

Multidimensional housing inequality across eight European countries: a cross-national cluster and risk analysis based on an online survey

Anna Conzatti

a.conzatti@exeter.ac.uk

University of Exeter

Peter Conradie

IMEC

Sofie Naeyaert

IMEC

Anne Kantel

Fraunhofer Institute for Systems and Innovation Research

Iska Brunzmeka

Fraunhofer Institute for Systems and Innovation Research

Hanna Janßen

Fraunhofer Institute for Systems and Innovation Research

Sabine Preuß

Fraunhofer Institute for Systems and Innovation Research

Gesche Huebner

University of Exeter

Research Article

Keywords: Housing vulnerability, Green space, Transport access, Housing policy

Posted Date: October 13th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10587000/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Multidimensional housing inequality across eight European countries: a cross-national cluster and risk analysis based on an online survey

Abstract

Housing inequalities remain widespread across Europe and intersect with environmental, infrastructural, and health-related vulnerabilities. While previous research has often examined dwelling quality, green-space access, and transport access separately, less is known about how these factors co-occur within households and how dwelling vulnerability relates to inequalities in environmental access and mobility.

Using survey data from eight European countries ($n = 7,169$), we conducted a K-Prototypes cluster analysis based on dwelling, green-space, and transport indicators and subsequently characterised the resulting clusters using socio-demographic, health and wellbeing, and contextual indicators. We identified four household profiles: (i) accumulated housing and environmental vulnerability; (ii) thermal and energy vulnerability; (iii) resource-advantaged with thermal strain; and (iv) dwelling problems and transport vulnerability.

We then used a category-based relative risk analysis to examine cross-domain co-occurrence. In the pooled sample, households with dwelling vulnerability were more likely to experience green-space inequality than households without dwelling vulnerability ($RR = 1.57$; 95% CI: 1.39–1.77), with point estimates above 1 in all national contexts. By contrast, dwelling vulnerability was not consistently associated with transport inequality. The pooled estimate was below 1 ($RR = 0.86$; 95% CI: 0.81–0.92), while country-specific estimates were generally close to 1 and varied across national contexts. These findings suggest that green-space inequality may be more consistently associated with dwelling vulnerability than transport inequality, which appears to be shaped by more context-specific processes.

Keywords

Housing vulnerability – Green space – Transport access – Housing policy

1. Introduction

Inequalities in households' living conditions are widespread across Europe and have significant implications for health, wellbeing, and life chances (Filandri and Olagnero 2014; European Parliament. Directorate General for Parliamentary Research Services. 2020). Adequate housing—recognised as a fundamental human right (United Nations General Assembly 1948)—is a central component of these living conditions, but it does not act in isolation (Ayala et al. 2022; Hick et al. 2022; James et al. 2024). Housing is embedded within wider environmental, infrastructural, and socio-economic systems that shape how people experience and use their homes, including access to green space and transport services.

European research has increasingly recognised the importance of multiple indicators for capturing differences in living conditions (Guio et al. 2021). Comparative studies document persistent cross-country differences in dwelling conditions, overcrowding, and access to adequate housing, particularly between Northern and Southern Europe and between long-standing and newer EU Member States (Norris and Shiels 2007; Filandri and Olagnero 2014; Matel 2025; Mandic and Cirman 2012). The development of EU deprivation indicators—including measures of housing quality, the ability to keep the home adequately warm, and the affordability of essential goods and services—has reinforced a shift from relying primarily on income or dwelling-based measures towards assessing deprivation through observed living conditions (Eurostat, 2009).

Alongside this shift, a growing body of research has highlighted inequalities extending beyond the dwelling itself to the wider living environment. Studies of neighbourhood and environmental conditions document socially patterned access to green and blue space, exposure to environmental burdens, and unequal access to public infrastructure and services across European cities and regions (Council of Europe Development Bank 2017; Willberg et al. 2024). More recent work has broadened this perspective further by conceptualising housing inequality as intersecting with a wider range of neighbourhood and service-related dimensions, including transport accessibility, access to healthcare and education, digital connectivity, and neighbourhood quality (Holmes et al. 2025;

Bentley et al. 2025; Tayefi Nasrabadi et al. 2025, 2024). Review work has increasingly challenged purely dwelling-centred interpretations of housing inequality. James et al. (2024), for example, show that housing inequality can be understood as an outcome, an experience, a product of public policy, and a social construct, frequently intersecting with inequalities in health, wealth, and social conditions. Similarly, Born et al. (2025) argue that housing inequalities extend beyond dwelling conditions and cannot be meaningfully analysed in isolation from other dimensions of inequality.

Among these wider dimensions, green-space access and transport accessibility are particularly relevant because they represent two key ways in which households interact with their surrounding environment (López-Gay et al. 2022; Boch et al. 2020; Browning et al. 2022; Endalew Terefe and Hou 2024; Smith et al. 2017; Frank et al. 2019). Access to green space has important implications for health, wellbeing, and opportunities for recreation, while transport accessibility shapes mobility, access to services, employment opportunities, and social participation (Browning et al. 2022; Endalew Terefe and Hou 2024; Cardinali et al. 2024). Both dimensions have been shown to be socially patterned across European contexts and therefore provide useful lenses to examine how housing-related disadvantages extend beyond the dwelling itself (Faka 2025; Wolff et al. 2020; Begic et al. 2026; Mejía Dorantes and Murauskaite-Bull 2022).

This literature suggests that housing inequality should not be understood solely as a characteristic of the dwelling itself, but as a multidimensional phenomenon shaped by the interaction between housing conditions, environmental resources, transport accessibility, and wider social circumstances. However, important limitations remain in understanding how wider housing inequalities intersect at the household level. First, there are dimensional limitations. Much of the existing literature examines single dimensions of inequality—such as affordability, housing quality, or tenure—separately from environmental conditions such as green space or transport access (Norris and Shiels 2007; Hick et al. 2022). Studies that examine neighbourhood quality, green space, or infrastructure typically adopt spatial or area-based perspectives, with limited attention to how environmental disadvantages intersect with housing conditions and socio-demographic vulnerabilities at the household level (Council of Europe Development Bank 2017; Willberg et al. 2024).

Second, there are analytical limitations in how multidimensional approaches are applied. Recognising that housing intersects with income and deprivation in living conditions (Dewilde 2022), recent research increasingly adopts multidimensional or non-binary approaches. These include joint analyses of housing discrimination and deprivation (Grotti et al. 2024), fuzzy-set measures of housing deprivation (Ayala et al. 2022), component-based analyses of overcrowding and housing conditions (Dudek and Wojewódzka-Wiewiórska 2024), and studies linking housing and environmental problems to health outcomes (Carmona-Derqui et al. 2023). Comparative and spatial research has also applied clustering approaches at the country level (Czerniak and Graca-Gelert 2024), within urban contexts (Walter et al. 2025), and through analyses of neighbourhood trajectories and policy change (Nelson et al. 2024). However, these studies often remain dwelling-centred, focus on specific dimensions, or are limited to particular contexts, such as individual cities, neighbourhoods, or national case studies, providing only partial insight into how multiple disadvantages co-occur and accumulate within households across the wider living environment.

This study addresses these gaps by examining multidimensional housing-related vulnerability across Europe. Using data from 7,169 households across eight European countries—Belgium (Flanders only), Czech Republic, Germany, Estonia, Finland, Great Britain, Italy and Romania—it combines a K-Prototypes cluster analysis with a category-based relative risk analysis to identify household vulnerability profiles and assess cross-domain associations. By adopting a cross-country household-level perspective, the study makes three contributions. First, it extends research on housing inequality beyond dwelling-centred approaches by examining dwelling conditions, green-space access, and transport-related disadvantages simultaneously at the household level. Second, it identifies recurring configurations of vulnerability across diverse European contexts, providing a comparative typology of multidimensional housing-related disadvantage. Third, it tests whether dwelling vulnerability is more strongly associated with some external domains of inequality than others, by comparing its relationships with green-space and transport inequalities. In doing so, the study provides new evidence on which dimensions of inequality tend to co-occur, and which appear to be shaped by different contextual processes.

2. Methods

This study used a two-stage analytical strategy. First, we conducted a K-Prototypes cluster analysis to identify household profiles based on different combinations of dwelling, green-space, and transport inequalities. We then described each cluster using socio-demographic, contextual, and health and wellbeing indicators that were not included in the clustering procedure. This allowed us to characterise the resulting household profiles and examine how housing-related vulnerability intersected with wider socio-demographic, contextual, and health and wellbeing differences. Because K-Prototypes assigns observations to clusters using an additive distance measure across all included variables, domains represented by more indicators can have a stronger influence on the clustering solution (Huang 1998).

To support transparency and reproducibility, the cluster analysis was preregistered (Conzatti and Huebner 2025) at <https://archive.org/details/osf-registrations-p3bd2-v1>.

As the survey included more detailed information on dwelling inequalities than on green-space or transport inequalities, dwelling-related variables were likely to contribute more strongly to cluster formation. This reflects the structure of the distance metric rather than a deliberate weighting of dwelling conditions. We therefore carried out a second analysis to assess whether dwelling inequalities co-occurred with green-space and transport inequalities. For this purpose, we created one vulnerability indicator for each domain—dwelling, green space, and transport—and estimated the relative risk (RR) of green-space and transport inequality among households with dwelling vulnerability.

2.1. Dataset

This study used data from the 2025 multinational survey (<https://doi.org/10.5281/zenodo.20538396>). The dataset was collected, cleaned, and anonymised for submission under Task 4.1 of the Horizon Europe HouseInc project by Ghent University before being shared with the research team at the University of Exeter.

The survey was administered online between March and May 2025 by a professional market research agency across Belgium (Flanders), Czech Republic, Germany, Estonia, Finland, Great Britain, Italy, and Romania. Because the first round yielded insufficient responses in Czech Republic and Romania, an additional data collection was conducted

in these two countries during July and August 2025. The questionnaire was developed in English and translated into Dutch, Italian, German, Romanian, Finnish, Estonian, and Czech using a back-translation procedure to ensure cross-national comparability.

The original dataset comprised 26,422 respondents. The sampling strategy was designed to support comparative cross-national analyses rather than to produce nationally representative samples. Prior to analysis, responses were screened for eligibility, consent verification, and attention-check performance, and respondents who did not meet these quality-control criteria were excluded. After removal of these participants, we retained 14,951 participants. Individuals identifying as Gender Queer ($n = 14$), Non-Binary ($n = 28$), or those who did not report their gender ($n = 9$) were excluded due to the very small size of these subgroups. After the location responses were harmonised (e.g., "Suomi" coded as Finland and "Bayern" as Germany), a further 12 respondents living outside the target regions were excluded, resulting in 14,888 respondents.

The initial sample was intentionally skewed towards lower-income groups. To improve comparability with national population distributions, a subsample was created to approximate the distributions of income, age, and gender reported by Eurostat. The sample was selected sequentially, first by income, then by age, and finally by gender. This process had a cumulative effect on sample size and meant that alignment with the target distributions was slightly reduced at each stage, as later steps sometimes required small changes in participant selection. Nevertheless, the procedure aimed to achieve the best possible overall alignment across all three characteristics, with the final sample showing the closest alignment with the target gender distribution. This resulted in a final analytical sample of 7,169 respondents. A more detailed overview of the stratification process can be seen at <https://osf.io/sj659/files/rdyzyx>.

2.2. Data analysis

The analysis had two stages: cluster analysis and relative risk analysis.

2.2.1 Cluster analysis

Cluster analysis was used to identify household vulnerability profiles based on different combinations of dwelling, green-space, and transport inequalities in the pooled multinational sample. This method preserves information across multiple dimensions of

housing-related vulnerability. It also helps identify recurring household profiles that may differ in their socio-demographic, contextual, and health and wellbeing characteristics.

The clustering model included variables representing characteristics of the dwelling and its immediate living environment that were considered to be relatively independent of the household occupying the dwelling. This approach was intended to identify profiles based primarily on housing-related factors rather than on the socio-demographic characteristics or health status of occupants. Variables such as age, income, employment, household composition, and health were therefore excluded from the clustering procedure and instead used to characterise the resulting clusters. The variables included in the clustering model are listed in Table 1.

