**Mihaela SIMIONESCU, PhD**
mihaela.simionescu@faa.unibuc.ro, mihaela.simionescu@ipe.ro
University of Bucharest, Bucharest, Romania
Romanian Academy, Institute for Economic Forecasting, Bucharest, Romania

**Bogdan OANCEA, PhD (corresponding author)**
bogdan.oancea@faa.unibuc.ro, bogdan.oancea@incdsb.ro
University of Bucharest, Bucharest, Romania
National Institute of Research and Development for Biological Sciences, Bucharest, Romania

# Explaining Regional Energy Poverty in Mediterranean Europe: A Multilevel Regression and Machine Learning Approach

**Abstract.** *Despite the European Union's extensive efforts to combat energy poverty (EP), especially amid climate change and recent energy crises, several member states continue to face substantial challenges. This paper analyses the determinants of EP in five vulnerable Mediterranean countries (France, Spain, Italy, Greece, and Portugal) using recent regional Eurostat data (NUTS 2, 2024). A multilevel regression framework is complemented by a machine learning analysis, combining fixed-effects logistic models, regularised estimators, and SHAP interpretability, to capture nonlinear interactions and test the robustness of the inferred relationships. The results show that regional economic growth, higher education, better living conditions, and greater labour productivity generally alleviate EP. In addition, inflation, unemployment, and financial instability exacerbated it. Internet access was found to mediate the link between economic growth and EP, but it does not fully explain structural disparities. The interaction between tertiary education and Internet use, although improving connectivity, has been associated with higher utility arrears, reflecting increased energy demand from digital activities. The findings highlight the multidimensional nature of EP and the importance of integrating digital and educational policies into regional development strategies.*

**Keywords**: *energy poverty, multilevel models, logistic regression, machine learning.*

**JEL Classification:** Q41, Q48, R11, I32, C38.

| Received: 21 November 2025 | Revised: 2 June 2026 | Accepted: 7 June 2026 |
|---|---|---|

## 1. Introduction

Over the past decade, a concerted effort by civil society and the scientific community has pushed energy poverty (EP) to the forefront of the European Union's political agenda. This urgent focus is determined by the fact that anywhere from 35 to 72 million EU citizens are grappling with EP. This leads to detrimental effects on public health, the environment, labour productivity, education, social cohesion, and overall well-being (Teixeira et al., 2024). Among the most susceptible are

DOI: 10.24818/18423264/60.2.26.04

households in Mediterranean countries. For instance, in 2023, a significant portion of homes in Spain (20.8%) and Portugal (20.8%) struggled to maintain adequate indoor warmth, with Greece (19.2%), France (12.1%), and Italy (9.5%) also reporting notable, though somewhat lower, percentages (Statista, 2025).

EP is a restriction in energy consumption that negatively impacts a household's basic energy needs, well-being, and health. Its primary drivers are income, energy demand, energy prices, and the energy efficiency of dwellings. However, to effectively identify the most vulnerable populations and tailor appropriate policy interventions, a more in-depth analysis at the regional and household level is essential.

This paper investigates the determinants of energy poverty in five vulnerable Mediterranean countries - France, Spain, Italy, Greece, and Portugal - using recent regional (NUTS 2) data for 2024. To do so, it combines multilevel econometric models with interpretable machine-learning techniques, enabling the analysis of both structural relationships and the robustness of the findings across flexible predictive settings. By integrating fixed-effects logistic regression, penalised regularisation, and SHAP-based attribution, the study captures both country-level structure and within-country heterogeneity in a transparent and reproducible framework. Its methodological contribution lies less in the prediction itself than in the joint use of hierarchical econometric modelling and explainable machine learning to identify the layered drivers of regional energy poverty in Mediterranean Europe.

While Bardazzi et al. (2023) used EU-SILC data and latent class analysis to identify “energy sufficient”, “energy poor”, and “energy vulnerable” households in Southern Europe, few studies have examined the causes of EP at the regional level. Some work has applied logistic regressions in national contexts, but the mechanisms through which economic, social, and technological factors interact remain unclear. The level of development clearly shapes EP, yet the role of economic growth as both a driver and an outcome of technological and social progress has rarely been explored. Recent evidence suggests that the digital economy can reduce EP, although results are mixed across regions (Simionescu and Cifuentes-Faura, 2024).

This study examines the determinants of EP in five Mediterranean countries using 2024 regional data and evaluates Internet access and higher education as mediating and moderating factors in the growth-EP relationship, measured by utility arrears and thermal comfort. Most prior work has taken a macroeconomic view, analysing EU-wide or national aggregates (e.g., Spain, Portugal, Greece, Italy, France), which obscures regional disparities. By contrast, this study’s regional approach provides finer insight into local vulnerabilities, bridging the gap between cross-country and single-country household analyses (Simionescu and Cifuentes-Faura, 2024).

Therefore, the paper contributes to the literature in several ways, making three key contributions. First, it identifies the main drivers of EP in five Mediterranean EU states and quantifies their effects using regional-level data, offering a more comprehensive framework for cross-country comparison. Second, it provides policy recommendations tailored to the specific economic and social contexts of these countries, addressing both affordability and access dimensions that EU strategies still lack. Third, it explores the mechanisms through which economic growth influences EP, emphasising the mediating role of Internet access and the moderating role of higher education, while recognising that additional structural factors also shape this relationship.

In addition to the econometric estimations, we also implemented a complementary machine-learning framework comprising penalised logistic regression, leave-one-country-out cross-validation, and SHAP-based interpretability (Lundberg & Lee, 2017). This approach quantifies the relative influence of regional socioeconomic factors and national structural characteristics on energy-poverty outcomes. In addition, it verifies the stability of associations under flexible predictive modelling approaches.

Given the framework of this study, the paper provides a short description of the determinants of EP. The next sections present data and methods, main results, and discussion to formulate specific conclusions.

