally the effective use of public transport.

Table 16: Multinomial regression evaluating the impact of car ownership

Predictor	Affluent	Constrained	Precarious
(Intercept)	0.14**	0.53	0.20**
Male[1]	0.99	0.81*	0.66**
36–49[2]	0.99	0.97	1.16
50–65[2]	1.24	0.77*	0.66*
65+[2]	1.40	0.44**	0.40**
Employment: Retired[3]	1.26	0.98	1.23
Employment: Other/Non-Active/Temp. Leave[3]	0.96	0.95	1.11
Education: Secondary[4]	1.00	1.15	1.18
Education: Higher Non-University[4]	0.98	0.93	1.03
Education: Higher University[4]	1.11	0.96	0.89
Living Area: Rural[5]	2.23**	0.82	1.42
Living Area: Town/Suburb[5]	1.56**	1.01	1.47*
Household Type: Single Parent[6]	1.25	1.15	1.63
Household Type: Single[6]	2.97**	0.83	0.70
Household Type: Other[6]	1.58**	0.84	0.97
Ownership: Owned, With Mortgage[7]	1.05	1.54**	1.21
Ownership: Rented: Social[7]	0.26**	2.03**	2.39**

Ownership: Rented: Private[7]	0.31**	1.72**	2.11**
Ownership: Free Residence[7]	0.54	1.64	2.17*
Income (1000s €)	1.71**	0.92	0.85*
Country: Czechia[8]	0.31**	1.90*	1.57
Country: Germany[8]	0.31**	0.80	0.70
Country: Estonia[8]	0.20**	1.85*	0.63
Country: Finland[8]	0.16**	0.97	0.15**
Country: United Kingdom[8]	0.25**	1.29	1.52
Country: Italy[8]	0.48**	1.70*	2.04*
Country: Romania[8]	0.29**	2.46**	2.33*
Amount of cars	1.10	0.82*	1.00

Reference cluster: Moderate
Reference Categories: [1]Gender: Female; [2]Age: 18-35; [3]Employment: Working; [4]Education: Primary/None; [5]Living Area: City Centre; [6]Household Type: Couple with Children; [7]Ownership: Owned, No Mortgage; [8]Country: Belgium.
Notes: Results displayed as odds ratios. Nagelkerke Pseudo R^2 = 0.310; Residual Deviance = 12,352.64; AIC = 12,520.64; * $p \leq 0.01$; ** $p \leq 0.001$. N = 5854

Car ownership patterns differ slightly across the housing clusters once other factors such as income, living area, and household type are taken into account. The number of cars owned does not significantly distinguish the Moderate, Affluent, or Precarious groups. However, respondents in the Constrained cluster show a lower likelihood of belonging to that group as the number of cars increases, with each additional vehicle associated with an 18% reduction in the odds of being in the Constrained cluster compared with the Moderate group.

Table 17: Multinomial regression evaluating the impact of public transport use

Predictor	Affluent	Constrained	Precarious
(Intercept)	0.18**	0.45**	0.20**
Male[1]	1.02	0.86	0.71**
36–49[2]	0.96	1.03	1.22
50–65[2]	1.21	0.80	0.72*
65+[2]	1.34	0.50**	0.41**
Employment: Retired[3]	1.14	0.93	1.12
Employment: Other/Non-Active/Temp. Leave[3]	0.97	1.00	1.12
Education: Secondary[4]	1.02	1.05	1.03
Education: Higher Non-University[4]	0.99	0.93	0.89
Education: Higher University[4]	1.18	0.94	0.86
Living Area: Rural[5]	2.21**	0.79	1.43*
Living Area: Town/Suburb[5]	1.57**	0.93	1.49**
Household Type: Single Parent[6]	1.29	1.29	1.43
Household Type: Single[6]	2.75**	0.90	0.76

Household Type: Other[6]	1.66**	0.89	0.97
Ownership: Owned, With Mortgage[7]	1.04	1.47**	1.13
Ownership: Rented: Social[7]	0.27**	1.93**	2.06**
Ownership: Rented: Private[7]	0.31**	1.72**	2.06**
Ownership: Free Residence[7]	0.55	1.59	1.70
Income (1000s €)	1.68**	0.88**	0.87*
Country: Czechia[8]	0.26**	1.49	1.31
Country: Germany[8]	0.28**	0.70	0.56
Country: Estonia[8]	0.17**	1.59*	0.70
Country: Finland[8]	0.14**	0.88	0.15**
Country: United Kingdom[8]	0.27**	1.17	1.65
Country: Italy[8]	0.47**	1.39	1.92*
Country: Romania[8]	0.28**	2.07**	2.25**
Public Transport Use: Yes[9]	0.81*	1.17	1.17

Reference cluster: Moderate
Reference Categories: [1]Gender: Female; [2]Age: 18-35; [3]Employment: Working; [4]Education: Primary/None; [5]Living Area: City Centre; [6]Household Type: Couple with Children; [7]Ownership: Owned, No Mortgage; [8]Country: Belgium; [9]Public Transport Use: No
Notes: Results displayed as odds ratios. Nagelkerke Pseudo R^2 = 0.300; Residual Deviance = 15,199.17; AIC = 15,367.17; * $p \leq 0.01$; ** $p \leq 0.001$. N = 7,169

When controlling for socio-demographic factors and dwelling location, differences in public transport use between the clusters are similarly limited. Individuals in the Affluent cluster are somewhat less likely to report using public transport compared to those in the Moderate group (19%, OR = 0.81, $p \leq 0.01$). In contrast, members of the Constrained and Precarious clusters show slightly higher odds of reporting public transport use (both OR = 1.17), although these differences are not statistically significant. After accounting for factors such as income, housing tenure, and urban–rural location, public transport thus does not clearly differentiate the Constrained, Precarious, and Moderate clusters.

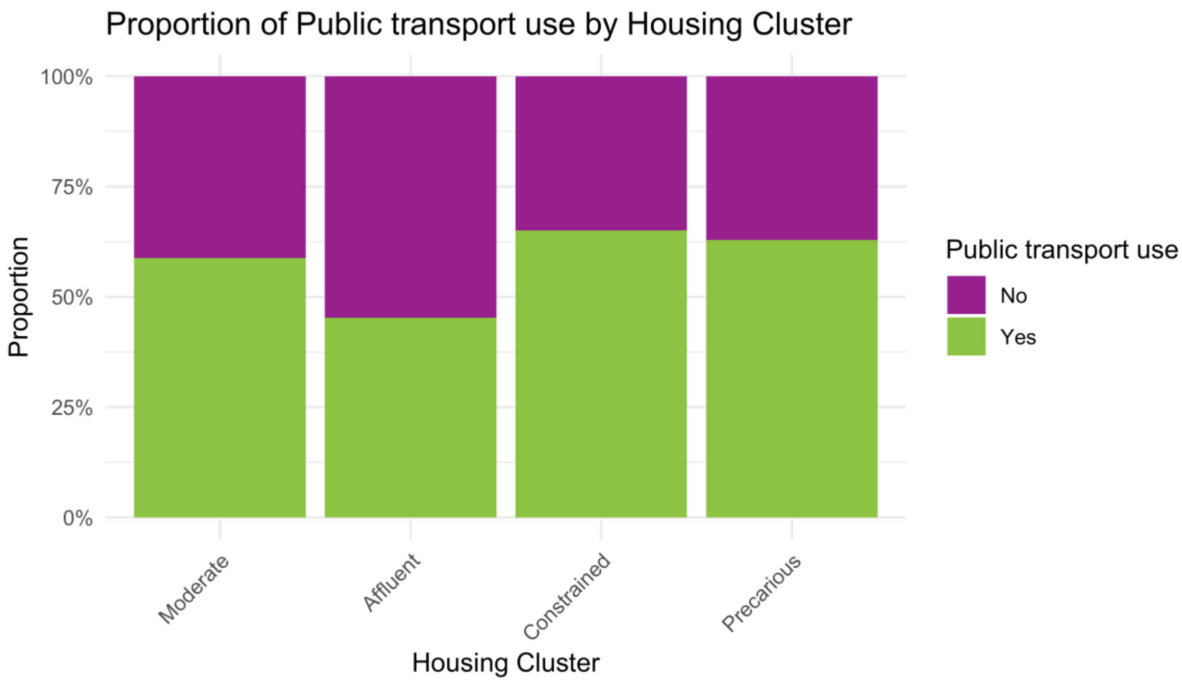

Figure 15: Proportion of respondents per cluster using public transport

We additionally examined the availability of public transport options, including bus, train, tram or metro, other modes, and the absence of any nearby public transport. For bus, train, other modes, and no availability, we found no statistically significant differences across clusters, with odds ratios close to one. Only tram or metro availability showed a significant result: households in the Affluent cluster were less likely to report access to a tram or metro (OR = 0.74, $p < 0.01$) compared to those in the Moderate group. Very few participants (between 2% and 3%) indicated "other" modes of public transport, and only a small share reported having no public transport nearby (3% in the Precarious and Constrained clusters, 4% in the Moderate cluster, and 5% in the Affluent cluster). These small group sizes limit the interpretability of these results. Overall, public transport availability appears broadly similar across housing clusters, with only a modest difference for tram or metro access among the Affluent cluster.

Table 18: Public transport availability across the four clusters

	Affluent	Constrained	Precarious
Public transport: Bus available	0.9	0.94	1.23
Public transport: Train available	0.93	1	1.12
Public transport: Tram/Metro available	0.74*	1.07	0.89
Public transport: Other mode available	0.66	0.86	0.82
Public transport: None available	1.24	1.03	0.99

Reference cluster: Moderate, Model results only display the variable of interest, but analysis includes the complete base model. Results displayed as odds ratios

Our final consideration is the use of public transport, covering buses, trains, trams or metros, and other modes. Use of public buses was more likely among Constrained households (OR = 1.22, $p < 0.01$) compared to the Moderate cluster. Those in the Precarious cluster reported similarly high odds (OR = 1.25), although this did not reach statistical significance ($p = 0.02$). Conversely, bus use was 17% less likely among the Affluent group compared to the Moderate cluster (OR = 0.83), but this result also fell just above the significance threshold ($p = 0.014$).

Train use showed a comparable pattern, with no significant differences for most clusters except for the Precarious group, which reported higher odds of train use (OR = 1.33, $p < 0.01$). Tram or metro use was significantly less common among Affluent households (OR = 0.73, $p < 0.01$), while no significant differences were found for the Constrained or Precarious clusters. Other modes of public transport were rarely reported, with incidence rates between 1% and 2% across clusters and no statistically significant variation. Overall, Constrained and Precarious households appear somewhat more reliant on bus and train travel respectively, while Affluent households use trams or metros less frequently when compared to our reference Moderate cluster. Recall that our results include our base set of control variables, so considers other additional factors such as income, dwelling location and type or household type.

Table 19: Public transport use among the four housing clusters

	Affluent	Constrained	Precarious
Public transport use: Bus	0.83	1.22*	1.25
Public transport use: Train	0.92	0.97	1.33*
Public transport use: Tram/Metro	0.73*	0.96	0.82
Public transport use: Other mode	0.57	0.9	0.73

Reference cluster: Moderate, Model results only display the variable of interest but includes the complete base model. Results displayed as odds ratios

4.7 HOUSING INEQUALITY AND COVID-19

Among homeowners (defined as those who owned their home, with or without a mortgage), about 19% (929 of 4,850) reported moving during the COVID-19 period, while 81% remained in the same dwelling. In contrast, tenants (defined as those living in social, private, or rent-free housing) were more mobile, with 42% (983 of 2,319) reporting a move during the same period. Only a very small share in either group declined to answer the question (±1%). These figures suggest that residential mobility during the pandemic was substantially higher among tenants than among homeowners, but this pattern is in line with the general increased ability for tenants to move, compared to those who own their own home.

Table 20: Summary table comparing proportions of owners and tenants who moved during Covid-19

	Moved during Covid-19			
Tenure Group	No	Yes	No Answer	Total
Owners	3,904 (80%)	929 (19%)	17 (<1%)	4,850
Tenants	1,324 (57%)	983 (42%)	12 (1%)	2,319

Given the small group of participants who did not answer the question, we performed our multinomial regression without them. Of the 2307 tenants, we found no relationship between moving during the COVID-19 pandemic and cluster membership, with no odds ratio being statistically significant. We see the same pattern when looking at owners (n=4833), with moving during covid-19 not being associated with cluster membership. As before, these results control for country, family type, dwelling location, etc.

