energies

MDPI

Article

Residential Heating Transitions in Hungary: Grid Profiles of Fuel Use and Newer Energy Equipment

Geza Toth [1,2,*] and Tekla Labelle Szép [1]

[1] Institute of World and Regional Economics, Faculty of Economics, University of Miskolc, 3515 Miskolc-Egyetemváros, Hungary; tekla.szep@uni-miskolc.hu
[2] Hungarian Central Statistical Office, Keleti Károly Utca 5–7, 1024 Budapest, Hungary
* Correspondence: geza.toth@ksh.hu

Abstract

We examine traditional heating fuel use and ownership of newer energy equipment in Hungary using 2022 census data. The traditional fuel-use index (OLD) combines reported use of wood, coal or lignite and fuels in the census "other fuels" category. The newer equipment ownership index (NEW) combines reported ownership of heat pumps, solar thermal collectors and photovoltaic panels. K-means clustering of 30,237 unique 1 km grid cells identifies three profiles distinguished mainly by traditional fuel use and a fourth with high values for both indices. This fourth profile accounts for 2.62% of cells and has a median of two occupied dwellings. It does not form a separate cluster when analysis is restricted to cells with at least six occupied dwellings. Global Moran's I is higher for traditional fuel use across all four neighbourhood definitions, while local bivariate categories vary with spatial weights. In the 17,146 cells with positive NEW in the mapping data, traditional fuel use shows stronger residual spatial dependence across all five regression weight specifications. Joint spatial regression finds positively correlated residual innovations. The results establish spatial co-occurrence, while evidence of household heating transitions requires information on actual equipment use.

Keywords: energy ladder; residential heating; fuel stacking; K-means; bivariate LISA; spatial error model; Hungary

1. Introduction

Households can retain one fuel or appliance after installing another. The classical energy ladder and the fuel-stacking literature provide complementary starting points for examining when households switch fuels or expand their options [1–4]. In Europe, households may retain fuelwood because it is affordable, locally available or embedded in daily routines [5].

Heating choices are also spatially structured. Nearby dwellings can share building characteristics, fuel supply networks and access to installers. Research on Hungarian municipalities documents spatial relationships in household energy practices and technology adoption [6]. Fine-scale analysis can show where reported fuels and equipment occur together and where their spatial patterns diverge.

We examine two overlapping indicators, reported use of selected fuels (OLD) and ownership of newer energy equipment (NEW). Piped natural gas and bottled LPG remain relevant to the heating transition but fall outside both indices. Heating oil can contribute to OLD because the census records it within a residual response category that also includes

Check for updates

Academic Editor: Andrea Frazzica

Received: 21 October 2026
Revised: 18 November 2026
Accepted: 20 November 2026
Published: 23 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

some biomass fuels [7]. The indices therefore capture selected dimensions of the heating mix. Low OLD alone gives no indication of whether an area relies on gas, district heat, electricity or another source.

We ask which grid profiles emerge from the two indicators, how their neighbourhoods differ, and how housing and education are associated with each indicator. The typology, global Moran's I and local spatial analyses use 30,237 unique mapped cells. The main regressions use the original regression sample of 17,146 cells with a positive NEW value in the mapping file. We also repeat the 6 km spatial error models on all 30,237 unique mapped cells to assess sensitivity to the sample restriction.

High values of OLD and NEW in the same grid are evidence of area-level co-occurrence. Establishing household fuel stacking requires information on the joint use of heating sources. The spatial error coefficient describes residual spatial dependence conditional on the chosen weights and covariates.

2. Fuel Stacking and the Geography of Residential Heating

2.1. From the Energy Ladder to Fuel Portfolios

The historical energy ladder describes a tendency for households to shift from traditional biomass towards more convenient fuels as income and access improve [1,2]. Even within that framework, uptake depends on infrastructure and equipment costs. Fuel stacking focuses on households' use of several fuels or appliances over the same period, including for the same end use [3,4]. Improved access may change which appliance is used most often without eliminating the others. A multidimensional energy profile distinguishes services, fuels and technologies and accommodates both substitution and continued use [8]. We use the ladder as historical context and the fuel-portfolio perspective to interpret the persistence and co-occurrence observed in the grid data.

2.2. Space Heating and Domestic Hot Water

Much of the early evidence for fuel stacking comes from cooking [3]. Evidence specific to space heating comes from household energy surveys in rural China, where Zhu et al. documented mixed fuel use and reported different transition rates for heating and cooking [9]. For Europe, research on fuelwood documents its role in coping with heating costs and its persistence as a cultural practice [5]. In Hungary, census-based research has documented spatial variation in residential heating fuels and equipment [10]. Together these studies motivate an analysis of local context, housing and equipment without assuming that an installation displaces an existing fuel.

Space heating and domestic hot water can share one system or be supplied by separate systems. Hybrid heat-pump assessments describe arrangements that retain a boiler alongside the heat pump [11]. For domestic hot water, Martinopoulos evaluated solar thermal systems across European locations and found that the modelled share of hot-water demand supplied by solar energy varied by location [12]. A solar collector may thus contribute to the hot-water supply while a different fuel continues to provide space heating. Eurostat reports space and water heating separately in its household energy statistics [13].