## 2. Literature review

Assessing EP is complex, as it reflects both household factors (income, energy use) and external conditions (prices, climate). Numerous single and composite indicators have been developed, each capturing different dimensions (Teixeira et al., 2024). Composite measures often hide detail through aggregation. Therefore, this study relies on Eurostat's primary indicators, arrears on utility bills, and the inability to keep the home adequately warm, to capture EP directly.

Many studies have examined EP through its socioeconomic and household determinants, often focusing on specific regional contexts (Mulder et al., 2023). Bollino and Botti (2017) found that low income, housing quality, geography, and population density are key drivers of EP. Women and immigrants were more affected, while older individuals tended to be less vulnerable. Primc et al. (2021) highlighted two main EP pathways: one linked to low household income and policy intervention, and another driven by high energy costs. They noted that even countries prioritising energy policy often struggle to achieve effective outcomes due to the scale and financial burden of EP programs.

Studies have identified both drivers and mitigators of EP (Churchill and Smyth, 2020). Economic growth generally reduces EP. The evidence from Latin America (Castro-Cárdenas and Ibarra-Yunez, 2023), Italian regions (Bardazzi et al., 2023), and broader cross-country analyses (Igawa and Managi, 2022) confirms a long-run cointegration between growth and lower EP.

Household attributes, particularly educational attainment, play a significant role in determining EP. In South Africa, factors such as spending behaviour, education, ethnicity, and housing quality have been identified as key drivers (Ismail and Khembo, 2015). Across European contexts, higher education consistently reduces the likelihood of EP, while smaller dwellings are associated with greater vulnerability (Legendre and Ricci, 2015). Larger or detached homes may indicate older and less energy-efficient structures, moderating the expected advantage.

Several additional factors exacerbate EP. Rising unemployment, particularly during economic downturns, intensifies EP (Halkos and Gkampoura, 2021), while social exclusion and poverty directly limit households' ability to maintain adequate warmth (Dong et al., 2022). Inflation further aggravates vulnerability by reducing purchasing power. Effective governance can mitigate its impact, as shown by Mawutor et al. (2024) for 40 African countries.

Technological progress, particularly digital innovation, plays a central role in reducing EP and achieving sustainability goals. Advances in technology support stable energy supply, clean energy access, and renewable integration (Paul and Sharma, 2019). In the digital era, innovation drives the transformation of energy systems through the expansion of the digital economy, improved efficiency, and greener consumption patterns (Bardazzi et al., 2023).

The digital economy shapes energy systems by applying information technologies to improve efficiency, innovation, and transition toward low-carbon development (Li et al., 2021). Internet technologies reduce dependence on traditional energy sources and promote industrial upgrading (Ishida, 2015). However, their overall effect on energy use remains debated. ICT expansion can raise consumption while enhancing efficiency (Usman et al., 2021), although gradual digital investment may lower energy demand. Empirical evidence also suggests that the digital economy supports the energy transition on both the production and consumption fronts (Zhang et al., 2023).

## 3. Data and methodology

The Mediterranean region has shown susceptibility to rising EP, with Greece experiencing the most severe effects. This research uses the latest available data from Eurostat and World Bank corresponding to 2024 and covers more NUTS 2 regions in five countries: Portugal (10 regions), Spain (23 regions), France (33 regions), Italy (24 regions), and Greece (16 regions). Table 1 shows descriptive statistics of the variables used in this study at the NUTS 2 level, as the well as country-level variables.

**Table 1. Variables and descriptive statistics**

| Variable | Notation | Data source | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|---|
| Arrears (mortgage or rent, utility bills or hire purchase) (% of population) | arrears | Eurostat | 15.10 | 13.90 | 0.4 | 72.2 |
| Inability to keep home adequately warm (% of population) | warm | Eurostat | 13.57 | 5.44 | 1.9 | 26.6 |

| Variable | Notation | Data source | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|---|
| Regional gross domestic product (PPS per inhabitant) | GDP | Eurostat | 31490.91 | 9176.30 | 16100 | 63400 |
| Share of population aged 25-64 who successfully completed superior studies (% of population) | tertiary education | Eurostat | 32.44 | 9.74 | 17.6 | 56.5 |
| At-risk-of-poverty rate (% of population) | poverty | Eurostat | 19.25 | 8.45 | 5.9 | 53.3 |
| Unemployment rate (people aged 15-74, %) | unemployment | Eurostat | 8.91 | 4.54 | 2 | 28.2 |
| Inability to face unexpected financial expenses (% of population) | unexpected expenses | Eurostat | 34.30 | 10.56 | 11.2 | 63.6 |
| Average number of rooms per person | rooms | Eurostat | 1.61 | 0.30 | 1.2 | 2.3 |
| Real labour productivity, index, 2015=100 | productivity | Eurostat | 98.1 | 10.94 | 18.6 | 119.9 |
| Households that have internet access at home (% of households) | Internet | World Bank | 93.858 | 2.95 | 84.74 | 98.8 |
| GDP per capita in constant 2021 international $ at country level | GDP country | World Bank | 48488.11 | 6249.66 | 36854.5 | 54017.9 |
| Inflation rate based on harmonised Indices of Consumer Prices at country level (%) | inflation country | Eurostat | 2.301 | 0.70 | 1.1 | 3 |
| Final energy consumption in households at country level in thousand tons of oil equivalent | energy country | Eurostat | 4093.05 | 5322.97 | 313.30 | 13863.2 |

*Source*: Authors' description.

Mixed-effects regressions (hierarchical or multilevel models) are used to explain arrears and the inability to keep homes adequately warm in Mediterranean regions in 2024, given that these regions belong to certain countries. The basic form of the linear multilevel model is given by:

$$EP_{ij} = \alpha_0 + \alpha_1 \cdot X_{ij} + u_j + v_{ij} \quad (1)$$

where $EP_{ij}$ is the energy poverty indicator (arrears/warm) for the $i$-th region in the $j$-th country, $\alpha_0, \alpha_1$ are fixed-effect parameters, X is the vector of predictors, $u_j$ is the random-effect for the $j$-th country following a normal distribution of null average and constant dispersion, $v_{ij}$ is the error for the *i*-th region in the *j*-th country following a normal distribution of constant dispersion.