Table 21: Chi-squared test comparing having moved across the four housing clusters

Category	Moderate	Affluent	Constrained	Precarious
Owners and tenants: Moved during COVID-19 - No	2594 (49.6%)	1108 (21.2%)	1005 (19.2%)	521 (10.0%)
Owners and tenants: Moved during COVID-19 - Yes	949 (49.6%)	290 (15.2%)	491 (25.7%)	182 (9.5%)
Pearson's Chi-squared Test: $X^2(3) = 54.10$, p-value = <0.001.				

Finally, we combined tenants and owners. The crosstabulation shows that most households who did not move belong to the Moderate (49.6%) or Affluent (21.2%) clusters, while the Constrained (19.2%) and Precarious (10.0%) groups make up smaller shares. Among those who did move, however, the share of Constrained households rises to 25.7%, while the Affluent group falls to 15.2%, suggesting that mobility during the pandemic was more common among less secure households. The difference is statistically significant ($\chi^2(3) = 54.10$, $p < 0.001$).

Table 22: Moving during the COVID-19 pandemic, renters, owners separately and renters and owners combined

	Affluent (OR)	Constrained (OR)	Precarious (OR)
Tenants: moved during COVID-19	0.76	1.08	0.79
Owners: moved during COVID-19	0.98	1.14	0.86
Owners and tenants: Moved during COVID-19	0.93	1.12	0.82

However, our regression results confirm that having moved during COVID-19 does not significantly alter the odds of belonging to any cluster once other socio-economic factors are controlled for, as the odds ratios remain close to one and not reaching level of statistical significance (Affluent OR = 0.93; Constrained OR = 1.12; Precarious OR = 0.82). Renters—both private and social—show substantially higher odds of being in the Constrained (OR = 1.68 and 1.90, respectively) and Precarious (OR = 2.20 and 2.16, respectively) clusters, while homeownership is strongly associated with the Affluent cluster.

Overall, the results suggest that residential mobility during COVID-19 was concentrated among tenants, who tend to face more insecure housing conditions and are more likely to belong to the Constrained or Precarious clusters. Moves among owners appear less frequent and are not statistically linked to cluster membership once tenure type is considered.

4.8 IMPACT ON HEALTH AND WELLBEING

4.8.1 SUBJECTIVE GENERAL HEALTH

We questioned participants on their subjective health, rated from 1 ("Very Good") to 5 ("Very Bad"). Within the representative sample, only 76 respondents selected "5 – Very Bad," while 11 declined to answer. To ensure sufficient group size, categories "4 – Bad" and "5 – Very Bad" were combined into a single "4 – Bad/Very Bad" category. Those who did not answer were excluded from this specific analysis, resulting in a final sample of n = 7,158.

Table 23: Multinomial regression evaluating the general health

Predictor	Affluent	Constrained	Precarious
(Intercept)	0.16**	0.36**	0.16**
Male[1]	1.02	0.87	0.72**
36–49[2]	0.99	0.94	1.08
50–65[2]	1.23	0.67**	0.57**
65+[2]	1.35	0.43**	0.35**
Employment: Retired[3]	1.11	0.85	0.97
Employment: Other/Non-Active/Temp. Leave[3]	0.93	0.93	0.99
Education: Secondary[4]	1.01	1.07	1.06
Education: Higher Non-University[4]	0.99	0.91	0.89
Education: Higher University[4]	1.16	0.98	0.91
Living Area: Rural[5]	2.37**	0.74*	1.34
Living Area: Town/Suburb[5]	1.62**	0.87	1.38*
Household Type: Single Parent[6]	1.30	1.31	1.46
Household Type: Single[6]	2.70**	0.88	0.71*
Household Type: Other[6]	1.66**	0.86	0.92
Ownership: Owned, With Mortgage[7]	1.05	1.45**	1.10
Ownership: Rented: Social[7]	0.26**	1.82**	1.78**
Ownership: Rented: Private[7]	0.31**	1.68**	2.00**
Ownership: Free Residence[7]	0.55	1.51	1.56

Income (1000s €)	1.69**	0.92	0.91
Country: Czechia[8]	0.24**	1.57	1.34
Country: Germany[8]	0.28**	0.71	0.56
Country: Estonia[8]	0.17**	1.58*	0.66
Country: Finland[8]	0.14**	0.86	0.14**
Country: United Kingdom[8]	0.26**	1.22	1.67
Country: Italy[8]	0.48**	1.41	1.93*
Country: Romania[8]	0.27**	2.19**	2.33**
General Health: Good[9]	0.97	1.32*	1.31
General Health: Fair	0.99	1.89**	1.98**
General Health: Bad/Very Bad	1.26	2.50**	3.84**

Reference cluster: Moderate
Reference Categories: [1]Gender: Female; [2]Age: 18-35; [3]Employment: Working; [4]Education: Primary/None; [5]Living Area: City Centre; [6]Household Type: Couple with Children; [7]Ownership: Owned, No Mortgage; [8]Country: Belgium; [9]General Health: Very Good:
Notes: Results displayed as odds ratios. Nagelkerke Pseudo R^2 = 0.310; Residual Deviance = 15,096.61; AIC = 15,276.61; * $p \le 0.01$; ** $p \le 0.001$. N = 7158

Subjective general health is strongly and significantly associated with housing cluster membership. Compared with respondents in very good health, those reporting poorer health are much more likely to belong to the Constrained and Precarious clusters. Respondents with fair health are 1.89 times more likely to be in the Constrained cluster (OR = 1.89, +89%, p ≤ 0.001) and 1.98 times more likely to be in the Precarious cluster (OR = 1.98, +98%, p ≤ 0.001) than in the Moderate cluster. Those with bad or very bad health show an even stronger association, being 2.50 times more likely to fall within the Constrained cluster (OR = 2.50, +150%, p ≤ 0.001) and 3.84 times more likely to be in the Precarious cluster (OR = 3.84, +284%, p ≤ 0.001) than in the Moderate cluster.

By contrast, we find that subjective general health is not predictive of membership of the Affluent cluster, compared to the Moderate cluster. Relative to respondents in very good health, those with fair health are 1.89 times more likely to be in the Constrained cluster than in the Moderate cluster.

By contrast, general health does not significantly predict membership of the Affluent cluster, where the odds remain close to one across all health categories (OR = 0.97, −3%; OR = 0.99, −1%; OR = 1.26, +26%; all ns). These results indicate that poorer self-reported health is concentrated among individuals in housing or financial difficulty, whereas health differences are minimal among those in more secure or affluent living conditions.

4.8.2 SUBJECTIVE WELLBEING (NOW AND IN FIVE YEARS)

We included two additional variables: wellbeing as judged at the time of the survey and expected wellbeing five years from now ("On which step of the ladder would you say you personally feel you stand at this time?" and "On which step do you think you will stand about five years from now?"). These two items were derived from the Cantril Ladder (Levin & Currie, 2014), a widely used measure of subjective well-being, specifically in the Gallup World Poll and reported in the World Happiness Report.

Table 24: Multinomial regression evaluating the wellbeing at this time

Predictor	Affluent	Constrained	Precarious
(Intercept)	0.23**	1.53	1.11
Male[1]	1.02	0.84*	0.70**
36–49[2]	0.98	1.00	1.16
50–65[2]	1.23	0.78*	0.70*
65+[2]	1.35	0.50**	0.41**
Employment: Retired[3]	1.13	0.90	1.09
Employment: Other/Non-Active/Temp. Leave[3]	0.94	0.94	1.03
Education: Secondary[4]	1.00	1.08	1.06
Education: Higher Non-University[4]	0.98	0.93	0.90
Education: Higher University[4]	1.15	0.98	0.92
Living Area: Rural[5]	2.35**	0.74*	1.32
Living Area: Town/Suburb[5]	1.62**	0.87	1.37*
Household Type: Single Parent[6]	1.29	1.25	1.38
Household Type: Single[6]	2.66**	0.85	0.68*
Household Type: Other[6]	1.65**	0.85	0.91
Ownership: Owned, With Mortgage[7]	1.04	1.44**	1.10
Ownership: Rented: Social[7]	0.26**	1.79**	1.82**
Ownership: Rented: Private[7]	0.30**	1.60**	1.88**
Ownership: Free Residence[7]	0.55	1.53	1.58
Income (1000s €)	1.70**	0.93	0.93
Country: Czechia[8]	0.24**	1.52	1.33
Country: Germany[8]	0.27**	0.66	0.51*
Country: Estonia[8]	0.16**	1.49	0.62
Country: Finland[8]	0.14**	0.93	0.16**
Country: United Kingdom[8]	0.26**	1.14	1.55
Country: Italy[8]	0.47**	1.38	1.91*
Country: Romania[8]	0.28**	2.35**	2.68**
Wellbeing at this time	0.96	0.86**	0.80**

Reference cluster: Moderate
Reference Categories: [1]Gender: Female; [2]Age: 18-35; [3]Employment: Working; [4]Education: Primary/None; [5]Living Area: City Centre; [6]Household Type: Couple with Children; [7]Ownership: Owned, No Mortgage; [8]Country: Belgium.
Notes: Results displayed as odds ratios. Nagelkerke Pseudo R^2 = 0.320; Residual Deviance = 15,072.99; AIC = 15,240.99; * $p \leq 0.01$; ** $p \leq 0.001$. N = 7,169

As with subjective general health, our results point towards a relationship between wellbeing now and in the future and cluster membership. The odds of being in the

Constrained cluster decrease by 14% for each one-point increase in wellbeing (OR = 0.86, p < 0.001), while for the Precarious cluster, the decrease is 20% (OR = 0.80, p < 0.001). Subjective wellbeing five years from now follows a similar but slightly weaker pattern: with each one-point increase in expected wellbeing, the odds of belonging to the Constrained cluster decrease by 11% (OR = 0.89, p < 0.001) and by 13% for the Precarious cluster (OR = 0.87, p < 0.001) compared to the Moderate reference group. The Affluent cluster again shows no statistically significant difference from the Moderate group (OR = 0.96 in both models). The overall results show that lower rates of wellbeing, both currently and five years from now is associated with membership of the Precarious or Constrained clusters. Moderate and Affluent clusters do not differ in this respect.

4.9 ACCESS TO GREEN SPACE

Our final area of interest is access to green space. As part of the survey, six questions were asked regarding access to different forms of green space (for example, a shared garden, or wooded area close). Participants could select multiple options.

Table 25: Access to green spaces per housing cluster

	Affluent	Constrained	Precarious
Green space: Private	3.10**	0.61**	1.13
Green space: Shared garden	0.78	1.28*	1.48**
Green space: Private balcony, patio or roof terrace	1.02	0.75**	0.73*
Green space: Shared roof terrace	1.35	1.13	1.36
Green space: Park, wood or wood	0.9	0.92	1.11
Green Space: None	0.85	1.48*	0.78

Reference cluster: Moderate; Model results only display the variable of interest but includes the complete base model. Results displayed as odds ratios. * p ≤ 0.01; ** p ≤ 0.001. N = 7,169

First, individuals with a private garden were much more likely (210%) to belong to the Affluent cluster (OR = 3.10, p ≤ 0.001) and less likely to be in the Constrained group (OR = 0.61, p ≤ 0.001), while there is no significant difference for the Precarious cluster (OR = 1.13). This again highlights that the Constrained group experiences more space shortage compared to the Precarious group, even after accounting for differences in residential location.

Access to a shared garden shows the reverse pattern, being more common among the Constrained (OR = 1.28, p ≤ 0.01) and Precarious (OR = 1.48, p ≤ 0.001) clusters, but slightly less common among the Affluent (OR = 0.78) compared to the Moderate group, however the differences between Moderate and Affluent clusters are not statistically significant.

Having a private balcony, patio, or roof terrace does not differ for the Affluent cluster (OR = 1.02) but is less likely among the Constrained (OR = 0.75, p ≤ 0.001) and Precarious (OR = 0.73, p ≤ 0.01) households when compared to the Moderate cluster. There are no significant differences for shared roof terraces or access to nearby parks or wooded areas, as odds remain close to one across all groups. However, the overall incidence of shared roof terrace access is very low (ranging from 1% to 3% across clusters), which limits the interpretability of this finding. Respondents reporting no access to any green space are more likely to belong to the Constrained cluster (OR = 1.48, $p \leq 0.01$) and slightly less likely to be in the Affluent group (OR = 0.85), with no clear difference for the Precarious cluster (OR = 0.78).