Although the clustering variables were selected to minimise dependence on occupant characteristics, complete independence was not possible. In particular, the measures of whether the home could be kept adequately warm or comfortably cool reflect not only dwelling characteristics but also household-related factors, including thermal preferences and financial resources. Nevertheless, these variables were retained because they provided the closest available indicators of heating difficulty and cooling discomfort in the dataset.

Before clustering, the data were cleaned and harmonised to ensure suitability for the K-Prototypes analysis, which combines numerical and categorical variables. Response categories selected by fewer than 25 respondents were merged into broader categories, binary responses (Yes/No), or a residual “Other” category to avoid sparse cells that could distort the clustering results. A complete list of category merges is provided in Table A 1 (Appendix A).

Table 1 Variables included in the cluster analysis.

Variable Name	Survey Variable	Vulnerability category		Type	Notes
SANITATION_ACCESS	SANITATION_ACCESS	Dwelling-level vulnerability	Housing condition	Categorical	Access to basic sanitation (sink, shower/bath, toilet).
KITCHEN_ACCESS	KITCHEN_ACCESS	Dwelling-level vulnerability	Housing condition	Binary	Whether the household has access to a kitchen.

SPACE_SHORT	SPACE_SHORT	Dwelling-level vulnerability	Housing condition	Categorical	Whether the respondent reports space shortage in the dwelling.
DWELL_PROB	DWELL_PROB	Dwelling-level vulnerability	Housing condition	Categorical	Presence of dwelling problems such as leaks, mould, or rot.
INADEQUATE_TRANSPORT (derived variable)	PUB_TRANSPORT_NO_USE_*	Transport vulnerability	Transport	Binary	Inadequate public transport access (e.g., distant, unsafe, unreliable, or unaffordable).
NO_TRANSPORT_OTHER (derived variable)	PUB_TRANSPORT_NO_USE_*	Transport vulnerability	Transport	Binary	No use of public transport for other reasons (e.g., preference for bike, car, or other).
GREEN_SPACE_*	GREEN_SPACE_*	Green space vulnerability	Green space	Categorical	Access to green space. Used for both cluster analysis and cluster description.
HAS_GREEN_SPACE (derived variable)	GREEN_SPACE_*	Green space vulnerability	Green space	Binary	Access to any green space, including gardens, balconies, or parks. Used for cluster description only.
NO_GREEN_SPACE (derived variable)	GREEN_SPACE_NONE	Green space vulnerability	Green space	Binary	No access to green space. Used for cluster description only.
EPC	EPC	Dwelling-level vulnerability	Housing condition	Ordinal	Self-estimated energy label from A (most efficient) to G (least).
AFF_HEAT	AFF_HEAT	Dwelling-level vulnerability	Housing condition	Categorical	Ability to afford adequate heating throughout the dwelling (yes, some rooms only, no).
COMF_COOL	COMF_COOL	Dwelling-level vulnerability	Housing condition	Categorical	Ability to keep the dwelling comfortably cool in summer (yes, some rooms only, no).
LACK_ENERGY_APP	LACK_ENERGY_APP	Dwelling-level vulnerability	Housing condition	Categorical	Lack of essential household appliances.

Clustering method

Household profiles were identified using the K-Prototypes algorithm, which is designed for datasets containing both numerical and categorical variables (Huang, 1998). K-

Prototypes was chosen to identify empirically recurring combinations of vulnerabilities rather than latent probabilistic classes. Unlike model-based approaches such as latent class analysis, it does not rely on distributional assumptions or probabilistic class membership and instead groups observations according to their similarity across the selected indicators. This similarity is quantified using a distance function that combines squared Euclidean distance for numerical variables and simple matching dissimilarity for categorical variables:

$$P = \sum_{l=1}^k \sum_{i=1}^n w_{i,l} d(X_i, Q_l) \quad (1)$$

where P is the objective function, $w_{i,l}$ is the element of the partition matrix $W_{n \times k}$ indicating whether observation i is assigned to cluster l , Q_l is the prototype of cluster l , and d is the mixed-type distance between observations and prototypes.

Distances were calculated using a combination of squared Euclidean distance for numerical variables and simple matching dissimilarity for categorical variables, weighted by a scaling parameter γ :

$$d(X_i, Q_l) = \sum_{j=1}^p (x_{ij}^r - q_{lj}^r)^2 + \gamma \sum_{j=p+1}^m \delta(x_{ij}^c, q_{lj}^c) \quad (2)$$

where the first term represents squared Euclidean distance across numerical attributes, and the second represents simple matching dissimilarity across categorical attributes.

Clusters were initialised randomly and updated iteratively until cluster assignments stabilised (Li et al. 2019). The final number of clusters was selected based on the elbow heuristic and the interpretability of the resulting cluster structure. Cluster cohesion and separation were assessed using the silhouette coefficient:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (3)$$

where $a(i)$ is the average dissimilarity of i with all other points in the same cluster, and $b(i)$ is the lowest average dissimilarity of i to any other cluster.

To support interpretation of the cluster structure, dimensionality-reduction techniques were also applied. Principal Component Analysis (PCA) was used to summarise patterns among the numerical variables. Uniform Manifold Approximation and Projection (UMAP) was applied using a precomputed Gower dissimilarity matrix to produce a two-dimensional embedding of the mixed-type dataset while preserving local similarities between households. Gower dissimilarities were also used to calculate the silhouette coefficient because the clustering dataset contained both numerical and categorical variables. These projections and statistics were used to assess cluster separation and internal structure.

Cluster description and statistical analysis

Once cluster membership had been established, the clusters were described using socio-demographic, contextual, and health and wellbeing indicators that were not included in the clustering algorithm (Table 2). Cross-tabulations and summary statistics were used to examine the distribution of these characteristics across clusters. To assess whether the distributions of these characteristics differed across clusters, categorical variables were tested using the chi-square test of independence, while continuous variables were tested using the Kruskal–Wallis test (Bhattacharyya 2014). To account for multiple comparisons, p-values were adjusted using the false discovery rate procedure.

Table 2 Variables used to describe the clusters.

Variable Name	Survey Variable	Category	Type	Notes
AGE	AGE	Socio-demographic	Ordinal	Calculated from year of birth; grouped into standard age bands (e.g. 18–19, 20–64, 65+).
GENDER	GENDER	Socio-demographic	Categorical	Low frequency responses (e.g., < 5%) may be merged.
EDU	EDU	Socio-demographic	Categorical	Education level. Grouping may be applied if categories are sparse.
EMP	EMP	Socio-demographic	Categorical	Employment status. Grouping may be applied if categories are sparse.
HHTYPE	HHTYPE	Socio-demographic	Categorical	Household type.
GEN_HEALTH	GEN_HEALTH	Health	Categorical	Self-rated general health.
CHRON_ILL	CHRON_ILL	Health	Categorical	Presence of chronic illness in the household;

				multiple responses permitted.
WELLBE	WELLBE	Health	Ordinal	Current and future life satisfaction on a 0–10 scale.
FAM_TOTAL	FAM_TOTAL	Socio-demographic	Continuous	Total number of household members, derived from age-group composition.
COUNTRY_R	COUNTRY_R	Socio-demographic	Categorical	Country of residence. Used for the multi-country level analysis.
COUNTRY_B	COUNTRY_B	Socio-demographic	Categorical	Whether respondent was born in country of residence.
ETH	ETH	Socio-demographic	Categorical	Perceived ethnic/racial similarity with majority population.
BORN_ETH_COMBO	COUNTRY_B, ETH	Socio-demographic	Categorical	Combined indicator of country of birth and perceived ethnicity, intended to capture intersections between migration background and perceived ethnic minority status.
LIVING_AREA	LIVING_AREA	Contextual	Categorical	Urbanicity of residence: city, suburb, rural.
FIN_COMF	FIN_COMF	Socio-demographic	Categorical	Perceived financial comfort. Used for cluster description.
OWNERSHIP	OWNERSHIP	Socio-demographic	Categorical	Type of tenure (e.g., rented, owned).
DWELLING	DWELLING	Contextual	Categorical	Type of dwelling (e.g., flat, detached).
DISC_SELF_1 to DISC_SELF_10	DISC_SELF_1 to DISC_SELF_10	Socio-demographic	Binary	Binary indicators of self-reported housing discrimination across the specified grounds.
CARS_OWN	CARS_OWN	Socio-demographic	Continuous	Number of cars owned. Indicator of household mobility resources; it may also partly reflect material resources.
HEAT_SRC	HEAT_SRC	Contextual	Categorical	Primary source of heating.
FINBURDEN_ECOST	FINBURDEN_ECOST	Socio-demographic	Categorical	Perceived financial burden from energy costs.
FIN_BURDEN_HOUSECOST	FIN_BURDEN_HOUSECOST	Socio-demographic	Categorical	Subjective burden of housing costs.
AIRCO	AIRCO	Contextual	Categorical	Availability of air conditioning.

2.2.2 Risk analysis

The cluster analysis was exploratory and reflected the full set of indicators included in the clustering model. Because dwelling-related indicators were more numerous than green-space or transport indicators, dwelling characteristics were likely to exert a greater influence on the clustering solution. To complement this analysis, we conducted a relative risk (RR) analysis in which each domain was represented by a single vulnerability measure. This approach ensured that each domain was represented by one analytical measure in the comparison of cross-domain associations.

Vulnerability categories

For the relative risk analysis, ordered vulnerability categories were developed for three domains: dwelling conditions, green-space access, and transport accessibility. These categories provided a common framework for comparing inequalities across domains while preserving different levels of vulnerability.

For dwelling conditions, a cumulative dwelling vulnerability score was calculated using survey indicators capturing dwelling problems, including damp, mould and leaks; insufficient living space; lack of sanitation or kitchen facilities; and heating difficulty and cooling discomfort. Each indicator was coded as 1 when the vulnerability was present and 0 otherwise, and summed to produce a cumulative score representing the number of dwelling-related disadvantages experienced simultaneously. Air-conditioning ownership was excluded because it was reported by more than two-thirds of households and does not represent a vulnerability in all climatic contexts. The resulting score was then classified into four ordered dwelling vulnerability categories, ranging from Category 1 (no dwelling vulnerability) to Category 4 (high or severe dwelling vulnerability), with intermediate categories representing increasing levels of disadvantage.

Green-space vulnerability was classified into five ordered categories reflecting decreasing levels of access to green space, ranging from Category 1 (access to private green space) to Category 5 (no access to any green space). Intermediate categories represented progressively lower levels of access, such as shared green space or nearby green space only.

Transport vulnerability was classified into seven ordered categories based on combinations of public transport availability, public transport use and its constraints, and car ownership. The two most severe original categories were merged to avoid very small cell counts and improve the stability of the estimates. The final classification ranged from Category 1, representing access to both public transport and a car, to Category 7, representing the most constrained mobility situation: no car ownership combined with either constrained public transport use or no public transport access. Intermediate categories represented varying combinations of transport access, reliance, and mobility constraints. Car non-ownership was therefore treated as contributing to transport vulnerability only when combined with constrained public transport use or lack of public transport access.

Relative risk calculation

For the RR analysis, the ordered vulnerability categories described above were further recoded into binary exposure and outcome variables.

Dwelling vulnerability was treated as the analytical exposure, without implying a causal relationship. Households were classified as exposed if they belonged to any dwelling vulnerability category above Category 1 (no dwelling vulnerability). Green-space inequality and transport inequality were treated as the outcomes, with households classified as experiencing the outcome if they belonged to categories representing moderate or higher vulnerability. Relative risks were calculated using standard 2 × 2 contingency tables (Table 3), comparing the probability of each outcome among households with and without dwelling vulnerability:

$$RR = \frac{a/(a + b)}{c/(c + d)} \quad (4)$$

where a represents households with dwelling vulnerability who also experienced the outcome, b households with dwelling vulnerability who did not experience the outcome, c households without dwelling vulnerability who experienced the outcome, and d households without dwelling vulnerability who did not experience the outcome.

Table 3 Structure of the 2×2 contingency table used to calculate relative risk (RR) of green-space or transport inequality according to dwelling vulnerability.