Particular attention should be assigned to the channels through which economic growth impacts EP indicators. Therefore, separate models are proposed to understand the role of mediators and moderators.

This paper investigates whether economic growth impacts EP via Internet access using a modified version of Wen and Ye's (2014) mechanism effect test.

$$Internet_{ij} = \alpha''_0 + \alpha''_1 \cdot GDP_{it} + \alpha''_3 \cdot energy_j + u''_j + v''_{ij} \quad (2)$$

$$EP_{ij} = \alpha'_0 + \alpha'_1 \cdot GDP_{ij} + \alpha'_2 \cdot Internet_{ij} + \alpha'_3 \cdot energy_j + u'_j + v'_{ij} \quad (3)$$

Initially, it is necessary to confirm the statistical significance of parameters $\alpha_1$ in the following equation:

$$EP_{ij} = \alpha_0 + \alpha_1 \cdot GDP_{ij} + \alpha_2 \cdot energy_j + u_j + v_{ij} \quad (4)$$

where $u_j, v_{ij}$, $u'_j, v'_{ij}$, $u''_j, v''_{ij}$ are the errors and $\alpha_0, \alpha_1, \alpha'_0, \alpha'_1, \alpha''_0, \alpha''_1$ the fixed-effect parameters.

First, we should check if GDP has a significant impact on EP. If the parameter $\alpha''_1$ and $\alpha'_2$ are significant, a mechanistic effect is present. Then, the significance of $\alpha'_1$ is verified and a partial mechanism is detected if this coefficient is significant. Otherwise, the mechanism is total.

This study also explains the moderating effects on the mediating role of Internet access in the relationship between economic growth and EP. The model uses tertiary education as moderator to examine its impact on the GDP-Internet access nexus, thereby impacting the indirect effect of GDP on EP via Internet access. More educated people are expected to have easier access to the Internet. Higher educated people usually have higher wages compared to unskilled workers and can afford more Internet access, while more educated people could have jobs that require Internet use. The models are constructed in two steps:

- The effect of GDP and of the moderator (*tertiary education*) on the mediator (*Internet*):

$$Internet_{ij} = \beta_0 + \beta_1 \cdot GDP_{ij} + \beta_2 \cdot tertiary\ education_{ij} + \beta_3 \cdot (GDP_{it} \cdot education_{ij}) + \beta_4 \cdot energy_j + \mu_j + \epsilon_{ij} \quad (5)$$

- The influence of mediator (*Internet*) and of the moderator (*tertiary education*) on EP:

$$EP_{ij} = \gamma_0 + \gamma_1 \cdot GDP_{ij} + \gamma_2 \cdot Internet_{ij} + \gamma_3 \cdot tertiary\ education_{ij} + \gamma_4 \cdot \left(GDP_{ij} \cdot tertiary\ education_{ij}\right) + \gamma_5 \cdot (Internet_{ij} \cdot tertiary\ education_{ij}) + \gamma_6 \cdot energy_j + \mu'_j + \epsilon'_{ij} \quad (6)$$

We are particularly interested in the significance of some parameters associated with interaction terms: $\beta_3$ and $\gamma_5$. A significant $\beta_3$ suggests a significant effect of higher education on the GDP-Internet access nexus.

We complemented the econometric analysis with a machine-learning framework adapted to a small hierarchical dataset. This additional analysis serves three purposes. First, it evaluates within-country associations between regional characteristics and energy-poverty outcomes. Second, it tests how well the models transfer across countries. Third, it quantifies predictor importance using SHAP-based explanations.

Two continuous indicators are analysed: arrears on utility bills and inability to keep the home warm. To enable stable probabilistic classification and maintain balanced classes, each outcome is median-thresholded into a balanced binary outcome. This thresholding is not meant to reproduce any official definition of EP. It provides a neutral, distribution-based split that facilitates model training and comparative evaluation across countries. Predictors include regional GDP per capita, tertiary education, poverty rate, unemployment, unexpected expenses, rooms per person, productivity, internet access, and macro controls (national inflation and energy prices). Country fixed effects are implemented as country dummy variables, given the small number of countries. These absorb between-country differences but are not interpreted. Continuous predictors are median-imputed and standardised using the following transformation:

$$X = \frac{X_{ij} - \text{median}(X_j)}{\text{MAD}(X_j)} \text{(implemented after median imputation)} \tag{7}$$

Our primary model is a logistic regression with country fixed effects:

$$\text{logit}\,\Pr(Y_i = 1) = \alpha + \delta_{c(i)} + \beta^{\top}\text{X}_i \tag{8}$$

where $c(i)$ denotes country and $\delta_{c(i)}$ are country intercepts (baseline dropped). This specification absorbs between-country structural heterogeneity and identifies the within-country association $\beta$ between regional characteristics and energy-poverty outcomes.

To mitigate small-sample instability and multicollinearity, we also fit ridge and lasso logistic regressions with 5-fold cross-validation tuning $C \in [10^{-2}, 10^{2}]$ by AUC. Generalisation across national contexts is assessed with leave-one-country-out cross-validation (LOCO-CV). In each fold, one country is held out for testing while training uses the remaining countries. Given the small number of countries, this cross-validation scheme approximates out-of-domain evaluation and avoids overly optimistic in-sample accuracy.

Model contributions are quantified using SHAP (SHapley Additive exPlanations) for the logistic model. For observation $i$, the prediction relative to a background expectation decomposes additively:

$$f(\mathrm{x}_i) - \mathbb{E}[f(\mathrm{X})] = \sum_j \phi_{ij} \tag{9}$$

and global importance is summarised by $\mathrm{Imp}(j) = \mathbb{E}_i[|\ \phi_{ij}\ |]$. As a sensitivity analysis, we estimated a shallow Gradient Boosting Machine (200 trees, depth 2) and computed TreeExplainer SHAP. In our small-N hierarchical settings, ensembles tend to split on high-signal categorical features. GBM-SHAP serves to test whether country effects dominate.