2.3. Energy Services, Poverty and the Built Environment

An energy-service perspective starts from outcomes such as adequate indoor warmth and hot water [14]. Continued fuelwood use can reflect affordability, reliability, the condition of the building or household preference [5]. Analysing these mechanisms would require direct information on costs, building performance and occupants' practices. The present study observes the geographic distribution of reported fuel use and equipment ownership.

2.4. Spatial Context and Technology Adoption

Neighbouring areas can share infrastructure, building stock, suppliers and information about new technologies [6,10]. Spatial association can arise from these common conditions or from diffusion between locations. The study first examines the spatial concentration of profiles defined by a cell's two indicators. It then uses bivariate local Moran's I to relate a cell's OLD value to NEW in neighbouring cells [15,16].

3. Data and Methods

The study uses data from the Hungarian Central Statistical Office's 2022 Population and Housing Census, with a reference date of 1 October 2022, aggregated to a 1 km grid [17,18]. The official database and housing questionnaire document the census results and collection instrument, respectively [7,17]. The initial mapping file contained 47,261 records. We retained one record per 1 km grid identifier because repeated records carried identical OLD and NEW values. After deduplication, the mapping file contains 30,237 unique mapped cells. Using the same grid identifier, we linked the original regression sample of 17,146 cells to the mapping file and the parent export of 30,897 unique grid cells. Section 5.1 describes how the samples overlap.

3.1. Variables and Preprocessing

OLD is the number of reported uses of wood-based fuels, coal or lignite, and fuels reported in the census residual category, per 100 occupied dwellings. Question 12 records piped natural gas, bottled LPG and electric heating separately, so those responses do not enter OLD. Heating oil contributes to OLD if reported in the residual category, which also includes examples such as straw and other plant material. The census provides no separate heating-oil count within that response. District heating is recorded in Question 11 and its respondents skip the fuel question. Respondents using more than two fuels are asked to report the two they use most [7].

NEW pools, dwelling counts for heat pumps, photovoltaic systems and solar thermal equipment, are reported in Question 13, per 100 occupied dwellings. This question allows multiple responses. NEW records equipment availability, including photovoltaic generation, and it does not identify actual heat provision or distinguish space heating from hot water. Electricity reported as a heating fuel does not itself enter NEW. The two indices may overlap and do not exhaust the heating mix. Because their numerators pool component counts, the values are composite indices of reported use or ownership. We capped both indices at 100 in the mapping and regression data, including in the full-sample sensitivity analysis. The clustering used these values without standardisation.

An occupied dwelling may report wood and piped gas as its two most-used heating fuels. In that example, the wood report contributes to OLD and the gas report contributes to neither index. If the dwelling also reports a heat pump, that equipment contributes to NEW regardless of whether it is used as the main heating system. Thus, the groupings are based on the census variables and do not assign gas or LPG to an intermediate stage of a transition [7].

3.2. K-Means Clustering and Selection of the Number of Clusters

The typology was constructed using K-means with Euclidean distance. We compared two to eight clusters using the silhouette coefficient, Davies–Bouldin index, Calinski–Harabasz index and within-cluster sum of squares (WCSS). The silhouette was estimated on a fixed random subset of 10,000 cells. The other diagnostics used the full sample. The three-cluster solution had the highest silhouette (0.575 versus 0.545 for four clusters). The four-cluster solution had the lowest Davies–Bouldin index (0.638) and reduced WCSS

by 37.1% relative to three clusters. It retained a separately identifiable high-NEW profile alongside three levels of OLD.

The four-cluster reference solution used k-means++ with 20 initialisations and random seed 42. No spatial constraints were entered into the clustering. We separately assess the influence of the dwelling denominator through the distribution of occupied dwelling counts, dwelling-weighted index means, and refits restricted to cells with at least six and at least twenty occupied dwellings.

3.3. Stability Assessment

We repeated the four-cluster fit with 12 random seeds on the full dataset and on 12 random 70% subsamples. We calculated the adjusted Rand index (ARI) against reference labels for the same cells. We matched centroids by minimum total Euclidean distance and recorded their displacements in index points.

3.4. Spatial Weights and Neighbourhood-Based Cluster Assessment

We used Euclidean distance between grid centres. The baseline used binary links at distances of up to 6 km, with 1,402,398 directed links, a mean of 46.38 neighbours and a range of 1 to 112. A 4 × 4 transition matrix counts directed neighbour links from each profile to the four profiles. We compare the fraction of links to the same profile with the profile's share of all cells. The sensitivity analysis uses 3 km, 9 km and a symmetrised eight-nearest-neighbour graph. The latter retains a link whenever either cell selects the other among its eight nearest neighbours, with distance ties resolved by the order returned by the nearest-neighbour search.

3.5. Global and Bivariate Local Spatial Association

Global Moran's I uses row-standardised weights and is calculated separately for OLD and NEW. The 6 km baseline uses 999 global permutations. Bivariate local Moran's I relates focal OLD to the spatial lag of NEW. Local categories use 999 conditional permutations and a two-sided pointwise pseudo-p threshold of 0.05. The two-sided probability doubles the smaller tail probability, with tied values counted in both tails. The maps are exploratory and the local probabilities are unadjusted for multiple comparisons. Cells without neighbours receive no local classification. Section 4.5 defines the four configurations.