	Outcome present	Outcome absent
Exposed (dwelling vulnerability)	a	b
Unexposed (no dwelling vulnerability)	c	d

To examine how cross-domain associations varied across increasing levels of dwelling vulnerability, relative risks were estimated separately for the low, moderate, and high/severe dwelling-vulnerability categories, using households in Category 1 (no vulnerability) as the reference group. Finally, to assess whether associations between dwelling vulnerability and other inequality domains varied across national contexts, relative risks were also estimated separately for each country. Country-stratified relative risks were estimated using the same binary exposure and outcome definitions as the pooled relative risk analysis. Thus, within each country, households with one or more dwelling vulnerabilities were compared with households reporting no dwelling vulnerability, and the corresponding binary indicators of green-space and transport inequality were defined using the same thresholds as in the pooled sample. Ninety-five per cent confidence intervals were calculated using the Katz log method (Katz et al. 1978).

2.3. Deviation from the preregistration

Deviations from the preregistration were limited to three areas. First, although both country-specific and pooled cluster analyses were initially planned, only the pooled multinational analysis was retained because the primary objective of the study was to identify cross-national household vulnerability profiles. Second, an additional relative risk analysis based on the vulnerability categories developed for dwelling conditions, green-space access, and transport accessibility was introduced to examine whether dwelling vulnerability co-occurred with green-space and transport inequalities. Third, principal component analysis (PCA), Uniform Manifold Approximation and Projection (UMAP), and silhouette statistics were used to support the interpretation and evaluation of the clustering solution. These modifications were made in response to data

characteristics identified during data preparation and analysis and do not alter the substantive interpretation of the findings.

3. Results

This section reports the cluster analysis, the distribution of vulnerability categories, and the relative risk analysis.

3.1. Cluster results

Cluster analysis using the K-Prototypes algorithm identified a four-cluster solution based on the elbow criterion and interpretability (Figure 1). The four-cluster solution was selected because the cost curve showed diminishing improvements beyond $k = 4$, while retaining greater interpretive granularity than the three-cluster solution. The Gower-based silhouette coefficient was positive but low for $k = 4$ (approximately 0.06), lower than for $k = 3$ (approximately 0.07–0.08), but higher than for $k = 5$, for which the coefficient was negative. These results indicate substantial overlap between clusters in mixed-feature space. Therefore, the clusters should not be interpreted as sharply separated population groups. Rather, they provide an exploratory typology of recurring vulnerability configurations within a high-dimensional mixed survey dataset. PCA and UMAP visualisations provided partial support for this interpretation: PCA projections showed limited but visible separation between clusters (Figure 2a), while UMAP showed clearer local differentiation (Figure 2b).

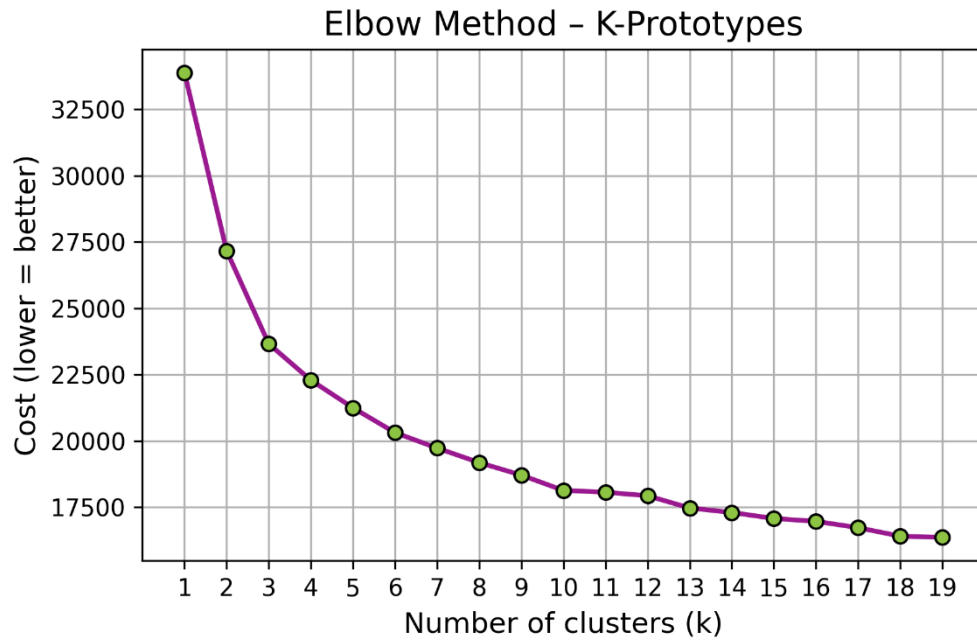


Figure 1 Elbow plot used to inform selection of the number of clusters in the K-Prototypes analysis.

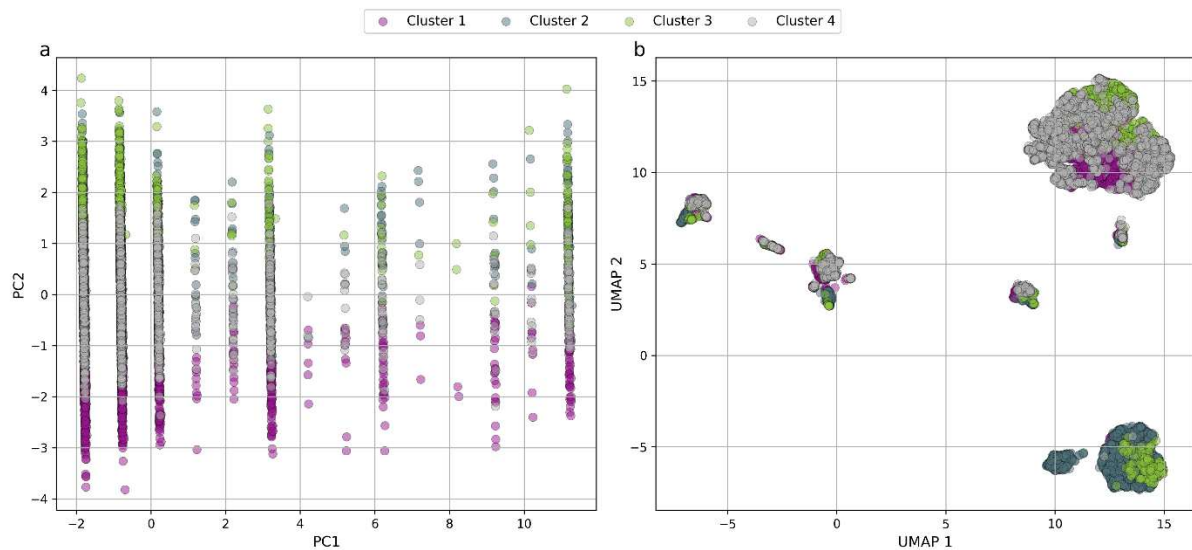


Figure 2 PCA and UMAP projections of household clusters: (a) PCA projection based on the numerically encoded clustering variables, annotated with the proportion of variance explained by each component; and (b) UMAP projection based on Gower dissimilarities.

The distribution of households across clusters is shown in Table 4. Cluster 2 was the largest group ($n = 2,608$; 36.4%), followed by Cluster 1 ($n = 2,320$; 32.4%), while Clusters 3 ($n = 1,148$; 16.0%) and 4 ($n = 1,093$; 15.2%) were smaller. Compared with the three-

cluster solution, which was dominated by a single cluster comprising 54.1% of households ($n = 3,879$), the four-cluster solution yielded a more even distribution across groups. In contrast, the five-cluster solution introduced a smaller cluster comprising 9.9% of households ($n = 713$).

Table 4 Number of households in each cluster.

Cluster number	Household count
Cluster 1	2,320
Cluster 2	2,608
Cluster 3	1,148
Cluster 4	1,093
Total	7,169

To support interpretation of the clustering structure, a principal component analysis (PCA) was conducted. The first five principal components explained 54.2% of the total variance, with the first four components capturing 48.6% (Table 5). Examination of the dominant loadings indicated that the principal dimensions corresponded to dwelling energy efficiency (PC1), mobility resources (PC2), household appliance adequacy (PC3), private green-space access (PC4), and cooling comfort (PC5). Variables related to dwelling problems displayed lower communalities. Cluster centroids in PCA space (Table 6) broadly aligned with these dimensions; notably, Cluster 3 had the highest score on PC1, consistent with stronger housing energy performance, while Cluster 4 had the highest score on PC2, reflecting the strong contribution of car ownership to this component.

To examine whether the clusters differed on the descriptive variables (Table 2), categorical variables were tested using the chi-square test of independence, while continuous variables were tested using the Kruskal–Wallis test (Bhattacharyya, 2014). The full results are reported in Appendix B. Most variables differed significantly across clusters after false discovery rate adjustment (Table B 1 and Table B 2), providing further support for the interpretability of the cluster profiles.

Table 5 Strongest contributor to each principal component (PC1–PC5). Each variable listed is the one with the highest absolute loading on its respective principal component, based solely on PCA results.

Principal component	Variance explained [%]	Dominant variable	Loading	Communality	Interpretation
PC1	18.7	EPC: very high energy efficiency	0.82	1.00	Housing energy performance
PC2	15.8	Car ownership	0.82	1.00	Car ownership / household mobility resources
PC3	8.1	No lack of energy appliances	0.34	0.22	Appliance availability / material comfort
PC4	6.0	Private garden access	0.27	0.21	Private green space
PC5	5.6	Adequate cooling comfort	0.35	0.22	Ability to keep comfortably cool

Table 6 Cluster positioning in PCA space (mean PC scores).

Cluster number	PC1 Housing energy performance	PC2 Car ownership / household mobility resources	PC3 Appliance availability / material comfort	PC4 Private green space	PC5 Ability to keep comfortably cool
Cluster 1	-1.10	-0.07	0.05	0.02	0.08
Cluster 2	0.25	-0.91	-0.00	0.04	-0.06
Cluster 3	1.48	0.69	-0.19	-0.05	-0.02
Cluster 4	0.19	1.58	0.09	-0.00	-0.01

3.1.1 Cluster characteristics

Table 7 summarises the prevalence of the clustering indicators within each cluster. Socio-demographic, financial, contextual, and health and wellbeing variables were not used to generate the clusters; instead, they were examined subsequently to characterise and interpret the cluster profiles.

Based on the variables included in the clustering procedure, the four clusters revealed distinct configurations of dwelling, green-space, and transport vulnerability (Table C 1). The accumulated housing and environmental vulnerability cluster (Cluster 1) showed the broadest concentration of dwelling-related disadvantage, with comparatively high levels of damp, mould, leaks, and space shortage, alongside more limited access to private and nearby green space. The thermal and energy vulnerability cluster (Cluster 2) was

characterised primarily by heating difficulty and cooling discomfort, while structural dwelling problems were less prevalent than in Clusters 1 and 4. The resource-advantaged with thermal strain cluster (Cluster 3) displayed the lowest prevalence of structural dwelling problems and the strongest access to private green space, although cooling discomfort remained common. Finally, the dwelling problems and transport vulnerability cluster (Cluster 4) combined persistent dwelling problems, including damp and mould, with weaker transport access than the other clusters. Based on these patterns, the four clusters were interpreted as representing: (1) accumulated housing and environmental vulnerability; (2) thermal and energy vulnerability; (3) resource-advantaged households with thermal strain; and (4) dwelling problems and transport vulnerability.

The accumulated housing and environmental vulnerability cluster showed the greatest concentration of multiple vulnerabilities, whereas the thermal and energy vulnerability cluster and the dwelling problems and transport vulnerability cluster exhibited more specific but still overlapping forms of disadvantage. The resource-advantaged with thermal strain cluster was comparatively advantaged in terms of dwelling quality and green-space access, although cooling discomfort remained evident (see Figure 3 and Table 7).

Table 7 Summary of cluster characteristics based on variables included in the clustering analysis.