## 4. Results

Before estimating the substantive models, we examined the correlation structure of the predictors in order to avoid including highly correlated variables in the same specification. We then estimated null random-intercept models to test whether the regional observations were sufficiently clustered within countries to justify a multilevel approach. The likelihood-ratio tests strongly support the inclusion of random intercepts. For arrears on utility bills, the LR statistic is 175.32, while for the inability to keep the home adequately warm, it is 59.82; in both cases, the p-value is below 0.01. The intraclass correlation coefficients are also sizeable, equal to 0.88 for arrears and 0.52 for inability to keep the home warm. These results indicate that a substantial share of the variation in energy-poverty indicators is located at the country level, which supports the use of hierarchical modelling.

Table 2 reports the mixed-effects models for arrears on utility bills. Across specifications, regional GDP, tertiary education, labour productivity, and the average number of rooms per person are associated with lower arrears. These variables can be interpreted as markers of stronger regional economic conditions, better human capital, and better living standards.

On the contrary, unemployment, poverty, the inability to face unexpected expenses, and national inflation are associated with higher arrears. Internet access is also positively associated with arrears in some specifications, suggesting that digital inclusion alone does not offset underlying financial vulnerability.

Real labour productivity, regional growth, and tertiary education appear to be the main protective factors against utility arrears by increasing the utility of household income. However, one possible interpretation is that greater connectivity may be associated with more energy-intensive consumption patterns or higher household service costs in financially constrained settings. One possible interpretation is that higher connectivity may be associated with more energy-intensive household consumption patterns, particularly in financially constrained settings.

Overall, the results indicate that arrears are shaped by both regional socioeconomic disadvantage and broader national macroeconomic pressures.

**Table 2. Mixed-effects regressions to explain arrears on utility bills in Mediterranean regions in 2024**

| Variable | Coefficients | | | | | | |
|---|---|---|---|---|---|---|---|
| GDP | -0.002*** (0.0007) | - | - | - | - | - | - |
| unemployment | - | 0.80*** (0.129) | - | - | - | - | - |
| productivity | - | - | -0.15*** (0.051) | - | - | - | - |
| tertiary education | - | - | - | -0.38*** (0.084) | - | - | - |
| poverty | - | - | - | - | 0.24*** (0.05) | - | 0.23*** (0.05) |
| Internet | - | - | - | - | 0.70*** (0.19) | - | 0.64*** (0.19) |
| rooms | - | - | - | - | -7.30*** (1.94) | - | -7.96*** (1.93) |
| unexpected expenses | - | - | - | - | 0.14*** (0.04) | - | 0.13*** (0.04) |
| GDP country | - | - | - | - | - | 0.80*** (0.129) | - |
| inflation country | - | - | - | - | - | - | 5.54*** (1.01) |
| constant | 23.21*** (6.42) | 9.07 (6.18) | 33.31*** (8.886) | 28.35*** (7.136) | -54.60** (21.69) | 9.07(6.1) | 34.60 (35.06) |
| variance of the random intercepts for the countries | 0.79 | 0.833 | 0.69 | 0.71 | 0.88 | 0.83 | 0.73 |
| residual variance | 0.29 | 0.291 | 0.27 | 0.25 | 0.35 | 0.21 | 0.25 |
| Wald test (p-value) | 9.99 (0.0016) | 39.04 (0.00) | 8.54 (0.003) | 20.64 (0.000) | 11.20 (0.0004) | 39.04 (0.000) | 395.91 (0.00) |
| LR test vs. linear model: chi-square test (p-value) | 114.11 (0.00) | 148.88 (0.00) | 158.23 (0.00) | 151.23 (0.00) | 165.95 (0.00) | 148.88 (0.00) | 0.57 (0.98) |

*Source*: own calculations in Stata 15. Note: standard errors in brackets.

Table 3 presents the mixed-effects estimates for the inability to keep the home adequately warm. Regional GDP, tertiary education, Internet access, and housing space are associated with lower values of this indicator, suggesting that both material living conditions and access to resources improve thermal comfort.

On the contrary, poverty, unemployment, inflation, and the inability to face unexpected expenses increase the probability of thermal deprivation. These findings indicate that the capacity to maintain adequate warmth depends not only on income-related variables, but also on broader macroeconomic stability.

The analysis for 2024 reveals a distinct divide between structural stabilisers and inflationary shocks in the Mediterranean region. Factors such as regional economic growth, tertiary education, and improved living standards acted as buffers by significantly reducing the inability to keep homes adequately warm. Human capital and infrastructure are essential for climate resilience. On the other hand, the vulnerability of these regions to macroeconomic volatility represents a threat to energy security. This suggests that long-term development might improve energy access, but short-term economic instability can erode these gains, making unexpected financial expenses an important driver of immediate energy poverty.