4. Results

4.1. The Four-Cluster Heating Typology

The first three profiles primarily distinguish levels of OLD. C1 has low OLD and modest NEW, C2 has intermediate OLD and low NEW, and C3 combines high OLD with the lowest NEW centroid. C4 contains 793 cells (2.62%) and combines an OLD centroid of 94.01 with a NEW centroid of 77.56. Table 1 summarises the sizes and centroids of the four clusters. Figure 1 shows their spatial distribution.

Table 1. Characteristics of the four-cluster solution.

Cluster	Profile	Cells	Cells (%)	OLD Mean	NEW Mean
C1	Low OLD	4203	13.9	25.26	14.2
C2	Intermediate OLD	10,574	34.97	63.61	5.76
C3	High OLD	14,667	48.51	95.08	3.18
C4	High OLD and NEW	793	2.62	94.01	77.56

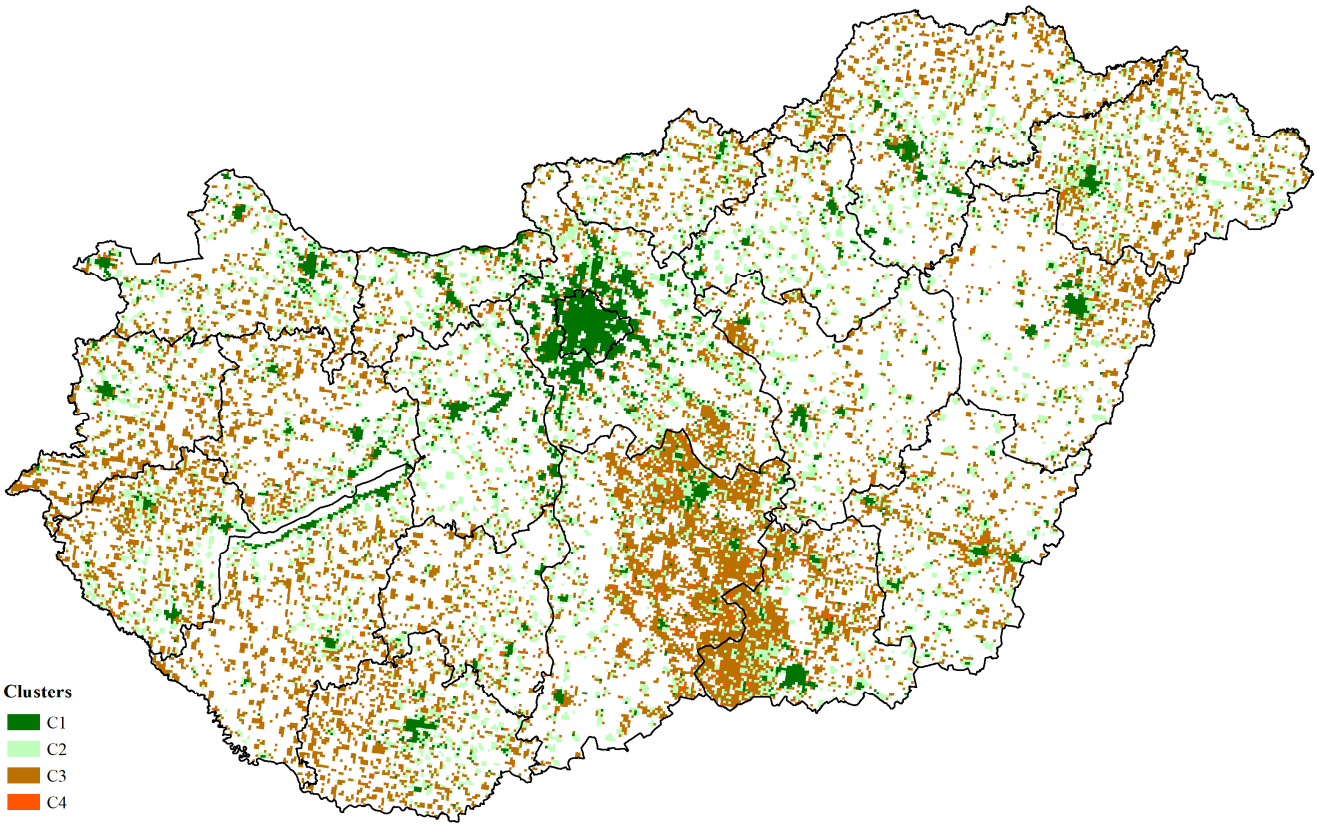

Figure 1. Distribution of the four heating and equipment profiles across unique 1 km grid cells.

C4 cells have a median of two occupied dwellings (interquartile range 1–3), compared with 22 across all cells. Of the 793 C4 cells, 691 (87.1%) have at most five dwellings and 312 (39.3%) have only one. Both indicators are positive in all 312 one-dwelling C4 cells. In those cells the records refer to the same occupied dwelling and indicate traditional fuel use alongside ownership of at least one item of newer equipment. In multi-dwelling cells the amounts attributable to the same dwelling cannot be determined from the aggregate counts. The data also do not identify which equipment drives the C4 centroid or how the equipment is used.