Domain	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Dwelling quality	Highest prevalence of dwelling problems, including damp, mould and leaks; frequent space shortage	Fewer structural dwelling problems than Clusters 1 and 4	Lowest prevalence of dwelling problems; most favourable dwelling quality	Dwelling problems relatively common, particularly damp and mould
Thermal strain	High prevalence of heating difficulty; widespread cooling discomfort	Strongest concentration of thermal strain, with very high heating difficulty and high cooling discomfort	Dwelling quality relatively favourable, but cooling discomfort remains common	Heating difficulty and cooling discomfort both common
Space and appliances	Lowest rooms-per-person; frequent appliance deprivation	Moderate space and appliance constraints	Largest dwellings; lowest appliance deprivation	Moderate appliance deprivation

Green-space access	Least access to private and nearby green space	Limited access to private gardens but better access to nearby green space	Strongest access to private gardens	Good access to private gardens, but weaker overall green-space access than Cluster 3
Transport access	Generally good public-transport access	Relatively favourable public-transport access	Moderate transport access	Weakest public-transport access, despite comparatively strong car ownership

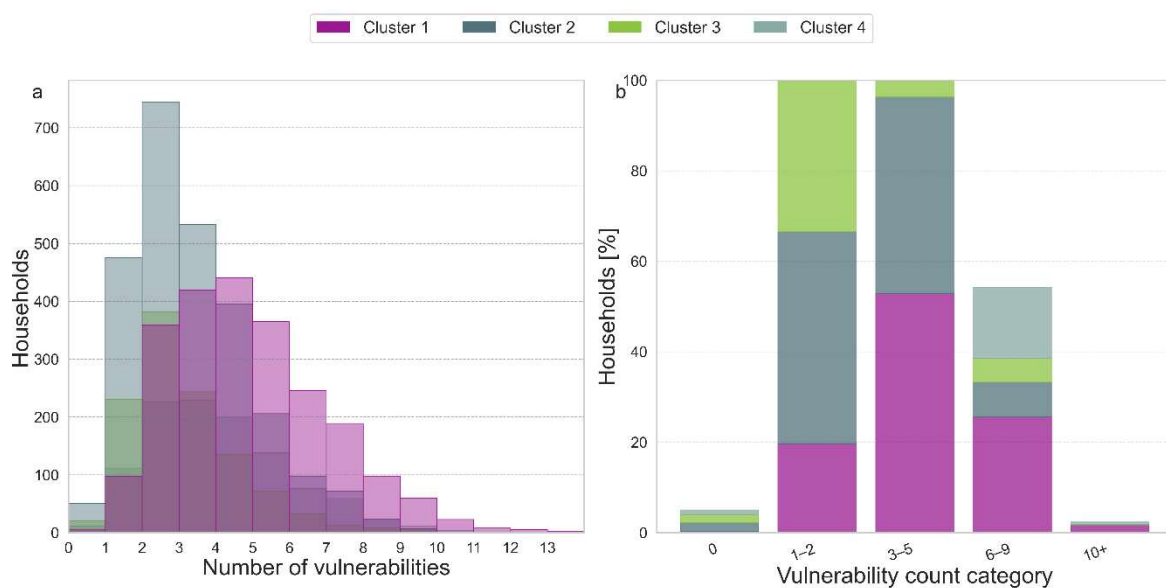


Figure 3 a) Distribution of the number of vulnerability indicators per household used in the cluster analysis, by cluster. (b) Percentage distribution of households across vulnerability count categories by cluster.

3.1.2 Cluster description

The cluster solution was subsequently examined using a separate set of socio-demographic, contextual, and health and wellbeing variables that were not included in the clustering procedure. Differences across clusters were assessed using chi-square tests for categorical variables and Kruskal–Wallis tests for continuous variables, with full results reported in Appendix B. Table 8 provides a substantive summary of the resulting cluster profiles. Detailed descriptive statistics of the cluster profiles are reported in Table C 2–Table C 4 in Appendix C.

The accumulated housing and environmental vulnerability cluster (Cluster 1) was characterised by the highest levels of financial strain and housing-related discrimination, alongside lower employment and a relatively large proportion of single-person households. Contextually, households in this cluster were more likely to live in urban areas and reported the lowest levels of private green-space access. This cluster also displayed the least favourable health profile, with the highest prevalence of poor self-rated health and the lowest wellbeing scores.

The thermal and energy vulnerability cluster (Cluster 2) was characterised by an older population and moderate levels of employment and financial comfort. Households were predominantly located in urban settings and occupied an intermediate position across most contextual indicators. Self-rated health was generally more favourable than in Clusters 1 and 4, although household chronic illness remained relatively common.

The resource-advantaged with thermal strain cluster (Cluster 3) emerged as the most socio-economically advantaged group. Households reported the highest levels of employment and financial comfort and were more likely to consist of family households. They were disproportionately located in suburban and rural areas, reported the highest levels of private green-space access and car ownership, and exhibited the most favourable overall health profile, including the highest wellbeing scores and lowest prevalence of household chronic illness.

The dwelling problems and transport vulnerability cluster (Cluster 4) occupied an intermediate socio-economic position but displayed clear signs of vulnerability. Households reported greater financial strain than those in Clusters 2 and 3 and a comparatively high prevalence of household chronic illness. They were more commonly located in suburban and rural areas and showed the weakest transport access of all clusters. Self-rated health was less favourable than in Clusters 2 and 3, and household chronic illness was comparatively common.

These findings indicate that housing-related vulnerabilities intersected with socio-demographic, contextual, and health and wellbeing differences in distinct ways, producing recognisable profiles of multidimensional disadvantage and advantage across households.

Table 8 Summary of cluster profiles across socio-demographic, contextual, and health and wellbeing domains.

Domain	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Socio-demographic	High share of single-person households (33.8%); substantial older-adult presence; lower employment; highest discrimination (~42%); strong financial strain	Older population; mixed financial comfort (“coping” common); moderate employment	Highest employment and financial comfort; comparatively high education; predominantly family households	Mixed household types; moderate employment; higher financial strain
Contextual	More urban profile, lowest private green-space access, and fewer mobility resources	More urban profile, with moderate contextual resources	Greater suburban and rural representation, highest access to private green space, and highest car ownership	More suburban and rural representation, with mixed contextual conditions
Health	Highest burden of poor self-rated health; lowest wellbeing scores	Moderate health profile; chronic illness remains notable (44%)	Most favourable health profile; highest wellbeing; lowest prevalence of household chronic illness	High prevalence of household chronic illness; below-average wellbeing
Overall profile	Most structurally disadvantaged: multiple overlapping vulnerabilities	Thermal and energy vulnerability with moderate socio-economic strain	Socio-economically advantaged, with favourable dwelling and green-space conditions but continued thermal strain	Dwelling problems and transport vulnerability

3.2. Vulnerability categories

This section describes the distribution of households across the dwelling, green-space, and transport vulnerability categories developed for the relative risk analysis. These

categories summarise increasing levels of vulnerability within each domain and provide the basis for the subsequent analysis of cross-domain associations.

3.2.1 Dwelling vulnerability

Figure 4 shows the pooled prevalence of the dwelling-vulnerability indicators included in the cumulative dwelling vulnerability score. In the pooled sample, 25.8% of respondents reported some cooling discomfort, 22.9% reported severe cooling discomfort, and 23.5% reported difficulty affording adequate heating. Figure 5 shows the distribution of households across the four dwelling vulnerability categories. Most households (57.1%) fell into Category 2, indicating exposure to one or two dwelling vulnerabilities, while 24.0% were in Category 3, with three or four vulnerabilities. Smaller shares were found

at the extremes: 12.3% of households were in Category 1, with no dwelling vulnerabilities, and 6.7% were in Category 4, with five or more vulnerabilities.

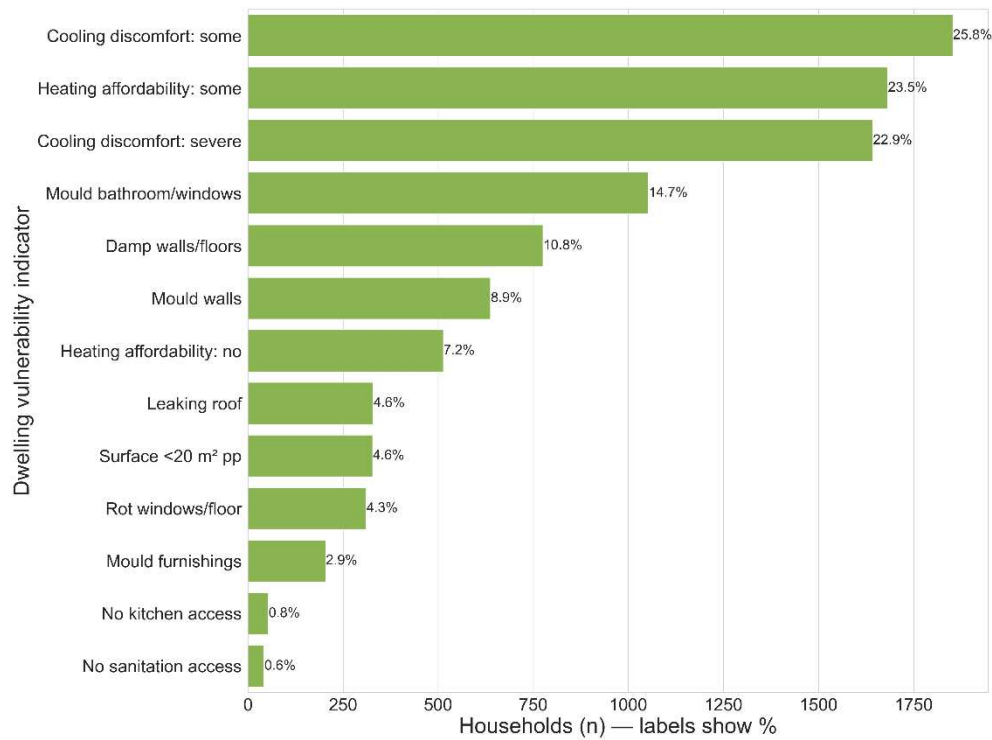


Figure 4 Prevalence of dwelling vulnerability indicators in the pooled sample (N = 7,169).

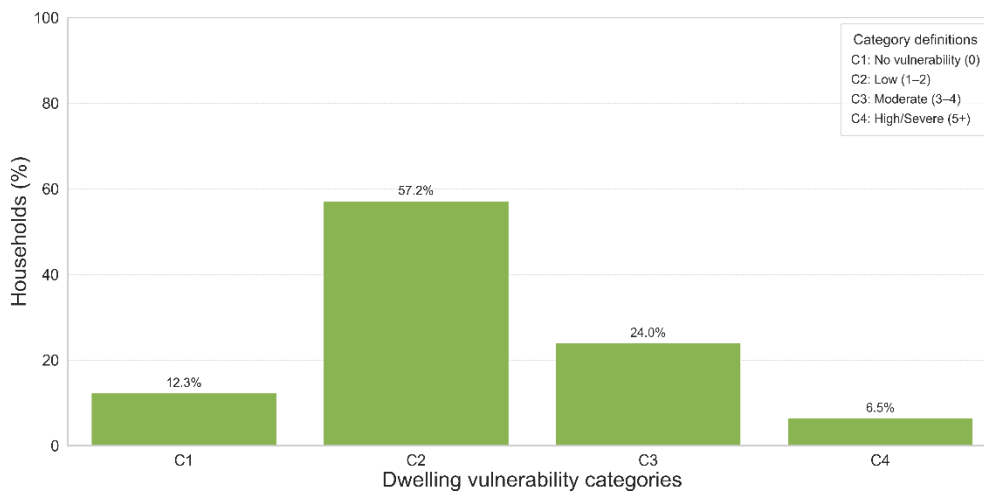


Figure 5 Distribution of dwelling vulnerability categories.

3.2.2 Green-space categories

Half of households (50.0%) reported access to a private garden (Category 1), while 32.0% had access only to a private balcony, patio, or roof terrace (Category 2). A small minority (5.4%) reported no access to any green space (Category 5) (Figure 6; Table 9).