In summary, access to green space differs clearly between the housing clusters, even when controlling for location, i.e., whether households live in urban, suburban, or rural areas. Affluent households are much more likely to have private gardens, while those in the Constrained and Precarious clusters more often depend on shared or smaller outdoor areas. The Constrained cluster stands out for its limited access to both private outdoor space and balconies, as well as a higher likelihood of having no green space at all, which also corresponds to their smallest average dwelling size and associated m2 availability per person. The Precarious cluster, by contrast, reports somewhat better access to shared gardens. Access to parks or wooded areas does not vary significantly across clusters.

5 DISCUSSION AND CONCLUSION

5.1 GOAL OF D4.1

The general goal of this deliverable is to sketch a better picture of the relationship between sociodemographic variables and related contextual variables and how they relate to four housing related clusters. By linking these clusters to other variables of interest, the results give an idea of how housing conditions reflect and reinforce wider (social) inequalities. Beyond this, this deliverable documents our sample in some detail, while also describing the data cleaning, stratification and imputation procedures applied in D4.1 and T4.1.

The housing clusters themselves were based on 1) a count of dwelling problems (i.e.: presence of mould), 2) subjective level of EPC, 3) surface area divided by OECD-equivalent household size and 4) a subjective assessment of space shortage, resulting in four housing clusters, Moderate, Affluent, Constrained and the Precarious cluster. The Moderate group represents average households with stable situations, medium sized dwellings and few problems. The Affluent group consists of older and financially secure households (see Figure 13) living in large and comfortable homes (see Figure 11 and Figure 10), often as outright owners in suburban or town areas. The Constrained group includes households in small dwellings with space shortage, lower income and a strong presence in cities, often with children. The Precarious group faces the most severe strain, with many dwelling problems, very low subjective EPC scores and the lowest income, combined with poorer housing quality and less secure work. In all cases, we used the Moderate cluster as reference cluster.

5.2 EFFECTS OF SOCIODEMOGRAPHIC VARIABLES

Our analysis was characterised by initially performing a multinomial regression analysis containing a set of sociodemographic variables. This aims to establish a set of base predictors for cluster membership. Following this, individual numerical or categorical variables were added to the model which allows isolation of the effects under investigation, ranging from general feelings of wellbeing to car ownership. From our initial sociodemographic variables, we identify three predictors that appear to have a consistent impact: income, ownership status and country of residence.

First, each additional one thousand euro in income (OECD equivalised) is linked with a 68% increase in the odds of belonging to the Affluent cluster (OR 1.68), compared to our reference cluster, Moderate. For the Constrain cluster, income is linked with a 12% decrease in the odds (OR 0.88), while we find a similar impact for the Precarious cluster at about 13% (OR 0.87). An odds ratio of 0.88 means that the odds are multiplied by 0.88 for each additional one thousand euro of income. In percentage terms, this corresponds to a reduction of 12%, calculated as 1 – 0.88 = 0.12. In other words, for each additional one thousand euro of income, the likelihood of belonging to the Constrained cluster decreases by about twelve percent relative to the Moderate cluster.

Ownership also shows strong predictive power. Compared with outright owners, social renting is linked with a 74% decrease in the odds of being in the Affluent group (OR 0.26) and a near doubling of the odds of being Constrained (OR 1.96) and Precarious (OR 2.08). Private renting follows the same pattern. It is linked with a 69% decrease in the odds of being Affluent (OR 0.31) and increases of 72% and 107% for the Constrained and Precarious groups (OR 1.72 and OR 2.07). By contrast, owning with a mortgage shows a 47% increase in the odds of being Constrained (OR 1.47) but no clear differences for the Affluent or Precarious groups. Free residence points in the same direction as renting, with higher odds for the Constrained and Precarious groups, although these estimates are less precise.

Finally, we see pronounced differences within the 8 countries. Compared with Belgium, our reference country, households in Czechia, Germany, Estonia and Finland show lower odds of being in the Affluent cluster, while Romania, Italy, United Kingdom and Estonia show higher odds of falling into the constrained or precarious clusters. Romania has the strongest pattern, with more than twice the odds of being in either the constrained or precarious group, compared to reference country Belgium. Finland stands out with much lower odds of being in the precarious cluster, which speaks to the generally higher quality of housing, as also seen in our descriptive data on housing quality (see Table 9).

It is worth nothing that, while there are some exceptions, generally, the strength of the demographic variables remains broadly similar in the analysis that follows. To illustrate, men do not differ from women in the odds of belonging to the Affluent or Constrained clusters, but they have lower odds of being in the Precarious cluster (OR 0.71). When examining the model that include wellbeing, the results remain the same: men do not differ from women in the odds of being in the Affluent cluster.

5.3 ADDITIONAL VARIABLES OF INTEREST

Zooming in on the broader relationships, we observe a strong and consistent association between both financial and thermal energy poverty and membership of the Constrained and Precarious clusters. Households experiencing financial energy poverty are substantially more likely to belong to these two groups, and a similar pattern is observed for thermal energy poverty. Respondents who report difficulties in keeping their homes adequately warm or cool are far more likely to fall into the Constrained or Precarious clusters, with the strongest association found among those unable to maintain either sufficient heating or cooling.

These findings indicate that energy poverty and wider forms of housing disadvantage frequently overlap, based on the definition of energy poverty applied in this study (see Section 4.5.1). However, alternative conceptualisations may lead to different conclusions. For example, the energy poverty triangle emphasises the interaction between household resources, energy prices and energy systems, and housing characteristics (Boardman, 2013). Likewise, financial energy poverty measures such as the Ten Percent Rule (Boardman, 1991) may yield different estimates and patterns.

Patterns of mobility during the COVID-19 period offer a different perspective. Tenants moved far more often than owners, but these moves do not explain differences between housing clusters once other social and economic characteristics are considered. While the descriptives suggest that housing mobility was more common among households in less secure living conditions, the regression analyses show that moving during the pandemic does not independently predict cluster membership. Instead, the effect of mobility appears to be tied to underlying differences between owners and tenants. Tenants, whether in private or social housing, show much higher odds of belonging to the Constrained or Precarious clusters, while ownership remains strongly linked to the Affluent cluster.

Of particular interest is that subjective health and wellbeing show some of the strongest associations with housing inequality. Respondents with poorer general health are much more likely to be in the Constrained or Precarious clusters than in the Moderate group, and this pattern strengthens as self-reported health worsens. The Affluent cluster does not differ from the Moderate group in this regard, suggesting that positive health outcomes are distributed similarly among households with stable or comfortable living conditions, while poorer health is concentrated among those facing housing strain. A similar pattern appears for current and expected wellbeing. Higher wellbeing is associated with a lower likelihood of belonging to the Constrained or Precarious clusters, both now and in the near future. These results indicate a close link between housing disadvantage, reduced expectations of future wellbeing and current health status.

The analysis of transport access and behaviour reveals smaller differences across clusters. Car ownership does not meaningfully distinguish most groups, apart from the Constrained cluster, where each additional car is linked to a lower likelihood of membership. Public transport use also varies only slightly. Affluent households are less likely to report using public transport, while Constrained and Precarious households show somewhat higher use of buses and trains, although many of these differences are small once income and residential location are controlled for. Access to public transport services shows similar patterns, with most modes not differing across clusters. The only clear finding is that tram or metro access is less common among the Affluent cluster, which aligns with their concentration in suburban or semi-rural areas rather than dense urban centres.

Access to green space shows clearer and more systematic differences. Private gardens are far more common among Affluent households, while the Constrained cluster has notably lower access to both gardens and balconies. Members of this group are also more likely to have no green space nearby, even after accounting for where they live. The Precarious cluster has slightly better access to shared outdoor space than the Constrained group, which reflects differences in their dwelling types. Access to parks or wooded areas does not differ across clusters, but this type of green space is more dependent on neighbourhood characteristics than on the dwelling itself. Overall, the results show that the Constrained cluster faces the most consistent form of space shortage, both indoors and outdoors.

Taken together, these findings show that the four housing clusters are meaningfully linked to broader social and material inequalities. The Moderate cluster reflects stable and relatively secure living conditions. The Affluent cluster shows clear advantages in housing quality, ownership, outdoor space and low exposure to energy poverty. By contrast, the Constrained and Precarious clusters face multiple disadvantages that span energy insecurity, poorer health, reduced wellbeing, limited outdoor space and a higher reliance on rental housing. The Constrained group is marked by a combination of smaller dwellings, space shortages and moderate-income limitations, while the Precarious group experiences the lowest incomes and the poorest housing quality, together with the highest levels of energy poverty.

The results demonstrate that housing conditions cannot be understood in isolation from broader social factors. Energy poverty, wellbeing, green space access and tenure all intersect with housing quality. Policies aimed at improving housing conditions for vulnerable groups will need to account for these overlaps, as addressing one aspect of disadvantage is unlikely to resolve the others. The clusters provide a structured way to understand these patterns and highlight the groups most likely to face persistent housing strain.

Because the analysis is based on cross-sectional survey data, the relationships described here should be understood as associations rather than causal effects. The data capture a single point in time, and households' living conditions, health, income, or wellbeing may all influence one another in ways that cannot be disentangled with this design. It is therefore not possible to determine whether poorer housing leads to worse wellbeing or whether households with lower wellbeing are more likely to end up in disadvantaged housing situations. The same applies to energy poverty, mobility, and access to green space: the observed links show how these factors cluster together across households, but they do not establish the direction of influence. The results should thus be read as a structured description of patterns within the sample and the wider inequalities they reflect, rather than as evidence of underlying causal processes.

Nonetheless, D4.1 provides a detailed view of how different aspects of living conditions group together within the population and offers a comparative lens for understanding housing inequality across countries. By combining objective indicators, subjective evaluations, and sociodemographic characteristics, the cluster-based framework distinguishes between households experiencing different combinations of housing quality, energy poverty, mobility constraints, and wellbeing outcomes. This makes it easier to compare patterns across domains that are often examined separately, such as energy affordability, wellbeing, and access to outdoor space, and to identify which household profiles face multiple and overlapping forms of strain.

The cluster approach also provides a practical basis for policy discussion and coordination at EU level. By moving beyond single indicators, it allows for more differentiated targeting of policy measures and for assessing whether existing instruments align with the profiles of households facing the most persistent forms of housing strain. The analysis further shows that housing-related inequalities are not static but can be expressed differently under changing circumstances. During the COVID-19 period, for example, residential mobility was more common among tenants, who are more likely to belong to the Constrained or Precarious clusters. This underlines the importance of recurring monitoring rather than one-off assessments.

When applied over time, the approach can support EU-level housing and affordability policies by tracking changes in the distribution of housing disadvantage, assessing whether policy instruments continue to reach households with higher levels of housing inequality, and highlighting where coordination gaps persist across housing, energy, transport, and social policy domains. Even without establishing cause and effect, the findings provide a structured and policy-relevant foundation for identifying priorities, monitoring inequalities, and informing future research using longitudinal or experimental designs.

6 REFERENCES

Andridge, R. R., & Little, R. J. A. (2010). A Review of Hot Deck Imputation for Survey Non-response. *International Statistical Review*, *78*(1), 40–64. https://doi.org/10.1111/j.1751-5823.2010.00103.x

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*, *20*(1), 40–49. https://doi.org/10.1002/mpr.329

Boardman, B. (1991). *Fuel poverty: From cold homes to affordable warmth*. (Original work published Belhaven Press)

Boardman, B. (2013). *Fixing Fuel Poverty* (0 edn). Routledge. https://doi.org/10.4324/9781849774482

Bouzarovski, S., & Petrova, S. (2015). A global perspective on domestic energy deprivation: Overcoming the energy poverty–fuel poverty binary. *Energy Research & Social Science*, *10*, 31–40. https://doi.org/10.1016/j.erss.2015.06.007

Bouzarovski, S., Thomson, H., & Cornelis, M. (2021). Confronting Energy Poverty in Europe: A Research and Policy Agenda. *Energies*, *14*(4), 858. https://doi.org/10.3390/en14040858

Brown, M. A., Soni, A., Lapsa, M. V., Southworth, K., & Cox, M. (2020). High energy burden and low-income energy affordability: Conclusions from a literature review. *Progress in Energy*, *2*(4), 042003. https://doi.org/10.1088/2516-1083/abb954

Buuren, S. van, & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

Della Valle, N., Maduta, C., D'Agostino, D., & Koukoufikis, G. (2025). Unpacking energy vulnerability in the European Union: Linking thermal discomfort with adaptive capacity. *Energy Research & Social Science*, *129*, 104376. https://doi.org/10.1016/j.erss.2025.104376

Donders, A. R. T., Van Der Heijden, G. J. M. G., Stijnen, T., & Moons, K. G. M. (2006). Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, *59*(10), 1087–1091. https://doi.org/10.1016/j.jclinepi.2006.01.014

Eremenco, S. L., Cella, D., & Arnold, B. J. (2005). A Comprehensive Method for the Translation and Cross-Cultural Validation of Health Status Questionnaires. *Evaluation & the Health Professions*, *28*(2), 212–232. https://doi.org/10.1177/0163278705275342

Fell, M. J. (2017). Energy services: A conceptual review. *Energy Research & Social Science*, *27*, 129–140. https://doi.org/10.1016/j.erss.2017.02.010

Groves, R. M. (2011). *Survey Methodology* (F. J. Fowler, M. Couper, J. M. Lepkowski, E. Singer, & R. Tourangeau, Eds; 2nd ed (Online-Ausg.)). Wiley.