**Table 3. Mixed-effects regressions to explain inability to keep home adequately warm in Mediterranean regions in 2024**

| Variable | Coefficients | | | | | |
|---|---|---|---|---|---|---|
| GDP | -0.002*** (0.0004) | - | - | - | - | - |
| productivity | - | - | - | - | 0.046 (0.103) | - |
| tertiary education | - | -0.214*** (0.065) | - | - | - | - |
| poverty | - | - | 0.151*** (0.055) | - | - | 0.175*** (0.065) |
| Internet | - | - | - | -1.159*** (0.235) | -0.504*** (0.170) | - |
| rooms | - | - | - | -12.032*** (2.886) | - | -4.019* (2.231) |
| unexpected expenses | - | - | 0.234*** (0.043) | - | - | 0.189*** (0.049) |
| unemployment | - | - | - | - | 0.629*** (0.088) | - |
| GDP country | - | - | - | - | - | -0.001*** (0.0003) |
| inflation country | - | - | - | - | - | 0.545*** (1.005) |
| constant | 21.214*** (2.089) | 21.099*** (3.036) | 3.330** (1.685) | 142.849*** (24.934) | 49.692** (21.299) | 103.304*** (40.330) |
| variance of the random intercepts for the countries | 0.945 | 0.544 | 0.824 | 0.874 | 0.812 | 0.083 |
| residual variance | 0.120 | 0.128 | 0.223 | 0.109 | 0.645 | 0.044 |
| Wald test (p-value) | 22.38 (0.000) | 10.74 (0.001) | 118.49 (0.000) | 33.36 (0.000) | 75.22 (0.000) | 335.63 (0.000) |
| LR test vs. linear model: chi-square test (p-value) | 36.09 (0.000) | 56.98 (0.000) | 55.64 (0.000) | 37.02 (0.000) | 26.37 (0.000) | 0.13 (0.999) |

*Source*: own calculations in Stata 15. Note: standard errors in brackets.

Table 4 examines whether Internet access mediates the relationship between regional economic growth and energy poverty. The results show that regional GDP is negatively associated with both arrears and inability to keep the home warm, while it is positively associated with Internet access. This suggests that economically stronger regions tend to have both lower energy poverty and better digital connectivity.

Internet access itself is significantly related to arrears, but not to the thermal-comfort indicator. Therefore, the evidence supports only partial mediation, and only for some dimensions of energy poverty. This means that Internet access is one channel through which growth may reduce energy poverty, but it is clearly not the only one.

The results point to a tension between regional economic advancement and the broader national management of resources. Although regional growth is associated with lower energy poverty, it also coincides with an expansion of the digital infrastructure. However, more national energy consumption exacerbates EP because of increased costs or infrastructure strain that outweighs the gains in connectivity. The evidence of a partial mechanism suggests that other mediators like income inequality, regional policy changes, or energy efficiency standards may also play significant roles in how wealth translates into actual energy security for households.

**Table 4. The mediating role of Internet access in the energy poverty economic growth nexus**

| Variable | arrears | Internet access | warm |
|---|---|---|---|
| GDP | -0.002*** (0.0006) | 0.0015*** (0.0002) | -0.002*** (0.0005) |
| Internet | 0.275* (0.145) | - | -0.036 (0.210) |
| Energy consumption country | 0.003 (0.002) | 0.002*** (0.0008) | 0.004*** (0.0001) |
| constant | -9.010 (19.07) | 87.546*** (1.101) | 22.221 (18.682) |
| Wald test (p-value) | 26.25 (0.000) | 39.58 (0.000) | 35.47 (0.000) |
| LR test vs. linear model: chi-square test (p-value) | 12.88 (0.0002) | 3.71 (0.027) | 5.72*** (0.0084) |

*Source*: own calculations in Stata 15. Note: standard errors in brackets.

Table 5 evaluates whether tertiary education moderates the relationship between economic growth, Internet access, and energy poverty. Tertiary education is directly associated with lower energy poverty, while its interaction with GDP suggests that the strength of this association varies with the level of economic development.

Tertiary education is also positively associated with Internet access, and the GDP × tertiary education interaction is significant in the Internet equation. This implies that more educated regions benefit more strongly from economic growth in terms of digital connectivity.

However, the interaction between Internet access and tertiary education is positively associated with arrears. This result may indicate that digital inclusion, when combined with higher educational attainment, can also be associated with energy-consuming behaviours that increase utility costs.

The findings suggest a nonlinear relationship, where tertiary education acts as a main catalyst for reducing EP. Its efficacy is moderated by macroeconomic conditions and digital integration. Higher education reduces EP, but its interaction with GDP growth shows a diminishing marginal return. These results indicate that the benefits of education depend on the wider economic environment. The synergy between education and economic growth expands Internet access, but this digital advancement introduces a behavioural trade-off. The combination of higher education and connectivity actually correlates with increased utility arrears. This could indicate that "digitally empowered" educated households may have higher energy consumption patterns that outpace their financial management of utility costs.

**Table 5. The moderating role of tertiary education in energy poverty- economic growth nexus**

| Variable | Internet access | arrears | warm |
|---|---|---|---|
| GDP | 0.003*** (0.0006) | -0.005*** (0.0001) | -0.005*** (0.0001) |
| Internet access | - | -0.694* (0.363) | -0.708 (0.675) |
| tertiary education | 0.320*** (0.081) | -5.132** (2.119) | -3.885* (2.197) |
| GDP x tertiary education | 0.006*** (0.0002) | -0.003*** (0.0001) | -0.001** (0.0005) |
| Internet x tertiary education | - | 0.046** (0.023) | 0.035** (0.023) |
| Energy consumption country | 0.002*** (0.0004) | 0.004*** (0.0001) | 0.0004* (0.0002) |
| constant | 78.621*** (2.642) | 102.096* (59.234) | 102.139* (60.975) |
| Wald test (p-value) | 88.73 (0.000) | 70.70 (0.000) | 47.26 |
| LR test vs. linear model: chi-square test (p-value) | 19.34 (0.000) | 26.27 (0.000) | 8.00 (0.0023) |

*Source*: own calculations in Stata 15. Note: standard errors in brackets.

The machine-learning analysis is used as a robustness and interpretability exercise rather than as a replacement for the econometric models. Its main purpose is to assess whether the relationships identified in the multilevel analysis remain visible under alternative predictive specifications and to examine the extent to which the results are driven by country-level heterogeneity.