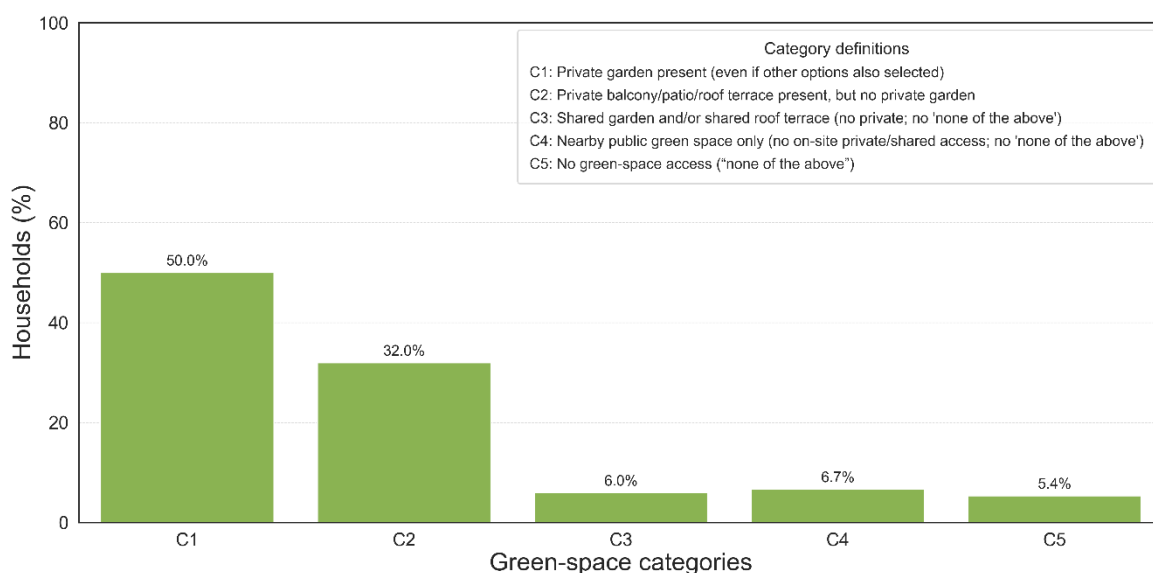


Figure 6 Distribution of households across green-space access categories.

3.2.3 Transport categories

Transport vulnerability was examined using the seven-category analytical typology. Category 1, representing public transport use and car availability, accounted for 41.3% of households. Category 5, representing car availability but non-use of public transport due to constraints, accounted for 24.5%. The most severe category, combining

households without a car and with either public-transport barriers or no public-transport access, accounted for 1.6% of households (Figure 7; Table 10).

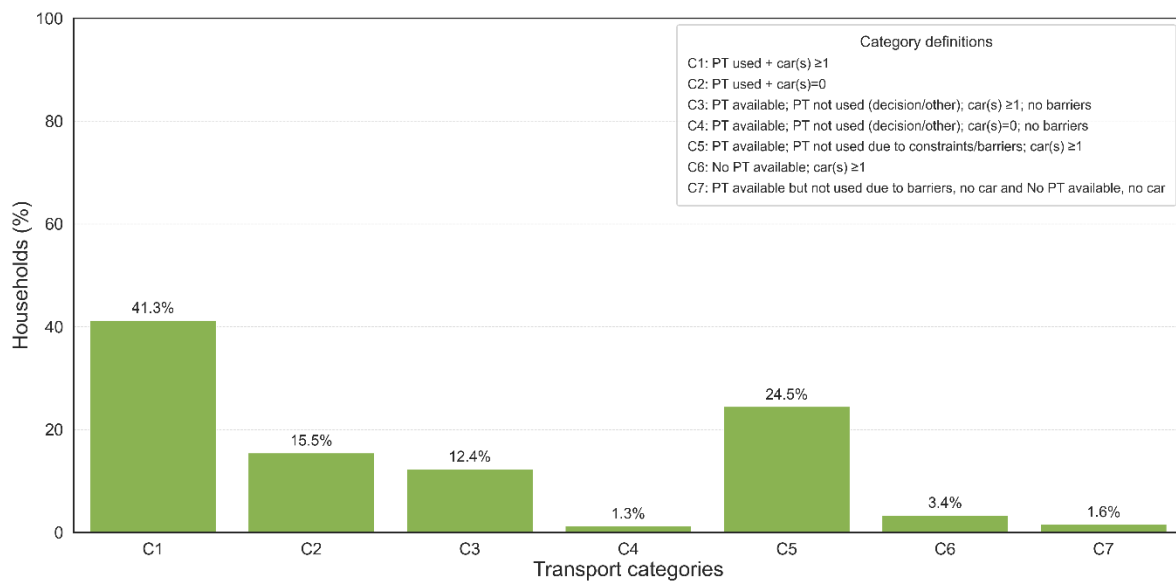


Figure 7 Distribution of households across transport vulnerability categories.

Table 9 Green-space access categories and household distribution.

Green space category	Vulnerability level	Category description	Households [n (%)]
Category 1	No vulnerability	Access to private garden	3586 (50.0 %)
Category 2	Mild vulnerability	- Access to private balcony/patio/roof terrace - No access to private garden	2293 (32.0 %)
Category 3	Moderate vulnerability	- Access to garden/roof terrace - No access to any private green space	427 (6.0 %)
Category 4	High vulnerability	- Access to nearby green space only - No access to private and shared green space	478 (6.7 %)
Category 5	Severe	No access to green space	385 (5.4 %)

Table 10 Transport vulnerability categories and household distribution.

Transport category	Vulnerability level	Category description	Households [n (%)]
Category 1	No vulnerability	- Access to public transport - Access to car(s)	2958 (41.3 %)
Category 2	Low vulnerability	- Access to public transport - No access to car(s) - Public transport used	1111 (15.5%)
Category 3	Mild vulnerability	- Access to public transport - Access to car(s) - Public transport not used due to decision (walk and/or cycle)	891 (12.4%)
Category 4	Moderate vulnerability	- Access to public transport - No access to car(s) - Public transport not used due to decision (walk and/or cycle)	92 (1.3%)
Category 5	High vulnerability	- Access to public transport - Access to car(s) - Public transport not used due to constraints	1759 (24.5%)
Category 6	Very high vulnerability	- No access to public transport - Access to car(s)	246 (3.4%)
Category 7	Severe vulnerability	- No car available - Either public transport barriers or no public transport access	112 (1.6%)

3.3. Risk analysis

Relative risk (RR) was estimated to assess whether households with dwelling vulnerability were more likely to experience green-space or transport inequality than households with no dwelling vulnerability.

3.3.1 Co-occurrence of dwelling, green-space, and transport inequalities

Binary RR estimates for the pooled sample are shown in Table 11. Households with dwelling vulnerability were more likely to experience green-space inequality than households with no dwelling vulnerability (RR = 1.57; 95% CI: 1.39–1.77). By contrast, dwelling vulnerability was associated with a lower likelihood of transport inequality in the pooled sample (RR = 0.86; 95% CI: 0.81–0.92).

Table 11 Relative risk of green-space and transport inequality according to dwelling vulnerability in the pooled sample, with Katz 95% confidence intervals.

Vulnerability	Households [n]				Risk exposed	Risk unexposed	RR	CI
	Exposed with outcome	Exposed without outcome	Unexposed with outcome	Unexposed without outcome				
Green space	995	3893	295	1986	0.20	0.12	1.57	1.39-1.77
Transport	2013	2875	1087	1194	0.41	0.48	0.86	0.81-0.92

3.3.2 Associations across dwelling-vulnerability severity levels

Relative risks across dwelling-vulnerability severity levels are shown in Table 12 and Table 13. To aid interpretation, these tables report both the relative risk and the number of households in each combination of dwelling-vulnerability level and green-space or transport-vulnerability category.

For green-space vulnerability, compared with households with no dwelling vulnerability, those in the low, moderate, and high/severe dwelling-vulnerability groups generally had elevated relative risks for the more severe green-space categories. The clearest pattern was observed among households with moderate dwelling vulnerability, for whom the relative risk was 2.51 for moderate green-space vulnerability and 1.99 for severe green-space vulnerability. Elevated risks were also observed in the low dwelling-vulnerability group, particularly for severe green-space vulnerability (RR = 1.83). In the high/severe dwelling-vulnerability group, relative risks were also above 1 for moderate and severe green-space vulnerability, although the number of households in these categories was small.

No comparable pattern was observed for transport vulnerability. Relative risks varied across dwelling-vulnerability levels and transport categories, with no clear increase as dwelling vulnerability became more severe. For example, households in the high/severe dwelling-vulnerability group were more likely to fall into the low transport-vulnerability category (RR = 2.56) and the severe category (RR = 2.18), but less likely to fall into the mild and high transport-vulnerability categories (RR = 0.47 and RR = 0.90, respectively). A similar mixed pattern appeared in the moderate dwelling-vulnerability group. More generally, the severity-gradient estimates should be interpreted cautiously because

several of the more severe dwelling-, green-space-, and transport-vulnerability categories contained small numbers of households.

Table 12 Relative risk and household counts of green-space categories across dwelling vulnerability levels. Rows indicate dwelling vulnerability levels and columns indicate green-space vulnerability categories*.

Green-space vulnerability levels										
Dwelling vulnerability level	Relative Risk					Household count				
	No vulnerability	Mild vulnerability	Moderate vulnerability	High vulnerability	Severe vulnerability	No vulnerability	Mild vulnerability	Moderate vulnerability	High vulnerability	Severe vulnerability
No vulnerability	1.00	1.00	1.00	1.00	1.00	1273	713	98	119	78
Low	0.83	1.09	1.33	1.42	1.83	1758	1291	216	281	237
Moderate	0.87	0.85	2.51	1.43	1.99	420	231	94	65	59
High / Severe	1.03	0.79	1.87	1.06	1.36	135	58	19	13	11

*Values above 1 indicate higher likelihood than the reference group, values below 1 indicate lower likelihood, and values equal to 1 indicate the same likelihood as the reference group. The reference group is households with no dwelling vulnerability.

Table 13 Relative risk and household counts for transport vulnerability categories across dwelling vulnerability levels. Values above 1 indicate higher likelihood than the reference group, values below 1 indicate lower likelihood, and values equal to 1 indicate the same likelihood as the reference group. The reference group is households with no dwelling vulnerability*.

Transport vulnerability levels														
Dwelling vulnerability level	Relative Risk							Household count						
	No vulnerability	Low vulnerability	Mild vulnerability	Moderate vulnerability	High vulnerability	Very High vulnerability	Severe vulnerability	No vulnerability	Low vulnerability	Mild vulnerability	Moderate vulnerability	High vulnerability	Very High vulnerability	Severe vulnerability

No vulnerab ility	1.00	1.00	1.00	1.00	1.00	1.00	1.00	945	249	374	29	578	75	31
Low	1.00	1.53	0.70	1.02	0.95	1.06	1.09	1575	630	433	49	908	132	56
Moderat e	0.99	1.75	0.46	0.99	0.99	1.20	1.25	357	166	66	11	219	32	18
High / Severe	0.83	2.56	0.47	1.00	0.90	0.90	2.18	81	66	18	3	54	7	7

*Values above 1 indicate higher likelihood than the reference group, values below 1 indicate lower likelihood, and values equal to 1 indicate the same likelihood as the reference group. The reference group is households with no dwelling vulnerability.

3.4. Country-stratified relative risks

Country-specific estimates are reported in Table 14. These estimates used the same binary definitions as the pooled relative risk analysis. For green-space inequality, point estimates were above 1 in all countries, indicating that households with one or more dwelling vulnerabilities were generally more likely to experience green-space inequality than households with no dwelling vulnerability, which formed the reference group. The largest point estimate was observed in Belgium (RR = 2.96; 95% CI: 1.31–6.70), followed by Great Britain (RR = 1.75; 95% CI: 1.25–2.45). The wide confidence interval for Belgium indicates lower precision, likely reflecting small cell counts in the country-specific analysis. In the remaining countries, point estimates also exceeded 1, although most confidence intervals included 1, indicating uncertainty in the country-specific estimates. For transport inequality, the country-specific estimates were weaker and less consistent. Most relative risks were close to 1, and several were below 1, indicating no clear or uniform association across countries. Confidence intervals included 1 in most cases, although the estimates for Finland (RR = 0.84; 95% CI: 0.74–0.95) and Romania (RR = 0.78; 95% CI: 0.61–0.99) were below 1.

Table 14 Country-stratified relative risks of green-space and transport inequality according to dwelling vulnerability, with 95% confidence intervals*.