Hills, Jo. (2012). *Getting the measure of fuel poverty: Final report of the Fuel Poverty Review* (CASEreport No. 72). Centre for Analysis of Social Exclusion, London School of Economics and Political Science. http://eprints.lse.ac.uk/43153

Jiglau, G., Hesselman, M., Dobbins, A., Grossmann, K., Guyet, R., Tirado Herrero, S., & Varo, A. (2024). Energy and the social contract: From "energy consumers" to "people with a right to energy". *Sustainable Development*, *32*(1), 1321–1336. https://doi.org/10.1002/sd.2727

Krugmann, H., & Goldemberg, J. (1983). The energy cost of satisfying basic human needs. *Technological Forecasting and Social Change*, *24*(1), 45–60. https://doi.org/10.1016/0040-1625(83)90062-8

Levin, K. A., & Currie, C. (2014). Reliability and Validity of an Adapted Version of the Cantril Ladder for Use with Adolescent Samples. *Social Indicators Research*, *119*(2), 1047–1063. https://doi.org/10.1007/s11205-013-0507-4

Li, J., Guo, S., Ma, R., He, J., Zhang, X., Rui, D., Ding, Y., Li, Y., Jian, L., Cheng, J., & Guo, H. (2024). Comparison of the effects of imputation methods for missing data in predictive modelling of cohort study datasets. *BMC Medical Research Methodology*, *24*(1), 41. https://doi.org/10.1186/s12874-024-02173-x

Martiskainen, M., Sovacool, B. K., Lacey-Barnacle, M., Hopkins, D., Jenkins, K. E. H., Simcock, N., Mattioli, G., & Bouzarovski, S. (2021). New Dimensions of Vulnerability to Energy and Transport Poverty. *Joule*, *5*(1), 3–7. https://doi.org/10.1016/j.joule.2020.11.016

Moore, R. (2012). Definitions of fuel poverty: Implications for policy. *Energy Policy*, *49*, 19–26. https://doi.org/10.1016/j.enpol.2012.01.057

Musil, C. M., Warner, C. B., Yobas, P. K., & Jones, S. L. (2002). A Comparison of Imputation Techniques for Handling Missing Data. *Western Journal of Nursing Research*, *24*(7), 815–829. https://doi.org/10.1177/019394502762477004

R Core Team. (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Shyu, C.-W. (2024). A framework of basic human energy needs for decent quality of life at household level: Policy implications for enhancing energy justice and meeting SDG7. *Journal of Asian Public Policy*, *0*(0), 1–15. https://doi.org/10.1080/17516234.2024.2446020

Simcock, N., Frankowski, J., & Bouzarovski, S. (2021). Rendered invisible: Institutional misrecognition and the reproduction of energy poverty. *Geoforum*, *124*, 1–9. https://doi.org/10.1016/j.geoforum.2021.05.005

Simcock, N., Jenkins, K. E. H., Lacey-Barnacle, M., Martiskainen, M., Mattioli, G., & Hopkins, D. (2021). Identifying double energy vulnerability: A systematic and narrative review of groups at-risk of energy and transport poverty in the global north. *Energy Research & Social Science*, *82*, 102351. https://doi.org/10.1016/j.erss.2021.102351

Thomson, H., Bouzarovski, S., & Snell, C. (2017). Rethinking the measurement of energy poverty in Europe: A critical analysis of indicators and data. *Indoor and Built Environment*, *26*(7), 879–901. https://doi.org/10.1177/1420326X17699260

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. https://doi.org/10.1002/sim.4067

7 ANNEXES

7.1 SOURCE FOR REPRESENTATIVE DATA

7.1.1 AGE AND AGE RESCALING

Eurostat[3] provides data on the ages of people in several age bins, starting from 0 and consisting of increments of 5 years. Age data for the UK was gathered from Office for National Statistics[4]. In both cases, the data was rescaled to account for the fact that our sample does not contain people in the range 0-18. The result thus considers the lack of minors and contains rescaled proportion.

Table 26: Proportion of people in every age bin (Eurostat)

Age bin	BE	CZ	DE	EE	FI	IT	RO	UK
0-4	4.8	5	4.5	4.8	4.2	3.4	4.8	NA
5-9	5.2	5.4	4.8	5.2	5	4.2	4.5	NA
10-14	5.7	5.5	4.6	5.7	5.7	4.7	5.6	NA
15-19	5.7	5.4	4.7	5.6	5.6	5	5.7	NA
20-24	5.5	4.7	5.3	4.8	5.5	5	5.2	NA
25-29	5.9	5	5.9	5.1	6.1	5.1	5.1	NA
30-34	6.4	6.5	6.4	7	6.8	5.4	6.9	NA
35-39	6.3	6.8	6.7	7.9	6.6	5.7	7	NA
40-44	6.5	7.3	6.4	7.2	6.5	6.2	7	NA
45-49	6.3	8.7	5.9	6.9	6.2	7.4	7.9	NA
50-54	6.4	7.3	6.7	6.8	5.6	8.1	7.7	NA
55-59	6.7	6.3	8.1	6.1	6.4	8.2	6.9	NA
60-64	6.8	5.6	7.6	6.4	6.4	7.3	5.7	NA
65+	21.7	20.5	22.4	20.5	23.4	24.3	20	NA

Table 27: Age data for every year

Age	BE	CZ	DE	EE	FI	IT	RO	UK
0	0.96	1.00	0.90	0.96	0.84	0.68	0.96	1.06
1	0.96	1.00	0.90	0.96	0.84	0.68	0.96	1.06

[3] Eurostat age bin data: https://ec.europa.eu/eurostat/databrowser/bookmark/b43501a7-4a8b-4aab-946e-bf7f946bbead?lang=en

[4] Age data for the UK: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland

2	0.96	1.00	0.90	0.96	0.84	0.68	0.96	1.06
3	0.96	1.00	0.90	0.96	0.84	0.68	0.96	1.06
4	0.96	1.00	0.90	0.96	0.84	0.68	0.96	1.06
5	1.04	1.08	0.96	1.04	1.00	0.84	0.90	1.16
6	1.04	1.08	0.96	1.04	1.00	0.84	0.90	1.16
7	1.04	1.08	0.96	1.04	1.00	0.84	0.90	1.16
8	1.04	1.08	0.96	1.04	1.00	0.84	0.90	1.16
9	1.04	1.08	0.96	1.04	1.00	0.84	0.90	1.16
10	1.14	1.10	0.92	1.14	1.14	0.94	1.12	1.21
11	1.14	1.10	0.92	1.14	1.14	0.94	1.12	1.21
12	1.14	1.10	0.92	1.14	1.14	0.94	1.12	1.21
13	1.14	1.10	0.92	1.14	1.14	0.94	1.12	1.21
14	1.14	1.10	0.92	1.14	1.14	0.94	1.12	1.21
15	1.15	1.08	0.94	1.12	1.12	1.00	1.14	1.15
16	1.15	1.08	0.94	1.12	1.12	1.00	1.14	1.15
17	1.15	1.08	0.94	1.12	1.12	1.00	1.14	1.15
18	1.15	1.08	0.94	1.12	1.12	1.00	1.14	1.15
19	1.15	1.08	0.94	1.12	1.12	1.00	1.14	1.15
20	1.11	0.94	1.06	0.96	1.10	1.00	1.04	1.21
21	1.11	0.94	1.06	0.96	1.10	1.00	1.04	1.21
22	1.11	0.94	1.06	0.96	1.10	1.00	1.04	1.21
23	1.11	0.94	1.06	0.96	1.10	1.00	1.04	1.21
24	1.11	0.94	1.06	0.96	1.10	1.00	1.04	1.21
25	1.17	1.00	1.18	1.02	1.22	1.02	1.02	1.29
26	1.17	1.00	1.18	1.02	1.22	1.02	1.02	1.29
27	1.17	1.00	1.18	1.02	1.22	1.02	1.02	1.29
28	1.17	1.00	1.18	1.02	1.22	1.02	1.02	1.29
29	1.17	1.00	1.18	1.02	1.22	1.02	1.02	1.29
30	1.27	1.30	1.28	1.40	1.36	1.08	1.38	1.38
31	1.27	1.30	1.28	1.40	1.36	1.08	1.38	1.38
32	1.27	1.30	1.28	1.40	1.36	1.08	1.38	1.38
33	1.27	1.30	1.28	1.40	1.36	1.08	1.38	1.38
34	1.27	1.30	1.28	1.40	1.36	1.08	1.38	1.38
35	1.27	1.36	1.34	1.58	1.32	1.14	1.40	1.34
36	1.27	1.36	1.34	1.58	1.32	1.14	1.40	1.34

37	1.27	1.36	1.34	1.58	1.32	1.14	1.40	1.34
38	1.27	1.36	1.34	1.58	1.32	1.14	1.40	1.34
39	1.27	1.36	1.34	1.58	1.32	1.14	1.40	1.34
40	1.29	1.46	1.28	1.44	1.30	1.24	1.40	1.28
41	1.29	1.46	1.28	1.44	1.30	1.24	1.40	1.28
42	1.29	1.46	1.28	1.44	1.30	1.24	1.40	1.28
43	1.29	1.46	1.28	1.44	1.30	1.24	1.40	1.28
44	1.29	1.46	1.28	1.44	1.30	1.24	1.40	1.28
45	1.25	1.74	1.18	1.38	1.24	1.48	1.58	1.21
46	1.25	1.74	1.18	1.38	1.24	1.48	1.58	1.21
47	1.25	1.74	1.18	1.38	1.24	1.48	1.58	1.21
48	1.25	1.74	1.18	1.38	1.24	1.48	1.58	1.21
49	1.25	1.74	1.18	1.38	1.24	1.48	1.58	1.21
50	1.28	1.46	1.34	1.36	1.12	1.62	1.54	1.36
51	1.28	1.46	1.34	1.36	1.12	1.62	1.54	1.36
52	1.28	1.46	1.34	1.36	1.12	1.62	1.54	1.36
53	1.28	1.46	1.34	1.36	1.12	1.62	1.54	1.36
54	1.28	1.46	1.34	1.36	1.12	1.62	1.54	1.36
55	1.35	1.26	1.62	1.22	1.28	1.64	1.38	1.37
56	1.35	1.26	1.62	1.22	1.28	1.64	1.38	1.37
57	1.35	1.26	1.62	1.22	1.28	1.64	1.38	1.37
58	1.35	1.26	1.62	1.22	1.28	1.64	1.38	1.37
59	1.35	1.26	1.62	1.22	1.28	1.64	1.38	1.37
60	1.37	1.12	1.52	1.28	1.28	1.46	1.14	1.21
61	1.37	1.12	1.52	1.28	1.28	1.46	1.14	1.21
62	1.37	1.12	1.52	1.28	1.28	1.46	1.14	1.21
63	1.37	1.12	1.52	1.28	1.28	1.46	1.14	1.21
64	1.37	1.12	1.52	1.28	1.28	1.46	1.14	1.21
65	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
66	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
67	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
68	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
69	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
70	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
71	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73

72	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
73	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
74	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
75	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
76	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
77	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
78	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
79	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
80	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
81	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
82	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
83	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
84	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
85	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
86	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
87	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
88	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
89	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73
90	0.84	0.79	0.86	0.79	0.90	0.93	0.77	0.73

7.1.2 UPPER LIMITS PER INCOME QUARTILE IN EURO

Income data was similarly gathered from Eurostat[5]. The data refers to household annual income, adjusted for family size. For the UK, Eurostat data was only available until 2018. To obtain more recent values, we adjusted the 2018 figures for inflation. Eurostat reports in euros, so cumulative eurozone inflation was applied: between 2018 and 2024, prices rose by 21.76% (meaning €100 in 2018 equalled €121.76 in 2024). For reference, the equivalent figure in GBP was 25.38% (£100 in 2018 equalled £125.38 in 2024). Since Eurostat reports in euros, we used the euro-based inflation adjustment. This provides a reasonable estimate in the absence of more recent OECD equivalised income data for the UK in the absence of detailed information on Eurostat.