4.2. Stability of the Four-Cluster Solution

Mean ARI was 0.974 across full-sample refits (minimum 0.970) and 0.972 across 70% subsamples (minimum 0.917). Mean matched centroid displacement was 3.87 and 3.43 index points, respectively. The average of the largest displacements per run was 12.01 and 10.60 points. The high partition-level agreement therefore accompanies more substantial movement of individual centroids.

4.3. Spatial Concentration of the Heating Profiles

Table 2 compares same-profile neighbour shares with national cell shares. The excess is greatest for C1 at 37.91 percentage points. C4 shows the weakest same-profile concentration. Most of its neighbouring cells belong to C3 (54.02%) or C2 (29.95%).

Table 2. Spatial concentration of the four heating profiles under the 6 km neighbourhood definition.

Cluster	Cell Share (%)	Same-Profile Neighbours (%)	Difference (pp)
C1	13.90	51.81	+37.91
C2	34.97	41.09	+6.12
C3	48.51	60.67	+12.17
C4	2.62	4.49	+1.87

4.4. Global Spatial Autocorrelation of the Two Heating Indicators

At 6 km, global Moran's I is 0.403 for OLD and 0.064 for NEW (permutation pseudo-$p = 0.001$ for each). Both indices are positively autocorrelated, with a stronger spatial pattern for OLD.

The ordering persists across 3 km, 9 km and eight-nearest-neighbour weights (Table 3). The magnitude depends on the neighbourhood definition.

Table 3. Sensitivity of global Moran's I and local categories to the spatial weights.

Weights	Median Neighbours	Isolated Cells	OLD Moran's I	NEW Moran's I	Significant Local (%)	Agreement with 6 km (%)
3 km	12	71	0.511	0.083	18.16	73.08
6 km	40	0	0.403	0.064	34.21	100.00
9 km	87	0	0.353	0.057	44.38	78.94
8 nearest neighbours	9	0	0.531	0.089	17.60	70.70

4.5. Bivariate Local Moran's I and Its Relationship to the Heating Typology

At 6 km, 10,343 of the 30,237 cells (34.21%) have a significant pointwise local association between focal OLD and neighbouring NEW. The remaining 19,894 cells are not assigned to a significant category. The first letter of each configuration refers to the focal cell's OLD value and the second to the weighted mean NEW of its neighbours. High and low are measured relative to the respective means in the analytical sample. Table 4 reports the distribution by heating profile, and Figure 2 shows the spatial pattern.

Table 4. Distribution of bivariate LISA categories within the four K-means heating profiles (% of each cluster).

Profile	Non-Significant	HH	LL	LH	HL
C1	55.82	0.00	5.66	38.52	0.00
C2	67.53	1.78	14.64	11.40	4.65
C3	67.50	9.29	0.00	0.00	23.22
C4	63.93	19.92	0.25	5.04	10.84

Note. HH indicates above-average focal OLD and above-average neighbouring NEW. LL indicates below-average focal OLD and below-average neighbouring NEW. LH indicates below-average focal OLD and above-average neighbouring NEW. HL indicates above-average focal OLD and below-average neighbouring NEW. Non-significant cells do not meet the two-sided pointwise pseudo-$p \leq 0.05$ threshold. In all four configurations the NEW value refers to neighbouring cells.