Country	N of households used	Green-space inequality RR	95% CI	Transport inequality RR	95% CI
Belgium	638	2.96	1.31–6.70	0.88	0.77–1.01
Czech Republic	597	1.28	0.93–1.76	1.00	0.73–1.37
Germany	1097	1.33	0.94–1.86	1.01	0.88–1.15
Estonia	809	1.19	0.93–1.53	0.94	0.74–1.21
Finland	1041	1.27	0.93–1.72	0.84	0.74–0.95
Great Britain	1296	1.75	1.25–2.45	0.92	0.79–1.07
Italy	996	1.34	0.94–1.93	0.97	0.87–1.08
Romania	695	1.31	0.94–1.84	0.78	0.61–0.99

* Dwelling vulnerability was defined as the presence of one or more dwelling vulnerabilities, compared with no dwelling vulnerability. Green-space inequality was defined as moderate, high, or severe green-space vulnerability, compared with no or mild green-space vulnerability. Transport inequality was defined using the merged transport typology applied in the relative risk analysis, with mild to severe transport vulnerability compared with no or low transport vulnerability.

4. Discussion

This study examined multidimensional housing inequality across eight European countries using (i) cluster analysis to identify recurring household vulnerability profiles and (ii) relative risk analysis to assess whether dwelling vulnerability co-occurred with green-space and transport inequalities. Together, the findings show that housing-related vulnerability can extend beyond dwelling conditions, although associations differed across domains. Dwelling vulnerability was positively associated with green-space inequality in the pooled analysis, with point estimates above 1 across countries, whereas its association with transport inequality was weaker and less consistent. These findings support multidimensional approaches to housing inequality while also showing that different domains of disadvantage do not necessarily overlap in the same way (Guio et al. 2021; James et al. 2024; Born et al. 2025).

4.1. How disadvantages cluster within households

The cluster analysis identified four interpretable household profiles, although overlap between clusters remained substantial. These profiles show that disadvantages do not accumulate in a single uniform way, but combine in different configurations across households. The subsequent descriptive analysis, using socio-demographic, contextual, and health and wellbeing indicators not included in the clustering procedure, showed how these profiles differed across socio-demographic, contextual, and health and wellbeing indicators.

The accumulated housing and environmental vulnerability profile (Cluster 1) illustrates how housing disadvantage can extend beyond dwelling conditions to encompass broader social and environmental inequalities. The concentration of financial strain, less favourable self-rated health and wellbeing, and limited environmental resources within this profile suggests that some households experience multiple and overlapping forms of disadvantage (Dewilde 2022). This interpretation is consistent with evidence linking poor housing quality and overcrowding to adverse health and wellbeing outcomes (Aldridge et al. 2021; Eiffener et al. 2025; Wilson 2023), as well as research showing that

access to environmental amenities is socially patterned across European contexts (McKee et al. 2020).

The thermal and energy vulnerability profile (Cluster 2) highlights that housing disadvantage is not always expressed through visible dwelling problems. Instead, vulnerability may arise through heating difficulty, cooling discomfort, and challenges in managing energy-related costs. This finding aligns with previous research showing that energy poverty can affect households whose homes appear relatively adequate on conventional housing-quality measures, reflecting the interaction of energy prices, household resources, and thermal comfort needs (Liddell and Morris 2010; Bouzarovski et al. 2022; Menyhért 2024). More broadly, it illustrates how single composite deprivation measures may obscure distinct pathways of vulnerability (Ayala et al. 2022; Dudek and Wojewódzka-Wiewiórska 2024).

The resource-advantaged with thermal strain profile (Cluster 3) combined comparatively favourable dwelling conditions and stronger access to green space with continued thermal strain, particularly cooling discomfort. The cluster description further indicated that this profile was associated with relatively strong socio-economic resources. This is important because it suggests that thermal strain is not confined to households experiencing the most severe material disadvantage. Rather, it may reflect factors beyond income alone, including dwelling characteristics and access to resources for adapting to hot weather. This finding aligns with emerging research on “cooling inequality,” which shows that climate exposure, building form, and uneven access to adaptation measures can shape vulnerability in ways that are not reducible to income alone (Falchetta et al. 2024; Zahnow et al. 2025).

The dwelling problems and transport vulnerability profile (Cluster 4) was characterised by persistent dwelling problems, thermal strain, and comparatively weaker transport access. The cluster description indicated less favourable health and wellbeing indicators and a comparatively high prevalence of household chronic illness, despite the cluster occupying an intermediate socio-economic position. This profile indicates that comparatively stronger socio-economic resources can coexist with dwelling problems, thermal strain, and weaker transport access. The greater suburban and rural representation within this cluster may partly contribute to its weaker transport access,

although the analysis did not test whether residential context explained this pattern. This profile is consistent with research showing that unfavourable dwelling conditions can coincide with health vulnerability, particularly where chronic illness is present (Filandri and Olagnero 2014; Aldridge et al. 2021).

4.2. Co-occurrence of inequalities across domains

While the cluster analysis describes how vulnerabilities accumulate within households, the relative risk analysis examined whether inequalities across domains tend to co-occur. The results indicated that households experiencing dwelling vulnerability were more likely to face inequalities in access to green space. This pattern is consistent with research showing that environmental amenities such as green space are unevenly distributed and often linked to socio-economic and housing inequalities (McKee et al. 2020; Willberg et al. 2024). In this sense, the findings reinforce the argument that housing inequality should not be understood only in terms of conditions inside the home, but also in relation to access to the external environments that support health and wellbeing.

By contrast, the relationship between dwelling vulnerability and transport inequality was weaker and less consistent than the relationship observed for green-space inequality. Households with dwelling vulnerability were slightly less likely to experience transport inequality in the pooled sample, although country-specific associations were generally weak and inconsistent. This suggests that housing disadvantage does not necessarily coincide with poorer transport accessibility, and that the relationship may be more context-dependent than for green-space inequality. The descriptive cluster analysis provides one possible explanation for this pattern. Clusters 1 and 2, which showed greater dwelling vulnerability, had more urban profiles, whereas Clusters 3 and 4 had greater suburban and rural representation. Given that public transport provision is often stronger in urban settings, households experiencing poorer dwelling conditions may nevertheless have comparatively good transport access. Conversely, households living in more favourable dwelling conditions may reside in areas where public transport provision is more limited and car dependence is greater. However, urbanicity was not included in the clustering procedure and its explanatory role was not formally tested. These contextual differences may nevertheless help explain why the association between dwelling vulnerability and transport inequality was weaker and less consistent

than that observed for green-space inequality (Larsson et al. 2022; Carroll et al. 2021; Van Dülmen et al. 2022).

More broadly, the findings demonstrate that vulnerability is not confined to the most economically disadvantaged households. Some households experienced cumulative disadvantage across multiple domains, while others faced more specific vulnerabilities related to energy affordability, thermal strain, or environmental access. These patterns highlight the limitations of single-indicator approaches to housing inequality and reinforce calls for more integrated analytical frameworks (Guio et al. 2021; James et al. 2024).

4.3. Policy implications

The findings suggest that policies addressing housing inequality should consider the broader living environment in which housing is embedded. First, the finding that households often experience multiple dwelling-related vulnerabilities highlights the need for comprehensive housing assessments and integrated improvement measures, rather than interventions focused on a single problem in isolation. While well-designed retrofit measures may help reduce heating difficulty and some moisture-related problems, they may not resolve other vulnerabilities such as leaking roofs, inadequate space, or limited access to basic facilities.

Second, the positive association between dwelling vulnerability and green-space inequality suggests that access to green space should be considered as part of housing-related wellbeing rather than as a wholly separate issue. Policies aimed at improving living conditions may therefore benefit from promoting more equitable access to private, shared, or nearby green space.

Third, the weaker and less consistent relationship between dwelling and transport inequality suggests that transport disadvantage should not be assumed to follow the same pattern as housing or green-space disadvantage. Policy responses in this area are likely to require greater attention to local context, particularly differences between urban, suburban, and rural settings. Policies aimed at reducing transport inequalities should therefore consider the spatial distribution of public transport provision, patterns of car dependence, and barriers to mobility alongside housing conditions.

4.4. Limitations

First, this study was based on cross-sectional survey data and therefore cannot show how households move into or out of vulnerability over time, or how changes in energy prices, housing conditions, health, or household circumstances may affect their situation. Longitudinal data would provide a clearer understanding of how vulnerability trajectories emerge and evolve.

Second, the cluster analysis depends on methodological decisions regarding variable selection, coding, and the number of clusters retained. Housing-related indicators were more numerous than indicators from the green-space and transport domains, which may have increased the influence of housing conditions on the resulting cluster structure. Although the four-cluster solution was selected using the elbow criterion and interpretability, with PCA and UMAP used to support interpretation, the low Gower-based silhouette coefficient indicates that vulnerability patterns were distributed along overlapping continua rather than forming sharply separated groups. The clusters should therefore be interpreted as heuristic summaries of recurring vulnerability configurations rather than as discrete population categories.

Third, the generalisability of the findings is limited by the survey sampling strategy. Participants were recruited through an online non-probability panel, and the sample cannot therefore be considered nationally representative within each participating country. The final subsample was selected to approximate country-level Eurostat distributions for income, age, and gender, which improved demographic balance across these characteristics. The findings should therefore be interpreted as evidence of associations and recurring vulnerability configurations within a large cross-national sample adjusted to approximate population distributions for income, age, and gender, rather than as estimates of national prevalence (Fülöp et al. 2022).

Despite these limitations, the study identified recurring patterns involving dwelling conditions, thermal strain, environmental access, transport, and wider social disadvantage. By combining cluster analysis with relative risk estimation, it provides new evidence on how multiple forms of inequality intersect within households and offers a foundation for future research and policy aimed at addressing multidimensional vulnerability.

5. Conclusions

This study shows that housing inequality across the eight-country sample is multidimensional, with dwelling problems, thermal strain, environmental access, transport conditions, financial strain, and health and wellbeing differences intersecting in different ways across households. Dwelling vulnerability was more consistently associated with green-space inequality than with transport inequality, suggesting that different domains of disadvantage do not always overlap in the same way.

These findings point to the need for integrated housing policy responses. Comprehensive housing assessments should consider multiple dwelling-related vulnerabilities rather than focusing on single problems in isolation. At the same time, policies addressing housing inequality should engage with the wider living environment, particularly access to green space, while recognising that transport disadvantage may require more context-specific local responses.

6. Acknowledgements

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

7. Funding

AC, PC, SN, AK, IS, HJ, SP, and GMH were supported by the Horizon Program “HouseInc – Inclusive, affordable and sustainable housing for marginalized communities”. The HouseInc 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 it contains.

8. Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors used AI solely to proofread and improve the clarity of some English-language passages. The authors reviewed and edited all suggested changes and take full responsibility for the final content of the manuscript.