Table 28: Income quartile cutoffs

Bin	BE	CZ	DE	EE	FI	UK	IT	RO
Q1	22,159	11,767	19,914	10,747	21,466	14,652	13,808	5,445
Q2	30,392	15,133	27,619	16,140	28,693	21,464	20,605	7,837
Q3	39,403	19,671	38,216	23,577	38,078	31,185	28,821	10,638

[5] Income quartile data from Eurostat: https://ec.europa.eu/eurostat/databrowser/view/ilc_di01$defaultview/default/table

7.1.3 GENDER PROPORTIONS

Finally, gender proportions data was also gathered from Eurostat[6].

Table 29:Gender shares per country

Country	Proportion Male (%)
BE	49.4
CZ	49
DE	49.3
EE	47.4
IT	48.9
RO	48.6
FI	49.5
UK	49

7.2 DETAILED DESCRIPTION OF THE COMPLETE DATASET

The pooled dataset includes 14 888 respondents across eight countries. The overall median age is 50 years, but there are notable differences across countries. Respondents in Belgium, Germany, Finland, the United Kingdom, Romania and Italy tend to be older (median ages 50–55), while those in the Czech Republic and Estonia tend to be younger (medians 44 and 46). This is reflected in the age group breakdown: Estonia and the Czech Republic have higher shares of 18–35-year-olds (31% each), whereas Belgium has relatively few in this category (17%) but a higher proportion of those 65 and older (25%).

Gender distribution is even across countries, though women make up a slight majority in all samples, with the highest share in Estonia (59%). Marked differences emerge in household income. Median equivalised income is highest in Belgium (€2,250) and the United Kingdom (€2,062), followed by Germany (€2024) and Finland (€1875). In contrast, the lowest medians are observed in Romania (€833), the Czech Republic and Estonia (€1,167 each). Income quartiles show that Romania is distinct, with 40% of its sample in the top quartile, while Finland and Czech Republic have relatively high proportions in the bottom quartile (43-44%).

6 Gender data from Eurostat: https://ec.europa.eu/eurostat/databrowser/view/demo_pjan/default/table?lang=en

Figure 16: Gender distribution and age across all countries for the full sample

Household structures also differ across countries. Couples with children are most common in Romania (46%) and Italy (42%), compared to only 21% in Finland. Single-person households are most frequent in Finland (38%), Estonia and Germany (both 28%), but much less common in Italy and Romania (14% each). Living with parents is particularly prevalent in Italy (13%), while it is very rare in Finland (2%).

Educational attainment shows clear contrasts. Bachelor's degrees are most common in Romania (38%) and the United Kingdom (30%), while Finland records the lowest proportion of respondents holding a master's degree or higher (11%). Finland stands out with a high proportion of respondents with higher non-university education (45%), while Italy and the Czech Republic have large shares with secondary education (40 and 44%). Urban–rural differences are pronounced. Romania and Estonia are highly urbanised in this dataset (79 and 63% living in cities), while Belgium has the lowest share of city residents (22%), with most respondents instead located in towns or suburbs. The United Kingdom also has a high suburban share (57% living in towns or suburbs).

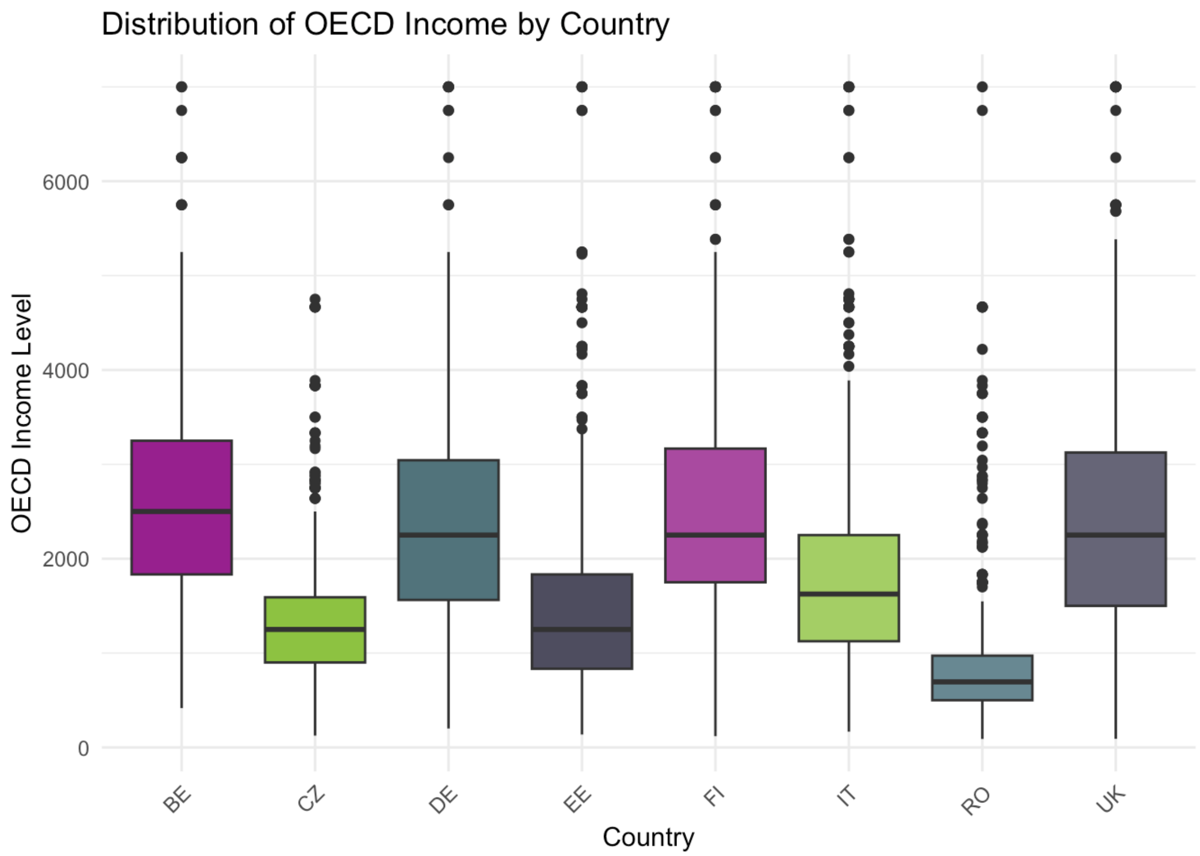

Figure 17: OECD equivalised income by country for the full sample

Cross-national differences are also evident in patterns of employment status. Full-time employment is highest in Romania (56%) and the Czech Republic (53%), but lowest in Finland, Italy and Belgium (38–39%). Retirees are most common in Belgium (33%) and Finland (26%), aligning with their older age profiles. Part-time work is more prevalent in Germany (15%), the United Kingdom and Italy (12% each) compared to Romania (3%). Unemployment is relatively high in Finland (10%) and Italy (8%), while Belgium and Germany report the lowest levels (1.6 and 2.7%).

Finally, most respondents indicated that they were born in their country of residence. However, this varies across countries. The United Kingdom has the largest share of participants indicating they were not born in their country of residence (11%), followed by Italy (6.9%). In contrast, Romania has almost no participants that were born outside their country of residence (less than 1%).

Table 30: General sample characteristics

Characteristic	Overall	BE	CZ	DE	EE	FI	IT	RO	UK
	N = 14,888[1]	N = 1,636[1]	N = 2,006[1]	N = 2,006[1]	N = 1,626[1]	N = 1,676[1]	N = 2,017[1]	N = 1,895[1]	N = 2,026[1]
Age	50 (36, 62)	55 (42, 66)	44 (32, 57)	51 (37, 64)	46 (33, 61)	50 (35, 64)	53 (41, 62)	50 (37, 58)	52 (38, 64)
Age Group									
18–35	3,494 (23%)	273 (17%)	621 (31%)	440 (22%)	505 (31%)	443 (26%)	357 (18%)	406 (21%)	449 (22%)

36–49	3,768 (25%)	366 (22%)	638 (32%)	491 (24%)	427 (26%)	367 (22%)	472 (23%)	534 (28%)	473 (23%)
50–65	4,868 (33%)	584 (36%)	523 (26%)	641 (32%)	426 (26%)	473 (28%)	835 (41%)	737 (39%)	649 (32%)
65+	2,758 (19%)	413 (25%)	224 (11%)	434 (22%)	268 (16%)	393 (23%)	353 (18%)	218 (12%)	455 (22%)
Gender									
Female	8,071 (54%)	935 (57%)	1,096 (55%)	1,039 (52%)	966 (59%)	905 (54%)	1,121 (56%)	978 (52%)	1,031 (51%)
Male	6,817 (46%)	701 (43%)	910 (45%)	967 (48%)	660 (41%)	771 (46%)	896 (44%)	917 (48%)	995 (49%)
Income (Equivalised, €)	1,500 (875, 2,250)	2,250 (1,750, 2,833)	1,167 (750, 1,500)	2,024 (1,321, 2,750)	1,167 (750, 1,750)	1,875 (1,250, 2,750)	1,310 (833, 1,875)	833 (500, 1,167)	2,062 (1,250, 2,833)
Income Quartile									
Q1	5,209 (35%)	565 (35%)	875 (44%)	697 (35%)	615 (38%)	729 (43%)	728 (36%)	342 (18%)	658 (32%)
Q2	3,548 (24%)	504 (31%)	433 (22%)	547 (27%)	385 (24%)	344 (21%)	493 (24%)	356 (19%)	486 (24%)
Q3	3,043 (20%)	317 (19%)	302 (15%)	443 (22%)	315 (19%)	316 (19%)	438 (22%)	440 (23%)	472 (23%)
Q4	3,088 (21%)	250 (15%)	396 (20%)	319 (16%)	311 (19%)	287 (17%)	358 (18%)	757 (40%)	410 (20%)
Household Type									
Couple No Children	4,228 (28%)	514 (31%)	535 (27%)	705 (35%)	343 (21%)	581 (35%)	493 (24%)	502 (26%)	555 (27%)
Couple With Children	5,077 (34%)	529 (32%)	732 (36%)	516 (26%)	533 (33%)	348 (21%)	838 (42%)	864 (46%)	717 (35%)
Living With Parents	1,081 (7.3%)	111 (6.8%)	203 (10%)	77 (3.8%)	126 (7.7%)	33 (2.0%)	272 (13%)	152 (8.0%)	107 (5.3%)
Non Family Household	225 (1.5%)	5 (0.3%)	51 (2.5%)	37 (1.8%)	35 (2.2%)	17 (1.0%)	23 (1.1%)	20 (1.1%)	37 (1.8%)
Single	3,361 (23%)	363 (22%)	334 (17%)	571 (28%)	457 (28%)	630 (38%)	274 (14%)	256 (14%)	476 (23%)
Single Parent	916 (6.2%)	114 (7.0%)	151 (7.5%)	100 (5.0%)	132 (8.1%)	67 (4.0%)	117 (5.8%)	101 (5.3%)	134 (6.6%)
Level of Education									
1 - Limited Education	1,874 (13%)	135 (8.4%)	126 (6.3%)	399 (20%)	237 (15%)	278 (17%)	499 (25%)	36 (1.9%)	164 (8.1%)
2 - Secondary	4,231 (29%)	629 (39%)	881 (44%)	307 (15%)	354 (22%)	187 (11%)	803 (40%)	647 (34%)	423 (21%)
3 - Higher Non-University	3,254 (22%)	151 (9.4%)	449 (22%)	633 (32%)	343 (21%)	760 (45%)	233 (12%)	225 (12%)	460 (23%)