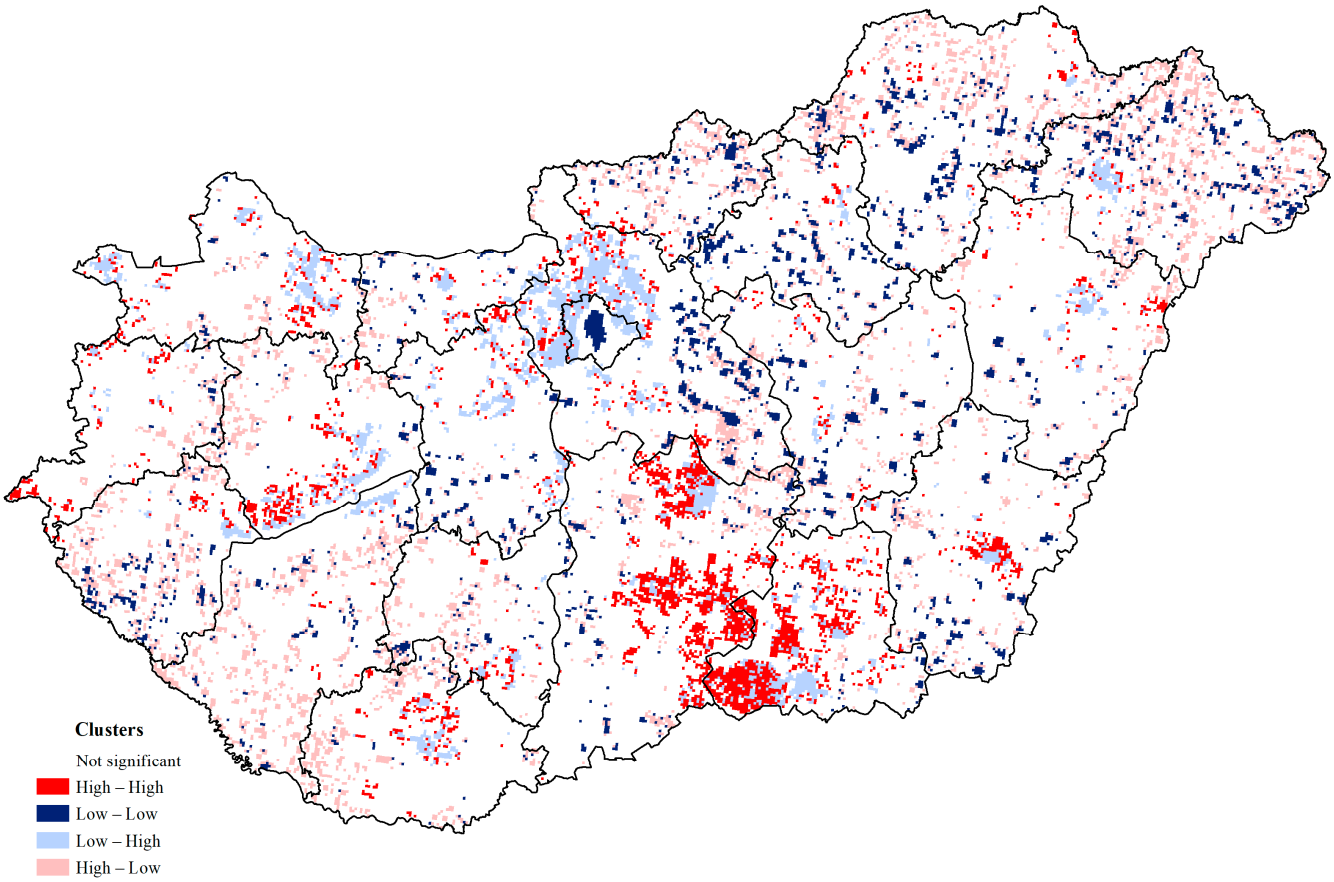

Figure 2. Pointwise bivariate local Moran categories for focal OLD and neighbouring NEW at 6 km.

In C1, 38.52% of cells are classified as LH, reflecting below-average focal OLD and above-average neighbouring NEW.

C2 has cells in all four significant quadrants. LL (14.64%) and LH (11.40%) are the most frequent significant categories.

In C3, 23.22% of cells are classified as HL and 9.29% as HH. High focal OLD therefore occurs in both low- and high-NEW neighbourhoods.

C4 has the largest HH share (19.92%). Taken together with the transition matrix, this shows that the HH category can occur where neighbouring cells have elevated NEW values and C4 itself remains locally uncommon.

The share classified as HH in C4 ranges from 8.32% under eight-nearest-neighbour weights to 25.73% at 9 km. The 6 km classification agrees with the alternatives for 70.70–78.94% of cells including non-significant observations. Cohen's kappa is 0.333–0.644. Agreement on the same significant quadrant among cells significant under either specification ranges from 27.69% to 57.72%. The local geography is more sensitive to weight choice than the global ordering of OLD and NEW.

The local percentages use all 30,237 cells as the denominator. The 71 isolated cells at 3 km are unclassified and counted as the absence of a significant category when computing agreement.

5. Spatial Regression Methods and Results

5.1. Regression Sample and Coverage

The original regression sample consists of the 17,146 unique mapped cells with a strictly positive NEW value in the mapping file. Matching cell identifiers verified that this filter reproduces the regression extract exactly. The other 13,091 mapped cells have a zero value

for this mapping indicator. The model uses the OLD and NEW columns in the regression export. The descriptive mapping export contains an earlier NEW series. OLD agrees across exports to rounding precision, while the two NEW series have a Pearson correlation of 0.802 on the matched mapped cells. We use the regression export consistently for all comparisons of regression inclusion and sample sensitivity. We matched each record to the parent export of 30,897 grid cells to check inclusion and recover the covariate labels. For the common regression specification, A is the share of dwellings built before 1960 and B is the share built after 2001; C–F represent construction materials, and G and I represent the lower and higher educational attainment shares, respectively. We omit the intermediate education category from the common specification. Table A1 summarises the three sample definitions and their analytical uses.

The original regression sample contains 55.5% of the 30,897 cells in the parent export and 95.6% of its occupied dwellings. Included cells have a median of 75 occupied dwellings, compared with four among the 13,751 excluded parent-export cells. Excluded cells have higher OLD and lower NEW on average. Among excluded cells with an observed NEW value, 85.3% record zero, versus 1.7% among included cells. The zero percentages in this comparison refer to NEW in the regression export, which differs from the mapping indicator used to define the sample. The main estimates therefore describe associations conditional on a positive mapped NEW value. The wider SEM assesses sensitivity to including mapped cells with a zero value.

The excluded group has 13,709 observed values for OLD, NEW and A–F and 13,751 for the dwelling count and education variables. Table 5 reports available-case summaries. Of the 13,751 parent-export cells outside the original regression sample, 13,091 are among the 30,237 unique mapped cells and 660 lie outside the mapping file.