9. References

- Aldridge, Robert W., Helen Pineo, Ellen Fragaszy, et al. 2021. 'Household Overcrowding and Risk of SARS-CoV-2: Analysis of the Virus Watch Prospective Community Cohort Study in England and Wales'. *Wellcome Open Research* 6 (December): 347. <https://doi.org/10.12688/wellcomeopenres.17308.1>.
- Ayala, Luis, Elena Bárcena-Martín, Olga Cantó, and Carolina Navarro. 2022. 'COVID-19 Lockdown and Housing Deprivation across European Countries'. *Social Science & Medicine* 298 (April): 114839. <https://doi.org/10.1016/j.socscimed.2022.114839>.
- Begic, Hamza, Francesco Bruzzone, Silvio Nocera, and Constantinos Antoniou. 2026. 'Analysis of Socio-Economic Vulnerability to Transport Disruptions in the European Union'. *Journal of Urban Mobility* 9 (June): 100235. <https://doi.org/10.1016/j.urbmob.2026.100235>.
- Bentley, Rebecca, Kate Mason, David Jacobs, et al. 2025. 'Housing as a Social Determinant of Health: A Contemporary Framework'. *The Lancet Public Health* 10 (10): e855–64. [https://doi.org/10.1016/S2468-2667\(25\)00142-2](https://doi.org/10.1016/S2468-2667(25)00142-2).
- Bhattacharyya, Helen. 2014. 'Kruskal–Wallis Test: Theory'. *Wiley StatsRef: Statistics Reference Online*, ahead of print. <https://doi.org/10.1002/9781118445112>.
- Born, Theo, Jenny Preece, and David Robinson. 2025. 'Housing Inequalities: Definitions, Understandings and Approaches'. *Housing, Theory and Society*, September 7, 1–20. <https://doi.org/10.1080/14036096.2025.2556704>.
- Bouzarovski, Stefan, Manon Burbidge, Amish Sarpotdar, and Mari Martiskainen. 2022. 'The Diversity Penalty: Domestic Energy Injustice and Ethnic Minorities in the United Kingdom'. *Energy Research & Social Science* 91 (September): 102716. <https://doi.org/10.1016/j.erss.2022.102716>.
- Browning, Matthew H. E. M., Alessandro Rigolon, Olivia McAnirlin, and Hyunseo (Violet) Yoon. 2022. 'Where Greenspace Matters Most: A Systematic Review of Urbanicity, Greenspace, and Physical Health'. *Landscape and Urban Planning* 217 (January): 104233. <https://doi.org/10.1016/j.landurbplan.2021.104233>.
- Cardinali, Marcel, Mariëlle A. Beenackers, Arjan Van Timmeren, and Uta Pottgiesser. 2024. 'The Relation between Proximity to and Characteristics of Green Spaces to Physical Activity and Health: A Multi-Dimensional Sensitivity Analysis in Four

- European Cities'. *Environmental Research* 241 (January): 117605. <https://doi.org/10.1016/j.envres.2023.117605>.
- Carmona-Derqui, Demetrio, Jonathan Torres-Tellez, and Alberto Montero-Soler. 2023. 'Effects of Housing Deprivation on Health: Empirical Evidence from Spain'. *International Journal of Environmental Research and Public Health* 20 (3): 2405. <https://doi.org/10.3390/ijerph20032405>.
- Conzatti, Anna, and Gesche Huebner. 2025. 'Identifying Patterns of Housing-Related Vulnerabilities Through Cluster Analysis in Eight European Countries'. Preprint, OSF, May 24. <https://doi.org/doi:10.17605/OSF.IO/P3BD2>.
- Council of Europe Development Bank. 2017. *Housing Inequality in Europe Tackling Inequalities in Europe: The Role of Social Investment*.
- Czerniak, Adam, and Patrycja Graca-Gelert. 2024. 'Varieties of Housing Inequality Patterns in Europe'. Paper presented at ERES: Conference. *30th Annual European Real Estate Society Conference*.
- Dewilde, Caroline. 2022. 'How Housing Affects the Association between Low Income and Living Conditions-Deprivation across Europe'. *Socio-Economic Review* 20 (1): 373–400. <https://doi.org/10.1093/ser/mwab003>.
- Dudek, Hanna, and Agnieszka Wojewódzka-Wiewiórska. 2024. 'Housing Deprivation Among Polish Households: Prevalence and Associated Factors'. *Real Estate Management and Valuation* 32 (2): 58–69. <https://doi.org/10.2478/remav-2024-0015>.
- Eiffener, Elodie, Rachel Murekatete, Anne-Sophie Merritt, et al. 2025. 'Crowded Housing, Indoor Environment and Children's Respiratory, Allergic and General Health in Sweden: A Cross-Sectional Study'. *BMJ Open* 15 (9): e106117. <https://doi.org/10.1136/bmjopen-2025-106117>.
- Endalew Terefe, Alene, and Yuting Hou. 2024. 'Determinants Influencing the Accessibility and Use of Urban Green Spaces: A Review of Empirical Evidence'. *City and Environment Interactions* 24 (December): 100159. <https://doi.org/10.1016/j.cacint.2024.100159>.
- European Parliament. Directorate General for Parliamentary Research Services. 2020. *Addressing Health Inequalities in the European Union: Concepts, Action, State of Play: In Depth Analysis*. Publications Office. <https://data.europa.eu/doi/10.2861/567478>.
- Faka, Antigoni. 2025. 'Mapping Green Space Inequalities in 26 European Cities'. *Land* 14 (12): 2362. <https://doi.org/10.3390/land14122362>.
- Falchetta, Giacomo, Enrica De Cian, Filippo Pavanello, and Ian Sue Wing. 2024. 'Inequalities in Global Residential Cooling Energy Use to 2050'. *Nature Communications* 15 (1): 7874. <https://doi.org/10.1038/s41467-024-52028-8>.

- Filandri, Marianna, and Manuela Olagnero. 2014. 'Housing Inequality and Social Class in Europe'. *Housing Studies* 29 (7): 977–93. <https://doi.org/10.1080/02673037.2014.925096>.
- Frank, Lawrence D., Nicole Iroz-Elardo, Kara E. MacLeod, and Andy Hong. 2019. 'Pathways from Built Environment to Health: A Conceptual Framework Linking Behavior and Exposure-Based Impacts'. *Journal of Transport & Health* 12 (March): 319–35. <https://doi.org/10.1016/j.jth.2018.11.008>.
- Fülöp, Gabor, Gabor Fulop, Francesco Avvisati, and Francesco Avvisati. 2022. 'The Analytical Value of Non-Probability Samples in the Context of TALIS: A Review of Current Practices in the Use of Non-Probability Samples in Comparative, Cross-National Research'. *OECD Education Working Paper No. 272*.
- Grotti, Raffaele, Helen Russell, Bertrand Maître, and Davide Gritti. 2024. 'The Experience of Housing Discrimination and Housing Deprivation Across Social Groups in Ireland'. *Social Indicators Research* 175 (1): 195–215. <https://doi.org/10.1007/s11205-024-03438-0>.
- Guio, Anne-Catherine, Eric Marlier, and Brian Nolan, eds. 2021. *Improving the Understanding of Poverty and Social Exclusion in Europe*. 2021 ed. Publications Office of the European Union.
- Hick, Rod, Marco Pomati, and Mark Stephens. 2022. 'Severe Housing Deprivation in the European Union: A Joint Analysis of Measurement and Theory'. *Social Indicators Research* 164 (3): 1271–95. <https://doi.org/10.1007/s11205-022-02987-6>.
- Holmes, Hannah, Katy Karampour, Robert Richardson, and Gemma Burgess. 2025. 'Housing, Poverty, and Digital Exclusion: The Multiple Ways in Which Digital Inequalities and Housing Inequalities Intersect'. *Housing Studies*, October 28, 1–21. <https://doi.org/10.1080/02673037.2025.2577727>.
- Huang, Zhexue. 1998. 'Extensions to the K-Means Algorithm for Clustering Large Data Sets with Categorical Values'. *Data Mining and Knowledge Discovery* 2: 283–304.
- James, Laura, Lyrian Daniel, Rebecca Bentley, and Emma Baker. 2024. 'Housing Inequality: A Systematic Scoping Review'. *Housing Studies* 39 (5): 1264–85. <https://doi.org/10.1080/02673037.2022.2119211>.
- Katz, D., J. Baptista, S. P. Azen, and M. C. Pike. 1978. 'Obtaining Confidence Intervals for the Risk Ratio in Cohort Studies'. *Biometrics* 34 (3): 469. <https://doi.org/10.2307/2530610>.
- Li, Chenlu, Xiaoxu Wu, Xiao Cheng, et al. 2019. 'Identification and Analysis of Vulnerable Populations for Malaria Based on K-Prototypes Clustering'. *Environmental Research* 176 (September): 108568. <https://doi.org/10.1016/j.envres.2019.108568>.

- Liddell, Christine, and Chris Morris. 2010. 'Fuel Poverty and Human Health: A Review of Recent Evidence'. *Energy Policy* 38 (6): 2987–97. <https://doi.org/10.1016/j.enpol.2010.01.037>.
- Mandic, Srna, and Andreja Cirman. 2012. 'Housing Conditions and Their Structural Determinants: Comparisons within the Enlarged EU'. *Urban Studies* 49 (4): 777–93. <https://doi.org/10.1177/0042098011405688>.
- Matel, Anna. 2025. 'Changing Profiles of Housing Deprivation in European Union Countries'. *Real Estate Management and Valuation* 33 (1): 47–60. <https://doi.org/10.2478/remav-2025-0005>.
- McKee, Kim, Adriana Mihaela Soaita, and Jennifer Hoolachan. 2020. "“Generation Rent” and the Emotions of Private Renting: Self-Worth, Status and Insecurity amongst Low-Income Renters'. *Housing Studies* 35 (8): 1468–87. <https://doi.org/10.1080/02673037.2019.1676400>.
- Menyhért, Bálint. 2024. 'Energy Poverty in the European Union. The Art of Kaleidoscopic Measurement'. *Energy Policy* 190 (July): 114160. <https://doi.org/10.1016/j.enpol.2024.114160>.
- Nelson, Ruth, Martijn Warnier, and Trivik Verma. 2024. 'Housing Inequalities: The Space-Time Geography of Housing Policies'. *Cities* 145 (February): 104727. <https://doi.org/10.1016/j.cities.2023.104727>.
- Norris, Michelle, and Patrick Shiels. 2007. 'Housing Inequalities in an Enlarged European Union: Patterns, Drivers, Implications'. *Journal of European Social Policy* 17 (1): 65–76. <https://doi.org/10.1177/0958928707071893>.
- Smith, Melody, Jamie Hosking, Alistair Woodward, et al. 2017. 'Systematic Literature Review of Built Environment Effects on Physical Activity and Active Transport – an Update and New Findings on Health Equity'. *International Journal of Behavioral Nutrition and Physical Activity* 14 (1): 158. <https://doi.org/10.1186/s12966-017-0613-9>.
- Tayefi Nasrabadi, Mahla, Taimaz Larimian, Andrew Timmis, and Tan Yigitcanlar. 2024. 'Mapping Four Decades of Housing Inequality Research: Trends, Insights, Knowledge Gaps, and Research Directions'. *Sustainable Cities and Society* 113 (October): 105693. <https://doi.org/10.1016/j.scs.2024.105693>.
- Tayefi Nasrabadi, Mahla, Taimaz Larimian, and Tan Yigitcanlar. 2025. 'Developing a Measurement Model to Assess Perceived Housing Inequality: Exploratory and Confirmatory Factor Analyses'. *Cities* 166 (November): 106276. <https://doi.org/10.1016/j.cities.2025.106276>.
- United Nations General Assembly. 1948. 'Universal Declaration of Human Rights (A/RES/217 A (III))'. December 10. <https://digitallibrary.un.org/record/666853>.

- Walter, Anita, Alice Wanner, and Ulrike Pröbstl-Haider. 2025. 'Clustering Comfort: A Cluster Analysis on Housing Conditions and Nature-Based Solutions in Polish Cities'. *Land* 14 (9): 1884. <https://doi.org/10.3390/land14091884>.
- Willberg, Elias, Christoph Fink, Robert Klein, Roope Heinonen, and Tuuli Toivonen. 2024. "Green or Short: Choose One" - A Comparison of Walking Accessibility and Greenery in 43 European Cities'. *Computers, Environment and Urban Systems* 113 (October): 102168. <https://doi.org/10.1016/j.compenvurbsys.2024.102168>.
- Wilson, Wendy. 2023. *Overcrowded Housing (England)*. House of Commons.
- Wolff, Manuel, Sebastian Scheuer, and Dagmar Haase. 2020. 'Looking beyond Boundaries: Revisiting the Rural-Urban Interface of Green Space Accessibility in Europe'. *Ecological Indicators* 113 (June): 106245. <https://doi.org/10.1016/j.ecolind.2020.106245>.
- Zahnow, Renee, Ali Rad Yousefnia, Mahnoosh Hassankhani, and Ali Cheshmehzangi. 2025. 'Climate Change Inequalities: A Systematic Review of Disparities in Access to Mitigation and Adaptation Measures'. *Environmental Science & Policy* 165 (March): 104021. <https://doi.org/10.1016/j.envsci.2025.104021>.

10. Appendix

The appendices provide supplementary information supporting the analyses presented in the main text. They include details of the data pre-processing and variable recoding procedures, statistical tests used to assess differences across clusters, and detailed cluster profiles that complement the summary results reported in the main manuscript.

10.1. Appendix A

This appendix documents the specific recoding procedures applied to variables with sparse categories (fewer than 25 responses per category). Categories were merged either to maintain statistical robustness or to align with the conceptual framing of the study.

Table A 1 Variables that were merged or combined during pre-processing.