4 - Higher University (Bachelors)	3,034 (20%)	454 (28%)	216 (11%)	272 (14%)	366 (23%)	258 (15%)	148 (7.3%)	714 (38%)	606 (30%)
5 - Higher University (Masters+)	2,445 (16%)	241 (15%)	329 (16%)	392 (20%)	326 (20%)	192 (11%)	333 (17%)	270 (14%)	362 (18%)
Missing	50	26	5	3	0	1	1	3	11
Living Area									
City	7,103 (48%)	364 (22%)	1,098 (55%)	952 (47%)	1,030 (63%)	860 (51%)	799 (40%)	1,494 (79%)	506 (25%)
Rural	2,783 (19%)	403 (25%)	421 (21%)	439 (22%)	272 (17%)	201 (12%)	408 (20%)	282 (15%)	357 (18%)
Town/Suburb	5,002 (34%)	869 (53%)	487 (24%)	615 (31%)	324 (20%)	615 (37%)	810 (40%)	119 (6.3%)	1,163 (57%)
Employment Status									
Employed Full Time	6,732 (45%)	639 (39%)	1,059 (53%)	877 (44%)	830 (51%)	629 (38%)	757 (38%)	1,065 (56%)	876 (43%)
Employed Part Time	1,467 (9.9%)	169 (10%)	163 (8.1%)	304 (15%)	175 (11%)	132 (7.9%)	233 (12%)	56 (3.0%)	235 (12%)
Household Work	584 (3.9%)	49 (3.0%)	49 (2.4%)	33 (1.6%)	31 (1.9%)	30 (1.8%)	221 (11%)	100 (5.3%)	71 (3.5%)
Other	103 (0.7%)	7 (0.4%)	12 (0.6%)	10 (0.5%)	14 (0.9%)	14 (0.8%)	9 (0.4%)	28 (1.5%)	9 (0.4%)
Retired	3,144 (21%)	533 (33%)	278 (14%)	471 (23%)	223 (14%)	444 (26%)	372 (18%)	369 (19%)	454 (22%)
Self Employed	794 (5.3%)	36 (2.2%)	141 (7.0%)	72 (3.6%)	88 (5.4%)	56 (3.3%)	162 (8.0%)	105 (5.5%)	134 (6.6%)
Student	644 (4.3%)	66 (4.0%)	118 (5.9%)	98 (4.9%)	79 (4.9%)	105 (6.3%)	81 (4.0%)	57 (3.0%)	40 (2.0%)
Temporary Leave	261 (1.8%)	18 (1.1%)	83 (4.1%)	40 (2.0%)	38 (2.3%)	20 (1.2%)	11 (0.5%)	33 (1.7%)	18 (0.9%)
Unable to Work	457 (3.1%)	93 (5.7%)	33 (1.6%)	47 (2.3%)	55 (3.4%)	74 (4.4%)	9 (0.4%)	14 (0.7%)	132 (6.5%)
Unemployed	702 (4.7%)	26 (1.6%)	70 (3.5%)	54 (2.7%)	93 (5.7%)	172 (10%)	162 (8.0%)	68 (3.6%)	57 (2.8%)
Born in Current Country									
No	783 (5.3%)	39 (2.4%)	121 (6.0%)	126 (6.3%)	90 (5.5%)	36 (2.1%)	140 (6.9%)	17 (0.9%)	214 (11%)
No Answer	11 (<0.1%)	1 (<0.1%)	2 (<0.1%)	2 (<0.1%)	2 (0.1%)	1 (<0.1%)	1 (<0.1%)	0 (0%)	2 (<0.1%)
Yes	14,094 (95%)	1,596 (98%)	1,883 (94%)	1,878 (94%)	1,534 (94%)	1,639 (98%)	1,876 (93%)	1,878 (99%)	1,810 (89%)

[1] Median (Q1, Q3); n (%)

7.2.1 HOUSING CHARACTERISTICS

The most common housing type in the overall sample is detached houses (28%), followed closely by large apartment blocks with ten or more units (26%). Semi-detached or terraced houses also represent a significant share (21%). Smaller apartment buildings with three to nine dwellings account for 16%, while those with only one or two dwellings constitute 8%. A very small share (less than 1%) report living in other housing types.

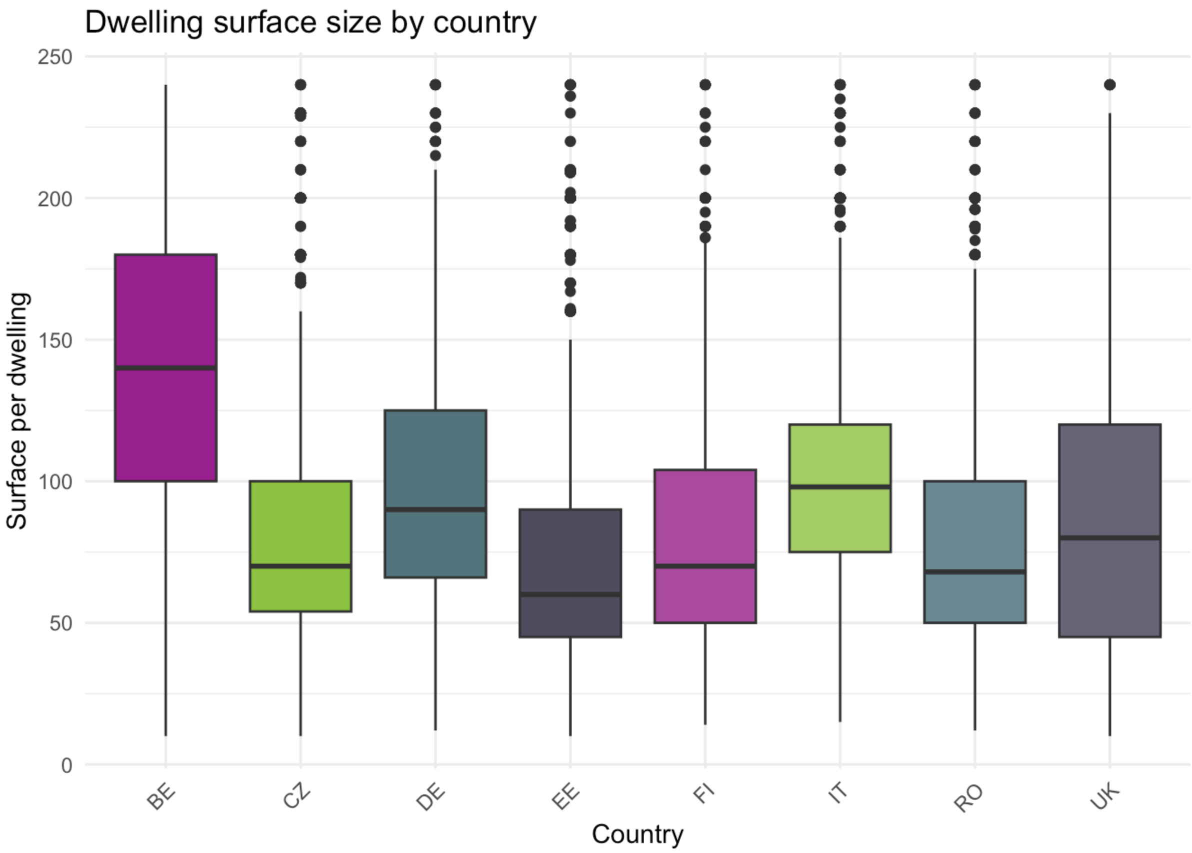

Figure 18: Dwelling size per country (full sample)

The distribution differs strongly by country. Semi-detached or terraced housing is dominant in the United Kingdom (49%) and substantial in Belgium (44%) and Finland (25%), but far less common elsewhere. Large apartment blocks with 10 or more dwellings are especially prevalent in Finland (42%), Romania (40%) and the Czech Republic (38%), while they are rare in the United Kingdom (5%) and Belgium (8%). Detached houses are more frequent in Belgium (34%), the United Kingdom (30%), and Romania (33%), while Finland stands out for its relatively low share of detached homes (19%).

Almost half of respondents report owning their home without a mortgage (44%), while 23% own with a mortgage and 22% rent privately. Social housing represents 8% of the sample, and a small minority (3%) live rent-free. Patterns vary across countries. Ownership without a mortgage is especially high in Romania (76%) and Italy (53%), while much lower in Germany (23%) and Finland (23%). Mortgaged ownership is most common in Belgium (36%) and Finland (29%). Private renting dominates in Germany (43%) and the Czech Republic (31%), contrasting with Romania (8%) and Belgium (14%). Social housing has a particularly high presence in Finland (17%) and the United Kingdom (13%) but is less common in our Southern and Eastern European sample.

The housing stock spans several generations of construction. Overall, most dwellings (42%) were built between 1971 and 2006, followed by 22% from 1941 to 1970. More recent buildings (constructed after 2007) make up 17%, while 12% predate 1940.

Country differences are clear. The United Kingdom has the highest proportion of pre-1940 housing (22%), followed by the Czech Republic (16%) and Germany (13%). Post-2007 housing is most common in Belgium (20%) and Finland (21%). Mid-century buildings (1941 – 1970) are particularly numerous in Germany (25%) and Italy (25%).

Around 28% of respondents report feeling that their dwelling is too small. Perceived shortages are most common in Romania (38%), Estonia (37%), and the Czech Republic (36%), and least common in Belgium (16%). Almost all respondents report having access to a kitchen (99%). Most dwellings have two or three rooms (46% overall). One- and two-room dwellings are most common in Romania Finland and Estonia, while Belgium and The United Kingdom stand out for their dwellings with five or more rooms. In Germany, the Czech Republic, and Italy, three- and four-room dwellings are most frequent. Reported dwelling sizes show large differences, with Belgium and Italy reporting the largest mean floor areas (141 and 102 m^2), and Estonia and Romania the smallest (75 and 80 m^2).

The mean self-reported value of dwellings across the sample is about €212,000. However, more than 9,000 cases are missing, largely because many respondents do not own their home and therefore could not provide a value. Values are highest in Belgium (€354,000) and Germany (€287,000), followed by the United Kingdom (€282,000). Romania reports the lowest average dwelling value at just under €79,000. A large share of households report that no major energy renovations have taken place (35%). At the same time, renovations are spread across different periods. About 15% of dwellings were renovated from 2020 onwards, while 10% underwent major works between 2015 and 2019, and 8% between 2010 and 2014. Earlier periods account for smaller shares. Belgium has the highest proportion reporting recent renovations from 2020 onwards (22%), while Italy and Germany have comparatively high shares of unrenovated dwellings (49% and 40% respectively).

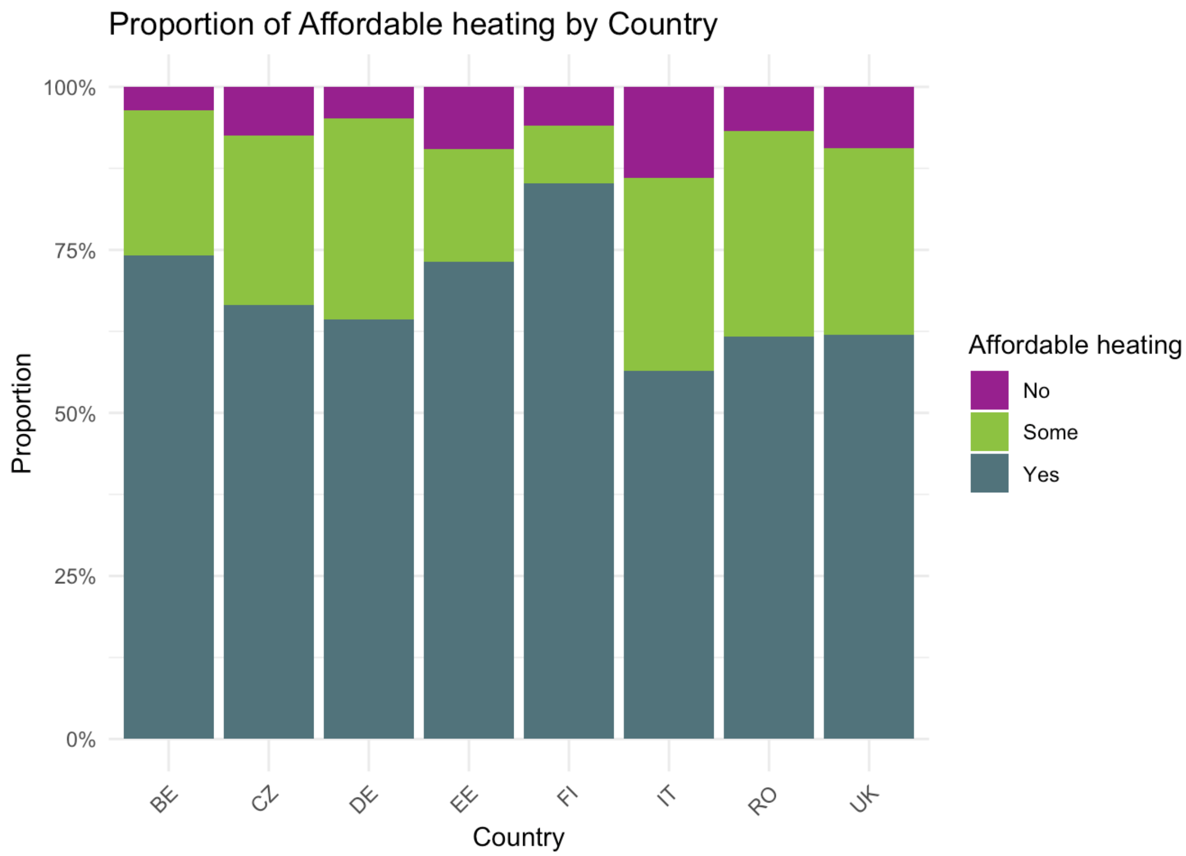

Figure 19: Proportions of people being to (in)adequately fuel their homes per country (full sample)