Table 5. Observed characteristics by inclusion in the national regression sample.

Characteristic	Included	Excluded
Grid cells	17,146	13,751
Occupied dwellings per cell, median	75	4
Occupied dwellings per cell, mean	223.47	12.86
OLD, mean	64.03	83.89
NEW, mean	13.48	3.15
A, mean	21.11	34.09
B, mean	15.00	7.52
G, mean	38.32	44.23
I, mean	39.52	30.65

5.2. Spatial Models and Correlation

OLD and NEW enter the models on a 0–100 scale. Across the 17,146 cells in the original regression sample, Pearson's r is 0.123 and Spearman's rank correlation is 0.056. Using the same regression-export indicators for all 30,237 unique mapped cells, the corresponding values are −0.037 and −0.317. The marginal association therefore depends on the sample used. The two OLS residual series have a Pearson correlation of 0.269 in the regression sample.

We compared Gaussian maximum-likelihood spatial lag (SAR) and spatial error models (SEM) against OLS. The baseline regression graph uses the 6 km distance-weight file. We mapped the polygon identifiers to unique grid identifiers, removed same-grid links, collapsed repeated cell pairs to a single binary link, and retained links between regression cells. We row-standardised after restricting the graph to the estimation sample. Isolated

cells remain in the regressions with zero spatial lag. The alternatives are 3, 5 and 7 km graphs constructed in the same way, and a symmetrised eight-nearest-neighbour graph using projected cell locations.

At 6 km, the SEM has the lowest AIC for each outcome among the separate specifications. For OLD, AIC is 152,862.5 for OLS, 147,177.2 for SAR and 146,999.6 for SEM. For NEW the values are 143,135.3, 142,871.6 and 142,594.0. The corresponding OLS R^2 values are 0.386 and 0.263.

We also fitted a spatial seemingly unrelated regression (spatial SUR) with equation-specific spatial error parameters and a contemporaneous covariance between the spatially filtered innovations. The fitted innovation correlation is 0.253. The joint AIC is 288,465.6, compared with 289,593.6 for the two independent SEMs combined. The joint fit assigns spatial error parameters of 0.795 for OLD and 0.402 for NEW. All slope signs agree with those from the separate baseline SEMs. This model allows shared unobserved shocks across equations. It does not identify reciprocal causal effects between OLD and NEW. Table 6 presents the coefficient estimates for the separate and joint models.

Table 6. Separate spatial error model coefficients and joint spatial SUR estimates at 6 km.

Term	SEM OLD	SEM NEW	SUR OLD	SUR NEW
Intercept	88.544	5.511	89.278	5.431
A	0.180	0.211	0.179	0.209
B	−0.005	0.283	−0.008	0.284
C	−0.154	0.163	−0.154	0.163
D	−0.148	−0.074	−0.155	−0.078
E	−0.775	−0.157	−0.786	−0.155
F	−0.320	−0.111	−0.325	−0.114
G	−0.056	−0.115	−0.058	−0.112
I	−0.300	0.208	−0.305	0.215
Spatial error λ	0.805	0.417	0.795	0.402

5.3. Sensitivity to Spatial Weights and Sample Coverage

Table 7 shows that the spatial error parameter is larger for OLD in all five specifications. Table 8 compares selected slopes across those specifications. Across weights, recent housing (B) is positively associated with NEW, and higher educational attainment (I) is negatively associated with OLD and positively associated with NEW. The magnitude of the spatial error parameters varies with the graph. The local LISA results in Section 4.5 show greater sensitivity to the direction of these regression patterns.

Table 7. Sensitivity of the separate SEM estimates in the original regression sample of 17,146 cells.

Weights	Median Neighbours	Isolated	OLD λ	NEW λ
3 km	8	232	0.670	0.272
5 km	17	22	0.766	0.367
6 km	25	7	0.805	0.417
7 km	32	2	0.832	0.480
Symmetrised 8-NN	9	0	0.759	0.329

Table 8. Selected SEM slopes across alternative spatial weights.

Weights	OLD B	NEW B	OLD I	NEW I
3 km	−0.034	0.275	−0.287	0.202
5 km	−0.006	0.282	−0.286	0.207
6 km	−0.005	0.283	−0.300	0.208
7 km	0.000	0.284	−0.318	0.211
Symmetrised 8-NN	−0.038	0.278	−0.232	0.206

We repeated the 6 km SEMs on all 30,237 unique mapped cells as a sensitivity check for the restriction to cells with a positive NEW value in the mapping file. We used the same regression-export outcomes and covariates in both samples. The OLD and NEW spatial parameters are 0.837 and 0.343. Recent housing has coefficients of −0.029 for OLD and 0.187 for NEW. Higher educational attainment has coefficients of −0.146 and 0.108, respectively. This check assesses the stability of conditional associations across different national cell samples.

6. Discussion

OLD has higher global autocorrelation than NEW across all four neighbourhood definitions. At the local level, different matrices change the mix and location of LISA categories. A 6 km radius holds physical distance constant while allowing a wide range in the number of observed neighbours. The nearest-neighbour rule sets a minimum number of links, with a variable physical reach. These specifications describe different spatial contexts. The estimates do not establish separate urban and rural effects.