The results of the logistic models with country fixed effects are reported below. Across specifications (logit, ridge, lasso, GBM), the largest share of predictive mass is consistently associated with country fixed effects. In the GBM sanity check, SHAP values concentrate on country dummies, with socio-economic variables receiving much smaller attributions. This indicates that macro-institutional differences between countries drive most of the observed variability in NUTS 2 energy-poverty outcomes. After absorbing country effects, regional variables such as poverty risk, unexpected expenses, rooms per person, and tertiary education retain directionally consistent associations with the outcomes, but their magnitudes are attenuated. Table 6 reports the logistic fixed-effects coefficients for the arrears and warm outcomes. It shows that, once country fixed effects are included, the coefficients of most regional variables become smaller and less stable. This suggests that much of the variation in the binary energy-poverty outcomes is explained by the country-level context rather than by within-country regional differences. Classical Wald inference on the unpenalised logit is unreliable under separation; we therefore report block-bootstrap and LOCO-jackknife uncertainty for the regularised FE-logit. The coefficients reported for the logistic fixed-effects models should not be interpreted in the same causal or magnitude sense as those from the linear multilevel regressions. Because the outcomes were median-thresholded and predictors were standardised, the estimated weights reflect relative contributions to classification in log-odds space rather than average marginal effects on the original scale. Moreover, dummy-variable country fixed effects absorb substantial between-country heterogeneity, so coefficient signs may differ from those obtained in the continuous-outcome models without implying inconsistency. The logistic framework thus serves as a robustness and predictive check, not as a direct replication of the econometric coefficients.

**Table 6. Logistic Fixed-Effects Coefficients for Arrears and Warm**

| Variable | Coef (arrear) | p-value (bootstrap) | p-value (jackknife) | Coef (warm) | p-value (bootstrap) | p-value (jackknife) |
|---|---|---|---|---|---|---|
| intercept | -0.03 | 0.89 | 0.98 | -0.33 | 0.56 | 0.68 |
| gdp | 0.27 | 0.20 | 0.39 | -0.14 | 0.68 | 0.74 |
| tertiary_education | 0.36 | 0.42 | 0.27 | 0.20 | 0.58 | 0.68 |
| poverty | 0.11 | 0.20 | 0.48 | 0.27 | 0.42 | 0.55 |
| unemployment | 0.86 | 0.00 | 0.05 | 0.49 | 0.32 | 0.27 |
| unexpected_expenses | 1.47 | 0.00 | 0.03 | 1.40 | 0.00 | 0.00 |
| rooms | -0.67 | 0.08 | 0.01 | -0.33 | 0.50 | 0.60 |
| productivity | -0.49 | 0.00 | 0.36 | 1.11 | 0.39 | 0.38 |
| gdp_country | 0.79 | 0.64 | 0.69 | -1.11 | 0.00 | 0.09 |
| inflation_country | 1.96 | 0.14 | 0.25 | 0.35 | 0.02 | 0.22 |
| energy_ country | 0.12 | 0.81 | 0.76 | 1.06 | 0.00 | 0.02 |
| country_2 | -1.10 | 0.66 | 0.39 | 0.08 | 0.39 | 0.99 |
| country_3 | 0.73 | 0.62 | 0.46 | 0.48 | 0.46 | 0.38 |
| country_4 | 0.15 | 0.65 | 0.56 | 0.28 | 0.56 | 0.33 |
| country_5 | -1.37 | 0.67 | 0.35 | -0.41 | 0.35 | 0.59 |

*Source*: own calculations in Python.

After absorbing country heterogeneity via fixed effects, the L2-regularised logistic coefficients for regional socioeconomic variables are small and statistically indistinguishable from zero. Country-block bootstrap and LOCO-jackknife uncertainty quantify this lack of within-country signal. The empirical two-sided p-values are large for most of the features, and the confidence intervals straddle zero. These results align with SHAP analyses, which attribute most predictive mass to country indicators, indicating that macro-institutional context dominates cross-regional differences.

Regarding the penalised models, ridge shrinks magnitudes but preserves signs, while lasso selects only a small subset of predictors, both reflecting high collinearity and a limited within-country signal. Penalisation improves model stability but does not substantially increase predictive accuracy, reinforcing the dominance of country-level differences.

The LOCO-CV AUC values are modest, showing that the models trained in some countries transfer only partially to the unseen ones. This aligns with substantial policy and institutional heterogeneity across Europe and cautions against naive cross-country extrapolation of fitted relationships. Figure 1 shows ROC curves from leave-one-country-out cross-validation. The reduced AUC values reveal that models calibrated on one group of countries generalise only weakly to others, highlighting strong cross-country heterogeneity. These results explain why European EP cannot be modelled with pooled cross-country data without accounting for heterogeneity.

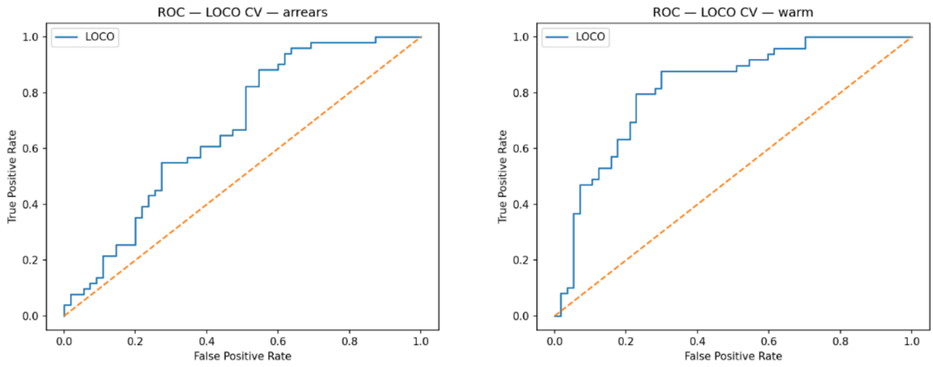


**Figure 1. LOCO-CV ROC Curve**
*Source*: Authors' own creation.

Table 7 shows high in-sample AUC values but noticeably lower LOCO-CV AUC values, especially for arrears. This indicates that models trained on a subset of countries do not generalise strongly to unseen countries, reinforcing the importance of cross-country heterogeneity.