Most respondents (67%) report being able to heat their entire dwelling to an adequate temperature. However, a quarter indicate they can only afford to heat some rooms to an adequate temperature, and the minority of 8% say they cannot afford adequate heating at all. Heating affordability challenges are most visible in Italy (14% cannot afford heating, and 29% can only heat some rooms), the United Kingdom (9% cannot afford heating, and 29% can only heat some rooms) and Romania (7% cannot afford heating, and 31% can only heat some rooms). In contrast, Finland has the lowest reported affordability problems, with 85% stating they can heat their entire dwelling adequately.

About half of respondents (49%) say they can keep their entire dwelling comfortably cool during summer. However, 27% report they can only cool some rooms, and 24% say they cannot keep their dwelling comfortably cool at all. Estonia stands out, with a majority (51%) indicating they cannot keep their dwelling comfortably cool. Finland also shows high levels of discomfort, with 30% unable to cool their home adequately and 22% able to cool only some rooms. In contrast, the United Kingdom (63%) and Belgium (59%) have the highest shares of respondents who report being able to keep their entire dwelling comfortably cool.

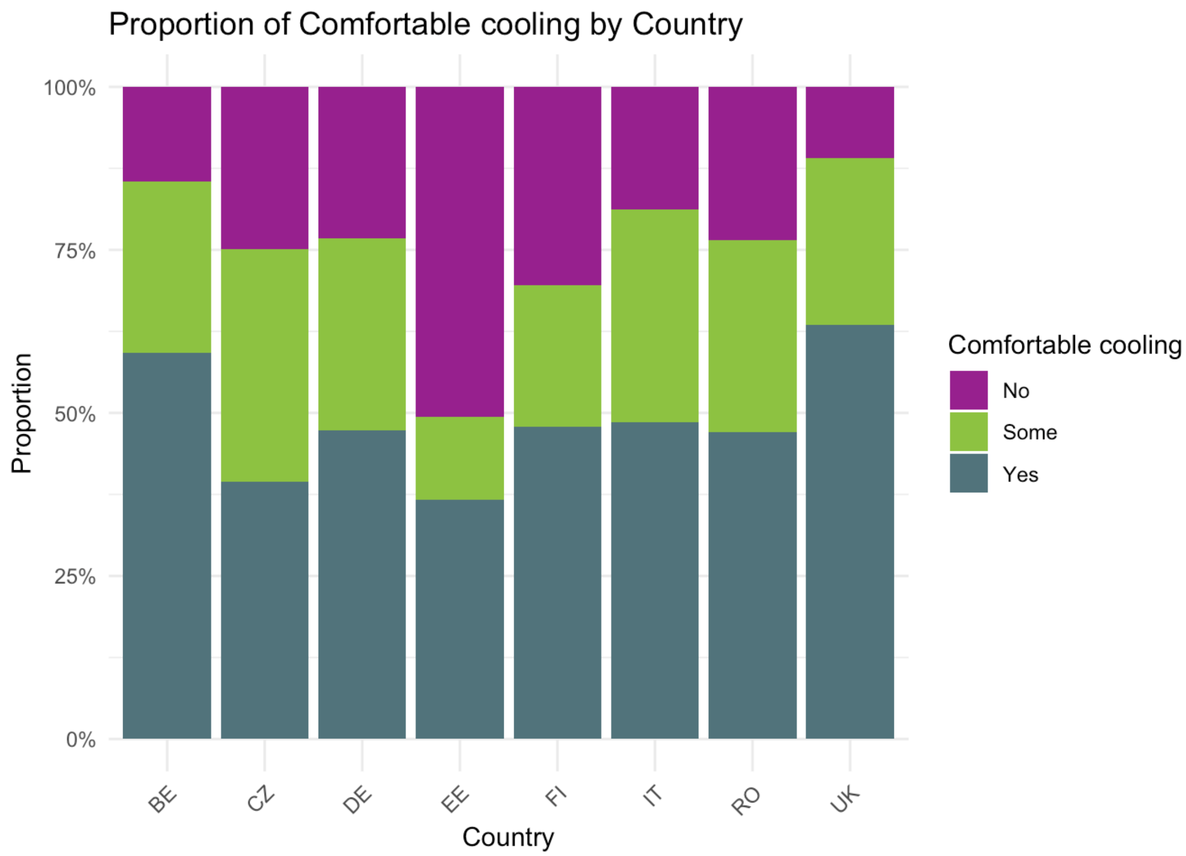

Figure 20: Proportions of people being to (in)adequately cool their homes per country (full sample)

Table 31: Housing characteristics of the whole sample

Characteristic	Overall N = 14,888[1]	BE N = 1,636[1]	CZ N = 2,006[1]	DE N = 2,006[1]	EE N = 1,626[1]	FI N = 1,676[1]	UK N = 2,026[1]	IT N = 2,017[1]	RO N = 1,895[1]
Type of Dwelling									
Detached House	4,200 (28%)	558 (34%)	566 (28%)	557 (28%)	477 (29%)	316 (19%)	613 (30%)	490 (24%)	623 (33%)
Flat: 10+ Dwellings	3,813 (26%)	137 (8.4%)	769 (38%)	357 (18%)	502 (31%)	709 (42%)	96 (4.7%)	494 (24%)	749 (40%)
Flat:1-2 Dwellings	1,187 (8.0%)	61 (3.7%)	108 (5.4%)	132 (6.6%)	283 (17%)	78 (4.7%)	152 (7.5%)	157 (7.8%)	216 (11%)
Flat:3-9 Dwellings	2,375 (16%)	154 (9.4%)	352 (18%)	636 (32%)	256 (16%)	137 (8.2%)	150 (7.4%)	510 (25%)	180 (9.5%)
Other	122 (0.8%)	3 (0.2%)	32 (1.6%)	12 (0.6%)	12 (0.7%)	13 (0.8%)	30 (1.5%)	8 (0.4%)	12 (0.6%)
Semi-detached / Terraced House	3,191 (21%)	723 (44%)	179 (8.9%)	312 (16%)	96 (5.9%)	423 (25%)	985 (49%)	358 (18%)	115 (6.1%)
Tenure status									

Free Residence	435 (2.9%)	41 (2.5%)	61 (3.0%)	23 (1.1%)	85 (5.2%)	17 (1.0%)	35 (1.7%)	99 (4.9%)	74 (3.9%)
Owned, No Mortgage	6,600 (44%)	726 (44%)	790 (39%)	467 (23%)	827 (51%)	381 (23%)	900 (44%)	1,075 (53%)	1,434 (76%)
Owned, With Mortgage	3,350 (23%)	588 (36%)	353 (18%)	416 (21%)	351 (22%)	493 (29%)	507 (25%)	454 (23%)	188 (9.9%)
Rented: Private	3,329 (22%)	223 (14%)	615 (31%)	868 (43%)	314 (19%)	506 (30%)	327 (16%)	323 (16%)	153 (8.1%)
Rented: Social	1,174 (7.9%)	58 (3.5%)	187 (9.3%)	232 (12%)	49 (3.0%)	279 (17%)	257 (13%)	66 (3.3%)	46 (2.4%)
Construction Period									
1 – Before 1940	1,725 (12%)	202 (12%)	316 (16%)	261 (13%)	177 (11%)	71 (4.2%)	436 (22%)	167 (8.3%)	95 (5.0%)
2 – 1941 - 1970	3,290 (22%)	344 (21%)	475 (24%)	504 (25%)	330 (20%)	355 (21%)	432 (21%)	505 (25%)	345 (18%)
3 – 1971 - 2006	6,259 (42%)	665 (41%)	710 (35%)	817 (41%)	743 (46%)	860 (51%)	572 (28%)	920 (46%)	972 (51%)
4 – 2007 - Later	2,543 (17%)	321 (20%)	295 (15%)	311 (16%)	261 (16%)	344 (21%)	391 (19%)	260 (13%)	360 (19%)
5 – Don't Know	1,071 (7.2%)	104 (6.4%)	210 (10%)	113 (5.6%)	115 (7.1%)	46 (2.7%)	195 (9.6%)	165 (8.2%)	123 (6.5%)
Perceived Space Shortage	4,206 (28%)	263 (16%)	727 (36%)	394 (20%)	594 (37%)	398 (24%)	559 (28%)	560 (28%)	711 (38%)
Kitchen Availability	14,782 (99%)	1,632 (100%)	1,995 (99%)	1,993 (99%)	1,600 (98%)	1,673 (100%)	2,008 (99%)	2,005 (99%)	1,876 (99%)
Surface Area (m²)	94 (50)	141 (49)	83 (43)	99 (43)	75 (46)	82 (43)	93 (58)	102 (40)	80 (43)
Total Rooms									
1	835 (5.6%)	17 (1.0%)	174 (8.7%)	65 (3.2%)	176 (11%)	220 (13%)	33 (1.6%)	36 (1.8%)	114 (6.0%)
2	2,976 (20%)	114 (7.0%)	483 (24%)	396 (20%)	444 (27%)	471 (28%)	186 (9.2%)	233 (12%)	649 (34%)
3	3,924 (26%)	291 (18%)	621 (31%)	560 (28%)	488 (30%)	384 (23%)	379 (19%)	629 (31%)	572 (30%)
4	2,824 (19%)	367 (22%)	327 (16%)	375 (19%)	261 (16%)	273 (16%)	413 (20%)	513 (25%)	295 (16%)
5	2,042 (14%)	326 (20%)	211 (11%)	283 (14%)	117 (7.2%)	209 (12%)	421 (21%)	317 (16%)	158 (8.3%)
6	1,143 (7.7%)	243 (15%)	84 (4.2%)	182 (9.1%)	71 (4.4%)	74 (4.4%)	274 (14%)	161 (8.0%)	54 (2.8%)
7	617 (4.1%)	129 (7.9%)	60 (3.0%)	87 (4.3%)	41 (2.5%)	27 (1.6%)	164 (8.1%)	80 (4.0%)	29 (1.5%)
8	378 (2.5%)	106 (6.5%)	32 (1.6%)	40 (2.0%)	24 (1.5%)	13 (0.8%)	105 (5.2%)	38 (1.9%)	20 (1.1%)