C4 is primarily a profile of sparsely occupied cells. Within the fixed C4 membership, weighting by occupied dwellings reduces mean OLD and NEW from 94.01 and 77.56 to 82.21 and 65.41. Among the 102 C4 cells with more than five dwellings, the unweighted means are 78.18 and 57.94. Among the 13 cells with at least twenty dwellings they are 61.57 and 60.81. Re-estimating K-means on all cells meeting either minimum yield, four clusters differentiated primarily by OLD, with maximum NEW centroids of 12.18 and 11.65, respectively. A refit can absorb surviving high-NEW cells into broader profiles when they become rare in the analysis sample.

The one-dwelling C4 cells establish co-occurrence of two recorded characteristics at the dwelling level, provided the reports are accurate. In particular, an occupied dwelling can report a traditional fuel and own photovoltaic equipment without using the electricity for space heating. The result therefore offers limited evidence on fuel stacking as a heating practice. For multi-dwelling cells, dwelling-level cross-tabulations would distinguish co-occurrence within a dwelling from aggregation across dwellings.

The separate and joint national regressions associate more recent housing with a higher NEW index and higher educational attainment with lower OLD and higher NEW. Traditional fuel use exhibits greater residual spatial dependence across all examined graphs. The exclusion of many small cells with low NEW shifts the marginal OLD–NEW correlation. This is important when interpreting the regression results as evidence about low-density areas. Income, renovation, fuel prices and network access could contribute to the observed patterns.

7. Conclusions

Across Hungary, reported use of traditional heating fuels was more geographically concentrated than ownership of newer energy equipment. The study measures these

two patterns separately for each 1 km grid cell. The traditional fuel-use indicator (OLD) combines reported use of wood, coal or lignite, and fuels recorded in the census category "other fuels". The newer equipment ownership indicator (NEW) combines reported ownership of heat pumps, solar thermal collectors and photovoltaic panels. Equipment ownership indicates what is installed in a dwelling. It provides no direct measure of what supplies space heating or hot water. Across all tested definitions of neighbouring cells, the spatial error models also indicated stronger residual spatial dependence for traditional fuel use after accounting for the housing and education variables included in the models. The profile combining high values on both indicators occurred predominantly in cells with few occupied dwellings. In such cells, a small number of dwellings can strongly influence the profile, so both indicators should be interpreted alongside the number of occupied dwellings.

Municipalities can use clusters of cells with high reported traditional fuel use to identify neighbourhoods for outreach and dwelling-level energy assessments. Combining the maps with dwelling counts, building condition and information on heating affordability can help prioritise where to offer advice or arrange site visits. Where both indicators are high, traditional fuel use and newer equipment ownership are both prominent at the grid-cell level. Before recommending a retrofit, an assessment should establish which systems individual households use for space heating and hot water, and whether traditional fuels remain in regular or occasional use alongside newer equipment.

National policymakers could assess whether renovation support and technical advice reach older dwellings and areas where ownership of newer equipment is relatively low. The association between more recent construction and higher equipment ownership supports monitoring installations in new buildings separately from installations made during retrofits. Programme evaluation should track whether households actually reduce their use of traditional fuels after installation, whether they continue to use existing systems, and how their heating costs and indoor warmth change. The grid profiles can help identify where to collect the dwelling-level evidence needed to answer these questions. At the EU level, the Energy Performance of Buildings Directive provides a framework for national building renovation plans and monitoring energy poverty [19]. A consistent reporting framework could pair Eurostat's end-use energy accounts with local indicators that distinguish fuel, equipment ownership, space heating, domestic hot water and building condition [13]. Cross-country comparisons should document differences in fuel categories, grid coverage and disclosure practices before comparing local profiles.

The 2022 census provides a cross-sectional picture. In cells with several dwellings, aggregation obscures which households jointly report fuels and equipment. Even in a single-dwelling cell, the equipment indicator does not measure frequency, energy delivered or end use. The fuel categories leave several important energy carriers outside the indices and group heating oil with other fuels. The main regression sample covers a large share of dwellings but disproportionately excludes small cells, and the local categories are sensitive to the spatial weights. The reported pointwise LISA probabilities are unadjusted for multiple comparisons. The regression coefficients describe conditional associations, with unmeasured influences such as household income, energy prices, network access and renovation histories potentially affecting the results.

Future work should follow the same dwellings across heating seasons and record their primary, auxiliary and backup appliances separately for space heating and hot water. Surveys and, where feasible, metered or expenditure data could record the fuel used, operating frequency and delivered energy. Linking those observations to renovation and installation dates, building energy performance, household resources, local fuel prices and network access would help distinguish genuine fuel replacement from seasonal or backup

use. Before-and-after comparisons with a suitable comparison group could test whether renovation programmes reduce solid-fuel use, improve thermal comfort or lower household costs. These data would support more accurate estimates of changes in emissions and energy affordability. A cross-tabulation of primary and secondary fuels with installed equipment and the services it supplies would also help quantify continued fuel use after installation.