**Table 7. Model Performance Summary**

| Target | Threshold | Train AUC logit | Ridge CV AUC | Lasso CV AUC | LOCO AUC |
|---|---|---|---|---|---|
| arrears | 9.7 | 0.98253119 | 0.97140496 | 0.97636364 | 0.67237077 |
| warm | 13.3 | 0.96813462 | 0.95198653 | 0.94835017 | 0.81740064 |

*Source*: own calculations in Python.

SHAP results from the logistic fixed-effects models show that, once country heterogeneity is controlled, regional socioeconomic variables contribute modestly to predictions. Because SHAP values are expressed in log-odds, their magnitudes cannot be directly compared to linear coefficients, but they remain directionally interpretable. Poverty risk, unexpected expenses, education, and housing space exert small within-country effects, while national institutional and macroeconomic factors dominate overall outcomes. Similarly, Gradient Boosting Machine SHAP patterns confirm this dominance: most predictive importance is assigned to country dummies, as the model prioritises highly discriminative features. The limited regional SHAP values thus reflect the overwhelming structural heterogeneity across Member States rather than a lack of substantive regional relevance.

**Table 8. SHAP Importance Values for Logistic FE Model**

| Variable | Mean abs SHAP Logistic FE | | Mean abs SHAP GBM | |
|---|---|---|---|---|
| | Arrears | Warm | Arrears | Warm |
| inflation_country | 1.53 | 0.27 | 1.30 | 1.51 |
| unexpected_expenses | 1.21 | 1.14 | 5.10 | 2.35 |
| gdp_country | 0.67 | 0.94 | 0.04 | 0.47 |

| Variable | Mean abs SHAP Logistic FE | | Mean abs SHAP GBM | |
|---|---|---|---|---|
| | Arrears | Warm | Arrears | Warm |
| unemployment | 0.56 | 0.32 | 0.10 | 0.34 |
| rooms | 0.49 | 0.24 | 0.18 | 0.16 |
| country_2 | 0.38 | 0.02 | 1.85 | 0 |
| tertiary_education | 0.28 | 0.15 | 0.36 | 0.83 |
| productivity | 0.24 | 0.55 | 0.68 | 0.87 |
| country_5 | 0.22 | 0.06 | 1.49 | 0 |
| gdp | 0.18 | 0.09 | 0 | 0.61 |
| country_3 | 0.18 | 0.11 | 0 | 0 |
| energy_country | 0.09 | 0.84 | 0.23 | 1.63 |
| poverty | 0.08 | 0.20 | 0.02 | 0.11 |
| country_4 | 0.05 | 0.09 | 0 | 0.02 |

*Source*: own calculations in Python.

SHAP values from the GBM model also reveal strong dominance of country dummies and confirm that nonlinear relationships do not override country-level structural effects, emphasising macro-institutional determinants.

SHAP results confirm that country effects account for a large share of predictive importance. Regional variables still contribute, especially unexpected expenses, poverty, housing conditions, and education, but their influence is modest relative to national structural factors. By decomposing predictions into country and regional contributions, SHAP clarifies this imbalance and shows consistent results across both linear and nonlinear models. These outcomes highlight that cross-country institutional heterogeneity is fundamental, and regional policies can only partially offset these structural constraints.

## 5. Discussion

Regional economic growth, education, Internet access, and productivity strengthen households' ability to manage energy costs and maintain warmth. National factors such as inflation, poverty, and unemployment also shape outcomes. Regional growth improves employment and stability, but national growth may distribute benefits unevenly, raising costs without matching income gains. Broader economic expansion can enhance thermal comfort by increasing purchasing power, as shown in Latin American evidence.

Higher education enhances financial stability, literacy, and awareness of energy efficiency, enabling households to manage utility costs and maintain adequate warmth. While both tertiary education and economic growth reduce EP, their joint effect is not linear. In advanced economies, EP shifts from affordability toward efficiency and sustainability concerns, where education matters more for behavioural adaptation than for basic energy access. Education also promotes digital inclusion, as more educated individuals value and effectively use Internet access for work, learning, and communication.

Economic growth fuels investment in telecommunications, expanding broadband and fiber infrastructure, and making Internet access more affordable. It also stimulates competition and innovation among providers, improving service quality. When combined with a well-educated population that both demands and supports technological advancement, this creates a strong synergy driving digital inclusion.

More rooms per person and higher labour productivity indicate stronger living standards and wages, improving households' ability to afford energy. On the other hand, inflation, unemployment, and poverty erode purchasing power and heighten EP, consistent with Paul and Sharma (2019) and Panzaru et al. (2025). Internet access, while increasing electricity use through greater device reliance, also promotes efficiency by providing information on energy-saving measures and enabling smart technologies that reduce waste. As shown by Zhang et al. (2023), digital expansion can thus both raise consumption and strengthen households' capacity to afford energy.

As a region becomes wealthier, there is often increased investment in infrastructure, including digital infrastructure. This can lead to better availability, affordability, and adoption of Internet services within the region, thereby enhancing internet access for its residents.

High national energy consumption reflects both inefficiencies in buildings and industries and the energy demands of advanced, digitised economies. Greater consumption can raise prices and reduce affordability for low-income households, while also signalling industrialisation and technological progress. Regional economic growth mitigates EP partly by expanding Internet access, which improves information flow and energy efficiency. However, this mediation is limited; growth also reduces EP directly through higher incomes, better housing, and policies that enhance affordability.

The machine-learning analysis supports the econometric results, confirming that country-level structural factors dominate NUTS 2 energy-poverty patterns. In addition, regional variables play a secondary role once the national context is considered. The combined framework of hierarchical econometrics and interpretable machine learning, using fixed-effects logistic, ridge, lasso, and gradient boosting, captures both linear and nonlinear dynamics under strong heterogeneity. Leave-one-country-out validation and SHAP analysis reinforce the robustness of findings and the predominance of macro-institutional effects, indicating structural rather than household-level drivers of variation.