9	149 (1.0%)	43 (2.6%)	14 (0.7%)	18 (0.9%)	4 (0.2%)	5 (0.3%)	51 (2.5%)	10 (0.5%)	4 (0.2%)
Perceived dwelling Value (EUR)	211,986 (155,029)	353,928 (129,079)	198,499 (114,316)	286,822 (159,420)	140,241 (117,488)	187,719 (116,752)	282,269 (175,771)	175,442 (125,893)	78,800 (76,465)
Missing	9.354	896	1.319	1.530	974	1.067	1.195	1.223	1.150
Major Energy Renovations									
1 - No	5,226 (35%)	462 (28%)	628 (31%)	808 (40%)	517 (32%)	623 (37%)	622 (31%)	984 (49%)	582 (31%)
2 - Before 2005	968 (6.5%)	92 (5.6%)	119 (5.9%)	153 (7.6%)	97 (6.0%)	68 (4.1%)	178 (8.8%)	122 (6.0%)	139 (7.3%)
3 - 2005-2009	911 (6.1%)	104 (6.4%)	167 (8.3%)	105 (5.2%)	122 (7.5%)	58 (3.5%)	112 (5.5%)	69 (3.4%)	174 (9.2%)
4 - 2010-2014	1,139 (7.7%)	165 (10%)	190 (9.5%)	127 (6.3%)	132 (8.1%)	106 (6.3%)	142 (7.0%)	102 (5.1%)	175 (9.2%)
5 - 2015-2019	1,557 (10%)	241 (15%)	222 (11%)	148 (7.4%)	176 (11%)	148 (8.8%)	185 (9.1%)	150 (7.4%)	287 (15%)
6 - From-2020	2,163 (15%)	352 (22%)	218 (11%)	260 (13%)	261 (16%)	199 (12%)	311 (15%)	271 (13%)	291 (15%)
7 - Yes: Unknown Year	1,120 (7.5%)	100 (6.1%)	234 (12%)	91 (4.5%)	148 (9.1%)	145 (8.7%)	149 (7.4%)	102 (5.1%)	151 (8.0%)
8 – Don't Know	1,804 (12%)	120 (7.3%)	228 (11%)	314 (16%)	173 (11%)	329 (20%)	327 (16%)	217 (11%)	96 (5.1%)
Affordability of Heating									
No	1,165 (7.8%)	59 (3.6%)	151 (7.5%)	98 (4.9%)	155 (9.5%)	99 (5.9%)	191 (9.4%)	283 (14%)	129 (6.8%)
Some rooms	3,700 (25%)	363 (22%)	521 (26%)	617 (31%)	281 (17%)	148 (8.8%)	579 (29%)	595 (29%)	596 (31%)
Yes	10,023 (67%)	1,214 (74%)	1,334 (67%)	1,291 (64%)	1,190 (73%)	1,429 (85%)	1,256 (62%)	1,139 (56%)	1,170 (62%)
Ability to keep cool									
No	3,580 (24%)	237 (14%)	499 (25%)	466 (23%)	823 (51%)	509 (30%)	221 (11%)	379 (19%)	446 (24%)
Some rooms	4,042 (27%)	431 (26%)	716 (36%)	590 (29%)	207 (13%)	364 (22%)	519 (26%)	658 (33%)	557 (29%)
Yes	7,266 (49%)	968 (59%)	791 (39%)	950 (47%)	596 (37%)	803 (48%)	1,286 (63%)	980 (49%)	892 (47%)

[1] Mean (SD); n (%)

Table 32: Geographic spread of the full sample (excluding Belgium)

CZ		DE		EE		FI		UK		IT		RO	
Region	N = 2,066[1]	Region	N = 2,066[1]	Region	N = 1,626[1]	Region	N = 1,676[1]	Region	N = 2,026[1]	Region	N = 2,017[1]	Region	N = 1,895[1]
Geographic Region		Geographic Region		Geographic Region		Geographic Region		Geographic Region		Geographic Region		Geographic Region	
Hlavní město Praha	322 (16 %)	Baden-Württemberg	240 (12 %)	Harju maakond	709 (50 %)	Central Finland Region	108 (6.5 %)	England	1,659 (88 %)	Abruzzi	46 (2.3 %)	Alba	28 (1.5 %)
Jihočeský kraj	79 (4.0 %)	Bayern	333 (17 %)	Hiiu maakond	16 (1.1 %)	Central Ostrobothnia Region	13 (0.8 %)	Northern Ireland	53 (2.8 %)	Basilicata	12 (0.6 %)	Arad	33 (1.8 %)
Jihomoravský kraj	196 (10 %)	Berlin	100 (5.1 %)	Ida-Viru maakond	57 (4.0 %)	Kainuu	23 (1.4 %)	Scotland	127 (6.7 %)	Calabria	44 (2.2 %)	Argeş	46 (2.5 %)
Karlovarský kraj	51 (2.6 %)	Brandenburg	72 (3.6 %)	Järva maakond	30 (2.1 %)	Kanta-Häme	54 (3.3 %)	Wales	43 (2.3 %)	Campania	145 (7.3 %)	Bacău	64 (3.5 %)
Kraj Vysočina	91 (4.6 %)	Bremen	23 (1.2 %)	Jõgeva maakond	32 (2.3 %)	Kymenlaakso	52 (3.1 %)		144	Emilia-Romagna	150 (7.6 %)	Bihor	29 (1.6 %)
Královéhradecký kraj	79 (4.0 %)	Hamburg	40 (2.0 %)	Lääne maakond	24 (1.7 %)	Lapland	40 (2.4 %)	[1] n (%)		Friuli-Venezia Giulia	40 (2.0 %)	Bistriţa-Năsăud	10 (0.5 %)
Liberecký kraj	73 (3.7 %)	Hessen	148 (7.5 %)	Lääne-Viru maakond	55 (3.9 %)	North Karelia	49 (3.0 %)			Lazio	187 (9.4 %)	Botoşani	28 (1.5 %)
Moravskoslezský kraj	242 (12 %)	Land Berlin	1 (<0.1%)	Pärnu maakond	118 (8.4 %)	North Ostrobothnia Region	112 (6.7 %)			Liguria	72 (3.6 %)	Brăila	38 (2.1 %)
Olomoucký kraj	122 (6.2 %)	Mecklenburg-Vorpommern	36 (1.8 %)	Põlva maakond	25 (1.8 %)	Northern Savo	82 (4.9 %)			Lombardia	371 (19 %)	Braşov	81 (4.4 %)
Pardubický kraj	106 (5.4 %)	Niedersachsen	208 (11 %)	Rapla maakond	37 (2.6 %)	Ostrobothnia Region	31 (1.9 %)			Marche	41 (2.1 %)	Bucureşti	287 (16 %)
Plzeňský kraj	107 (5.5 %)	Nordrhein-Westfalen	386 (20 %)	Saare maakond	32 (2.3 %)	Päijänne Tavastia	68 (4.1 %)			Molise	7 (0.4 %)	Buzău	55 (3.0 %)
Středočeský kraj	210 (11 %)	Rheinland-Pfalz	102 (5.2 %)	Tartu maakond	187 (13 %)	Pirkanmaa	169 (10 %)			Piemonte	160 (8.1 %)	Călăraşi	10 (0.5 %)
Ústecký kraj	175 (8.9 %)	Saarland	30 (1.5 %)	Valga maakond	24 (1.7 %)	Satakunta	62 (3.7 %)			Puglia	145 (7.3 %)	Caraş-Severin	19 (1.0 %)

Zlínský kraj	107 (5.5 %)	Sachsen	106 (5.4 %)	Viljandi maakond	39 (2.8 %)	South Karelia	28 (1.7 %)			Sardegna	70 (3.5 %)	Cluj	81 (4.4 %)
		Sachsen-Anhalt	54 (2.7 %)	Võru maakond	23 (1.6 %)	South Ostrobothnia Region	44 (2.7 %)			Sicilia	144 (7.3 %)	Constanța	63 (3.4 %)
		Schleswig-Holstein	61 (3.1 %)			Southern Savonia	31 (1.9 %)			Toscana	121 (6.1 %)	Covasna	12 (0.7 %)
		Thüringen	34 (1.7 %)			Southwest Finland	165 (9.9 %)			Trentino-Alto Adige	33 (1.7 %)	Dâmbovița	28 (1.5 %)
						Uusimaa	529 (32 %)			Umbria	18 (0.9 %)	Dolj	63 (3.4 %)
										Valle D'Aosta	4 (0.2 %)	Galați	62 (3.4 %)
										Veneto	169 (8.5 %)	Giurgiu	11 (0.6 %)
												Gorj	32 (1.7 %)
												Harghita	12 (0.7 %)
												Hunedoara	64 (3.5 %)
												Ialomița	24 (1.3 %)
												Iaşi	97 (5.3 %)
												Ilfov	50 (2.7 %)
												Maramureş	32 (1.7 %)
												Mehedinți	15 (0.8 %)
												Mureş	31 (1.7 %)
												Neamț	41 (2.2 %)

							Olt	36 (2.0%)
							Prahova	59 (3.2%)
							Sălaj	20 (1.1%)
							Satu Mare	26 (1.4%)
							Sibiu	24 (1.3%)
							Suceava	57 (3.1%)
							Teleorman	13 (0.7%)
							Timiş	82 (4.5%)
							Tulcea	12 (0.7%)
							Vâlcea	25 (1.4%)
							Vaslui	16 (0.9%)
							Vrancea	23 (1.3%)
Missing	46	32	218	16	144	38		56

7.3 ADDITIONAL DATA GATHERED

We additionally gathered extra information in two specific countries: Czechia and Romania, in order to increase the share of participants in the lower income brackets. This yielded 309 (CZ) + 322 (RO) participants from these two countries and also added a German and Italian respondent, for a total of 633 participants.

		Country			
Characteristic		CZ	DE	IT	RO
		N = 309[1]	N = 1[1]	N = 1[1]	N = 322[1]
Gender	633				
Female		154 (49.8%)	1 (100.0%)	1 (100.0%)	170 (52.8%)

Male		155 (50.2%)	0 (0.0%)	0 (0.0%)	152 (47.2%)
Age	633				
18–35		50 (16.2%)	1 (100.0%)	1 (100.0%)	62 (19.3%)
36–49		92 (29.8%)	0 (0.0%)	0 (0.0%)	121 (37.6%)
50–65		110 (35.6%)	0 (0.0%)	0 (0.0%)	87 (27.0%)
65+		57 (18.4%)	0 (0.0%)	0 (0.0%)	52 (16.1%)
Income Quartile	633				
Q1		137 (44.3%)	1 (100.0%)	0 (0.0%)	73 (22.7%)
Q2		94 (30.4%)	0 (0.0%)	1 (100.0%)	72 (22.4%)
Q3		47 (15.2%)	0 (0.0%)	0 (0.0%)	88 (27.3%)
Q4		31 (10.0%)	0 (0.0%)	0 (0.0%)	89 (27.6%)

[1] n (%)

Table 33: Additional data gathered to increase the Romanian and Czech sample

7.4 BASE MODEL FOR THE COMPLETE SAMPLE

Predictor	Affluent	Constrained	Precarious
(Intercept)	0.16**	0.46**	0.22**
Male[1]	0.97	0.91	0.80**
36–49[2]	1.03	1.06	1.21
50–65[2]	1.25*	0.79**	0.87
65+[2]	1.33	0.51**	0.51**
Employment: Retired[3]	1.11	0.88	0.92
Employment: Other/Non-Active/Temp. Leave[3]	1.04	1.03	1.25*
Education: Secondary[4]	1.09	1.08	1.01
Education: Higher Non-University[4]	1.05	1.03	0.98
Education: Higher University[4]	1.20	1.05	0.90
Living Area: Rural[5]	2.32**	0.83*	1.57**
Living Area: Town/Suburb[5]	1.58**	0.90	1.35**
Household Type: Single Parent[6]	1.53**	1.17	1.33*
Household Type: Single[6]	2.81**	0.86	0.78*
Household Type: Other[6]	1.61**	0.89	0.85
Ownership: Owned, With Mortgage[7]	1.07	1.28**	1.19
Ownership: Rented: Social[7]	0.31**	1.84**	2.00**
Ownership: Rented: Private[7]	0.35**	1.73**	2.04**
Ownership: Free Residence[7]	0.76	1.28	2.12**

Income (1000s €)	1.65**	0.88**	0.81**
Country: Czechia[8]	0.22**	1.47**	1.45*
Country: Germany[8]	0.30**	0.74*	0.63*
Country: Estonia[8]	0.19**	1.74**	0.89
Country: Finland[8]	0.14**	1.00	0.15**
Country: United Kingdom[8]	0.30**	1.25	1.64**
Country: Italy[8]	0.48**	1.38*	1.73**
Country: Romania[8]	0.25**	1.88**	1.79**

Reference cluster: Moderate
Reference Categories: [1]Gender: Female; [2]Age: 18-35; [3]Employment: Working; [4]Education: Primary/None; [5]Living Area: City Centre; [6]Household Type: Couple with Children; [7]Ownership: Owned, No Mortgage; [8]Country: Belgium.
Notes: Results displayed as odds ratios. Nagelkerke Pseudo R^2 = 0.280; Residual Deviance = 32,364.21; AIC = 32,526.21; * $p \le 0.01$; ** $p \le 0.001$. N = 14888

Table 34: Base multinomial regression model for the complete sample prior to stratification