Variable	Reason for merging	How categories were merged	Final variable used in analysis
EDU (Education)	Some education categories had fewer than 25 responses	Rare categories merged into “Other”	Education variable with “Other”
EMP (Employment)	Some employment categories had fewer than 25 responses	Rare categories merged into “Other”	Employment variable with “Other”
DISC_SELF* (Perceived discrimination items)	Multi-response format; needed binary indicators	Each item converted to a separate 0/1 flag; “Prefer not to say” recoded to 0	Ten binary indicators + composite variable ANY_DISC
PUBLIC_TRANSP_TRAIN and PUB_TRANS_TRAM_METRO	Low frequency when separated	Train and tram/metro access merged into single category	Train/Tram/Metro combined variable
GREEN_SPACE_1 and GREEN_SPACE_3	Both represent <i>private</i> outdoor green space	Presence in either category counted as private green access	Private green space indicator
GREEN_SPACE_2 and GREEN_SPACE_4	Both represent <i>shared</i> green space	Presence in either category counted as shared green access	Shared green space indicator

10.2. Appendix B

This appendix reports the statistical tests used to assess differences across the four clusters. Chi-square tests were used for categorical variables, while Kruskal–Wallis tests were used for continuous variables. False discovery rate (FDR) adjustment was applied to account for multiple testing.

Table B 1 Results of Chi-square tests assessing differences across the four clusters for all categorical variables included in the descriptive analysis.

Variable	Chi ²	df	p-value (raw)	p-value (FDR-adjusted)
HHTYPE	974.002	15	< 0.0001	< 0.0001
FIN_COMF	673.287	12	< 0.0001	< 0.0001
GREEN_SPACE_1	459.15	3	< 0.0001	< 0.0001
DWELLING	500	15	< 0.0001	< 0.0001
FINBURDEN_ECOST	450.981	12	< 0.0001	< 0.0001
OWNERSHIP	395.884	12	< 0.0001	< 0.0001

GREEN_SPACE_3	285.736	3	< 0.0001	< 0.0001
AFF_HEAT	256.917	3	< 0.0001	< 0.0001
COUNTRY_R	318.69	21	< 0.0001	< 0.0001
COMF_COOL	252.602	3	< 0.0001	< 0.0001
EMP	301.148	27	< 0.0001	< 0.0001
GREEN_SPACE_5	214.818	3	< 0.0001	< 0.0001
GREEN_SPACE_6	189.828	3	< 0.0001	< 0.0001
LIVING_AREA	194.392	6	< 0.0001	< 0.0001
DWELL_PROB_2	179.174	3	< 0.0001	< 0.0001
GEN_HEALTH	210.595	12	< 0.0001	< 0.0001
AGE_GROUP_LABEL	176.772	6	< 0.0001	< 0.0001
PUB_TRANS_TRAIN_TRAM	159.454	3	< 0.0001	< 0.0001
AIRCO	159.07	3	< 0.0001	< 0.0001
DWELL_PROB_4	146.114	3	< 0.0001	< 0.0001
SPACE_SHORT	135.021	3	< 0.0001	< 0.0001
DWELL_PROB_5	105.395	3	< 0.0001	< 0.0001
CHRON_ILL	122.59	12	< 0.0001	< 0.0001
DWELL_PROB_3	93.35	3	< 0.0001	< 0.0001
DISC_SELF_8	100.903	6	< 0.0001	< 0.0001
EDU	113.427	18	< 0.0001	< 0.0001
DWELL_PROB_1	72.32	3	< 0.0001	< 0.0001
ANY_DISC	68.581	3	< 0.0001	< 0.0001
DISC_SELF_6	48.861	6	< 0.0001	< 0.0001
DISC_SELF_10	42.88	6	< 0.0001	< 0.0001

DISC_SELF_4	38.663	6	< 0.0001	< 0.0001
PUB_TRANS_NONE	29.958	3	< 0.0001	< 0.0001
DISC_SELF_9	37.156	6	< 0.0001	< 0.0001
DISC_SELF_2	35.547	6	< 0.0001	< 0.0001
DWELL_PROB_6	23.962	3	0.00003	0.00003
GS_SHARED_ONLY	23.469	3	0.00003	0.00004
GREEN_SPACE_2	20.658	3	0.00012	0.00016
DISC_SELF_7	24.462	6	0.00043	0.00053
DISC_SELF_1	24.154	6	0.00049	0.00059
PUB_TRANS_BUS	17.058	3	0.00069	0.00081
DISC_SELF_3	21.316	6	0.00161	0.00185
DISC_SELF_5	18.077	6	0.00604	0.00676
PUB_TRANS_OTHER	8.803	3	0.03203	0.03501
GENDER	8.467	3	0.03729	0.03983
ETH	10.641	6	0.10011	0.10456
GREEN_SPACE_4	2.745	3	0.43269	0.44210
COUNTRY_B	4.464	6	0.61418	0.61418

Table B 2 Results of Kruskal–Wallis tests examining differences in continuous variables across the four clusters.

Variable	H	df	p-value (raw)	p-value (FDR-adjusted)
EPC_10	4707.32	3	< 0.0001	< 0.0001
CARS_OWN	5676.565	3	< 0.0001	< 0.0001

SPACE_PP	315.711	3	< 0.0001	< 0.0001
ROOM_PP	247.886	3	< 0.0001	< 0.0001

10.3. Appendix C

This appendix presents the tables describing the cluster profiles. Table C 1 reports the distribution of variables used in the clustering procedure and therefore describes the dimensions defining the cluster structure. Table C 2 to Table C 4 present the cluster profiles based on variables used for cluster description only. These descriptive variables were not included in the clustering procedure but are reported to show how the cluster profiles differ across socio-demographic, contextual, and health and wellbeing characteristics.

Table C 1 Cluster profiles based on variables included in the clustering procedure. Values indicate the percentage of households within each cluster reporting each characteristic.

Variable group	Vulnerability indicator	Cluster 1 (%)	Cluster 2 (%)	Cluster 3 (%)	Cluster 4 (%)
Space	Space shortage	33.8	22.5	17.3	27.4
Dwelling problems	Leaking roof	7	2.6	2.5	6.3
	Damp walls/floors/foundation	16.2	6.2	6.1	15.4
	Rot in window frames or floor	7.4	2.1	2.5	5.1
	Mould around windows/bathroom	19.9	9.2	11.6	19.9
	Mould on any walls	12.6	5.8	5.1	12.3
	Mould on furnishings	3.9	1.9	2.1	3.8
Thermal strain	Difficulty affording adequate heating	81.1	97.4	94.4	91
	Difficulty keeping comfortably cool in summer	56.2	85.1	73.3	69.3
Green space	Private garden (access)	39.2	43.9	71.9	64.6
	Private balcony/patio/roof terrace (access)	38.8	61.2	58.7	45.8
	Shared green space	16.3	13.7	10.3	14.2
	Accessible green space nearby (access)	37.2	57.8	48.7	44.7

	None of the above	10.6	3.3	1.6	3.3
Transport	Bus (access)	93.5	94.9	92.7	91.5
	Train (access)	58.4	69.9	63.2	49.1
	No public transport	3	3	4.5	6.3
Energy appliances	No heating	12.2	5.7	4.7	8.2
	No cooling	25.2	15.1	13.6	16.7
	No hot water	6.5	3.6	3.7	4.6
	No cooking appliances	5.2	2.5	2.5	2.8
	No other kitchen appliances	4.6	2.6	2.8	3.5
	No fridge	5.1	2.7	2.7	2.9
	No freezer	9.6	5.1	4.7	5.8
	No lighting	3.5	2.2	2.4	2.8
	No television	5.9	3.6	2.5	3.8
	No telephone	6	3.4	3.4	3.3
	No washing machine	7.5	3.8	2.7	3.6
	No tumble dryer	31.7	20.6	18	18.9

Table C 2 Socio-demographic cluster profiles based on descriptive variables not included in the clustering procedure. Values are percentages unless otherwise indicated.

Variable group	Vulnerability indicator	Cluster 1 (%)	Cluster 2 (%)	Cluster 3 (%)	Cluster 4 (%)
Age group	Adults (20-64)	70.1	65.5	80.7	81.2
	Older adults (65+)	29	33.6	17.5	16.9
	Teens (14-19)	0.9	1	1.8	1.8
Gender	Female	52.2	50	48.6	53.9
	Male	47.8	50	51.4	46.1
Education	Bachelors	1.6	21.1	25.3	19.1
	Doctorate	10.6	2	3.7	1.7

Employment	Lower Secondary	13.1	8.6	7.1	9.8
	Masters	3.4	17.3	21.6	16.4
	Primary	27.6	3.1	1.8	2.3
	Upper Secondary	23.8	24.6	20.6	28.3
	Vocational	1.6	23.3	19.7	22.3
	Employed Full Time	40.5	41.8	56.6	51.2
	Employed Part Time	7.5	8.1	10.1	11.3
	Household Work	4.3	3	2	4.1
	Other	0.6	0.5	0.3	0.5
	Retired	26.9	31.4	16.9	16.6
	Self Employed	5.1	4.4	5.5	5.4
	Student	3.9	3.6	5.4	5.2
	Temporary Leave	1.3	0.9	1	1.6
	Unable to Work	3.9	2.5	0.5	1.6
	Unemployed	6.1	4	1.6	2.6
Household type	Couple No Children	27.8	33.2	33.6	33.4
	Couple With Children	23.2	24.4	51.1	48.4
	Living With Parents	6	4.9	8.3	10.2
	Non-Family Household	2	1.2	1.2	1
	Single	33.8	31	3.7	4.5
	Single Parent	7.2	5.3	2.1	2.5
Country of residence	BE	6.5	9.6	12.6	8.4
	CZ	9.6	8.6	5.1	8.6
	DE	13.8	15.8	17.3	15.2
	EE	14.1	10.1	8.8	10.7
	FI	10.9	19.8	12.1	12.2
	GB	17.9	17.2	20.8	17.7
	IT	14	9.4	17.4	20.8

	RO	13.2	9.6	5.8	6.4
Country of origin same as country of residence	No	6.2	5.9	5.1	5.2
	No Answer	0.1	0	0	0
	Yes	93.7	94	94.9	94.8
Ethnicity same as country of residence	Don't Know	3.5	2.3	2.6	2.4
	No	7	5.8	6.4	6.2
	Yes	89.5	91.9	90.9	91.4
Discrimination	No	58.3	69.3	66.6	62.8
	Yes	41.7	30.7	33.4	37.2
Financial comfort	Finding it very difficult to live on current income	13.2	5.6	1.7	5.1
	Finding it difficult to live on current income	28.8	16.5	9.4	18.3
	Coping on current income	40.8	46.9	38.2	45.7
	Living comfortably on current income	14.4	26.6	42.4	27.9
	Living very comfortably on current income	2.8	4.3	8.3	3
Ownership	Owned, No Mortgage	41.1	45.7	49.5	43.6
	Owned, With Mortgage	16.6	20.5	34.8	31.1
	Rented: Private	9.7	9.7	3.1	3.8
	Free Residence	29.2	22.2	10	17.4
Mobility resources	Car [avg]	1	1	2.14	2.17

Table C 3 Contextual cluster profile based on descriptive variables not used in the clustering procedure.

Variable group	Vulnerability indicator	Cluster 1 (%)	Cluster 2 (%)	Cluster 3 (%)	Cluster 4 (%)
Air conditioning	No	78.8	72.9	59.1	67.5
Living area	City	51.1	52	36.9	34.5
	Rural	16.4	14.1	23.9	27.4

	Town suburb	32.5	33.9	39.2	38.2
Dwelling type	Detached House	21.2	22.7	42.7	37.9
	Flat: 10+ Dwellings	26.7	30.1	11.2	14
	Flat:1-2 Dwellings	11.6	7.9	3.9	5.4
	Flat:3-9 Dwellings	18.4	15.7	12.2	15.3
	Other	1.2	0.7	0.3	0.2
	Semi Terraced House	20.9	22.8	29.6	27.3

Table C 4 Health cluster profile based on descriptive variables not used in the clustering procedure.

Variable group	Vulnerability indicator	Cluster 1 (%)	Cluster 2 (%)	Cluster 3 (%)	Cluster 4 (%)
General health	Bad	10.6	7.2	2.7	5.9
	Very bad	0	0	0	0
Chronic illness	Yes	48.7	44	36.5	48.7
	Yes, household	17.2	14.6	17.2	23.9
Wellbeing ladder	Wellbeing [avg]	6.91	7.49	7.83	7.39