Author Contributions: Conceptualization, T.L.S.; Methodology, G.T.; Formal Analysis, G.T.; Resources, G.T.; Data Curation, G.T.; Writing—Original Draft, G.T. and T.L.S.; Writing—Review and Editing, G.T.; Visualisation, G.T.; Supervision, T.L.S. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Data Availability Statement: The official Census 2022 results and questionnaire are available from the Hungarian Central Statistical Office [7,17,18]. Enquiries about access to the analytical grid extracts and analysis outputs should be directed to the corresponding author. Access to census-derived grid data is subject to the applicable data access and statistical confidentiality requirements.

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

Appendix A. Analytical Samples

Table A1. Analytical samples and unit counts.

Sample	Cells	Use
Parent grid export	30,897	Reference population for sample coverage (Table 5)
Unique mapped cells	30,237	Typology, global Moran's I, LISA and 6 km SEM sample sensitivity
Original regression sample (mapped NEW > 0)	17,146	OLS, SAR, SEM and spatial SUR

References

1. Hosier, R.H.; Dowd, J. Household fuel choice in Zimbabwe: An empirical test of the energy ladder hypothesis. *Resour. Energy* **1987**, *9*, 347–361. [CrossRef]
2. Leach, G. The energy transition. *Energy Policy* **1992**, *20*, 116–123. [CrossRef]
3. Masera, O.R.; Saatkamp, B.D.; Kammen, D.M. From linear fuel switching to multiple cooking strategies: A critique and alternative to the energy ladder model. *World Dev.* **2000**, *28*, 2083–2103. [CrossRef]
4. van der Kroon, B.; Brouwer, R.; van Beukering, P.J.H. The energy ladder: Theoretical myth or empirical truth? Results from a meta-analysis. *Renew. Sustain. Energy Rev.* **2013**, *20*, 504–513. [CrossRef]
5. Stojilovska, A.; Dokupilová, D.; Gouveia, J.P.; Bajomi, A.Z.; Tirado-Herrero, S.; Feldmár, N.; Kyprianou, I.; Feenstra, M. As essential as bread: Fuelwood use as a cultural practice to cope with energy poverty in Europe. *Energy Res. Soc. Sci.* **2023**, *97*, 102987. [CrossRef]
6. La Belle, M.; Szép, T.; Tóth, G. Why neighbors matter in the energy transition: The diffusion of social practices, technologies, and knowledge between municipalities. *Renew. Sustain. Energy Transit.* **2025**, *7*, 100109. [CrossRef]
7. Hungarian Central Statistical Office. Census 2022: Housing and Personal Questionnaire, English Version. Available online: https://nepszamlalas2022.ksh.hu/media/HCSO_census_2022_english.pdf (accessed on 18 November 2026).
8. Kowsari, R.; Zerriffi, H. Three dimensional energy profile: A conceptual framework for assessing household energy use. *Energy Policy* **2011**, *39*, 7505–7517. [CrossRef]
9. Zhu, X.; Yun, X.; Meng, W.; Xu, H.; Du, W.; Shen, G.; Cheng, H.; Ma, J.; Tao, S. Stacked use and transition trends of rural household energy in mainland China. *Environ. Sci. Technol.* **2019**, *53*, 521–529. [CrossRef] [PubMed]
10. Tóth, G.; Szép, T.; LaBelle, M.C. A lakossági fűtési célú energiafelhasználás térbeli mintázata Magyarországon. *Területi Stat.* **2025**, *65*, 555–581. [CrossRef]

11. Department for Business, Energy & Industrial Strategy. *Hybrid Heat Pumps Study*; BEIS: London, UK, 2018. Available online: https://www.gov.uk/government/publications/hybrid-heat-pumps-study (accessed on 18 November 2026).
12. Martinopoulos, G. Domestic hot water solar thermal systems as an energy-security alternative in the EU: A parametric analysis. *Sustain. Energy Technol. Assess.* **2024**, *69*, 103921. [CrossRef]
13. Eurostat. Disaggregated Final Energy Consumption in Households—Quantities (nrg_d_hhq). Available online: https://ec.europa.eu/eurostat/databrowser/view/nrg_d_hhq/default/table?lang=en (accessed on 18 November 2026).
14. Sovacool, B.K. Conceptualizing urban household energy use: Climbing the "Energy Services Ladder". *Energy Policy* **2011**, *39*, 1659–1668. [CrossRef]
15. Anselin, L. Local indicators of spatial association—LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [CrossRef]
16. Anselin, L.; Syabri, I.; Kho, Y. GeoDa: An introduction to spatial data analysis. *Geogr. Anal.* **2006**, *38*, 5–22. [CrossRef]
17. Hungarian Central Statistical Office. Census 2022: Census Database. Available online: https://nepszamlalas2022.ksh.hu/en/database/ (accessed on 18 November 2026).
18. Hungarian Central Statistical Office. Census 2022: Final Data. Available online: https://nepszamlalas2022.ksh.hu/en/results/final-data (accessed on 18 November 2026).
19. European Parliament and Council of the European Union. Directive (EU) 2024/1275 of 24 April 2024 on the Energy Performance of Buildings (Recast). *Off. J. Eur. Union* **2024**, *L 2024/1275*. Available online: https://eur-lex.europa.eu/eli/dir/2024/1275/oj/eng (accessed on 18 November 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.