Concluding, the findings suggest that energy poverty in Mediterranean Europe is shaped by an interaction between regional socioeconomic conditions and national macro-institutional environments. Regional prosperity, education, and housing conditions tend to reduce vulnerability, but national inflation, labour-market weakness, and structural differences between countries remain central determinants.

## 6. Conclusions

Regional employment, income, education, and financial stability directly reduce utility arrears. Macroeconomic pressures, such as inflation, uneven national growth, and persistent poverty or unemployment, intensify them. The contrast between national and regional growth effects indicates that economic gains are not evenly distributed, and national policies may fail to address regional disparities.

Local prosperity directly combats EP and improves digital connectivity. High national energy consumption, despite potentially indicating overall development and internet access, can exacerbate EP in case of inefficiency and high energy prices. The relationship between economic growth and EP is multifaceted. Although Internet access plays a role, many other socioeconomic and policy factors are critical in determining whether economic prosperity effectively translates into less EP.

This analysis, while comprehensive in its national scope, might not fully capture the nuances of EP at the household level. The definition of EP, relying only on arrears on utility bills and the inability to keep a home adequately warm, is limited. A more holistic understanding could be achieved by incorporating other proxies or an aggregated index.

Future research should adopt a multidimensional approach to EP, extending analysis from the regional to the household level to identify specific vulnerabilities and tailored solutions. Broader definitions of EP, incorporating composite indicators and qualitative methods such as interviews or focus groups, can better capture its human dimension. Additionally, assessing the role of emerging technologies, such as smart grids, energy storage, and evolving market designs, will be essential for effective policies.

***Acknowledgements:*** *This work was partially supported by the Romanian National Research, Development, and Innovation Plan 2022-2027, project PNRR/2022/C9/MCID/I8 no. 1672173292, contract no. 760231.*

## References

[1] Bardazzi, R., Charlier, D., Legendre, B., Pazienza, M.G. (2023), *Energy vulnerability in Mediterranean countries: A latent class analysis approach. Energy Economics*, 126, 106883.

[2] Bollino, C.A., Botti, F. (2017), *Energy poverty in Europe: A multidimensional approach. PSL Quarterly Review*, 70(283), 473-507.

[3] Castro-Cárdenas, C., Ibarra-Yunez, A. (2023), *Understanding the link between energy poverty and economic growth in Latin America. Contaduría y administración*, 68(2), e388. https://doi.org/10.22201/fca.24488410e.2023.4629.

[4] Churchill, S.A., Smyth, R. (2020), *Ethnic diversity, energy poverty and the mediating role of trust: Evidence from household panel data for Australia*. *Energy economics*, 86, 104663, https://doi.org/10.1016/j.eneco.2020.104663.

[5] Dong, K., Dou, Y., Jiang, Q. (2022), *Income inequality, energy poverty, and energy efficiency: who cause who and how? Technological Forecasting and Social Change*, 179, 121622, https://doi.org/10.1016/j.techfore.2022.121622.

[6] Halkos, G.E., Gkampoura, E.C. (2021), *Evaluating the effect of economic crisis on energy poverty in Europe. Renewable and Sustainable Energy Reviews*, 144, 110981, https://doi.org/10.1016/j.rser.2021.110981.

[7] Igawa, M., Managi, S. (2022), *Energy poverty and income inequality: An economic analysis of 37 countries. Applied Energy*, 306, 118076, https://doi.org/10.1016/ j.apenergy.2021.118076.

[8] Ishida, H. (2015), *The effect of ICT development on economic growth and energy consumption in Japan. Telematics and Informatics*, 32(1), 79-88, https://doi.org/10.1016/j.tele.2014.04.003.

[9] Ismail, Z., Khembo, P. (2015), *Determinants of energy poverty in South Africa. Journal of energy in southern Africa*, 26(3), 66-78, https://doi.org/10.17159/2413-3051/2015/v26i3a2130.

[10] Legendre, B., Ricci, O. (2015), *Measuring fuel poverty in France: Which households are the most fuel vulnerable? Energy Economics*, 49, 620-628, https://doi.org/10.1016/j.eneco.2015.01.022.

[11] Li, Y., Yang, X., Ran, Q., Wu, H., Irfan, M., Ahmad, M. (2021), *Energy structure, digital economy, and carbon emissions: evidence from China. Environmental Science and Pollution Research*, 28, 64606-64629.

[12] Lundberg, S.M., Lee, S.-I. (2017), *A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems* (NeurIPS 30), 4765-4774.

[13] Mawutor, J.K.M., Gborse, F.C., Agbanyo, R., Sogah, E. (2024), *Inflation and energy poverty: threshold effect of governance quality in Africa. Journal of Economic Studies*, 51(1), 222-237, https://doi.org/10.1108/JES-12-2022-0620.

[14] Mulder, P., Dalla Longa, F., Straver, K. (2023), *Energy poverty in the Netherlands at the national and local level: A multi-dimensional spatial analysis*. *Energy Research & Social Science*, 96, 102892, https://doi.org/10.1016/j.erss.2022.102892.

[15] Panzaru C, Belea S, Jianu L. (2025), *Delayed Taxation and Macroeconomic Stability: A Dynamic IS–LM Model with Memory Effects*. *Economies*, 13(7), 208, https://doi.org/10.3390/economies13070208.

[16] Paul, M., Sharma, P. (2019), *Inflation rate and Poverty: Does poor become poorer with inflation? Available at SSRN*: https://ssrn.com/abstract=3328539 or http://dx.doi.org/10.2139/ssrn.3328539.

[17] Primc, K., Dominko, M., Slabe-Erker, R. (2021), *30 years of energy and fuel poverty research: A retrospective analysis and future trends. Journal of Cleaner Production*, 301, 127003, https://doi.org/10.1016/j.jclepro.2021.127003.

[18] Simionescu, M., Cifuentes-Faura, J. (2024), *The digital economy and energy poverty in Central and Eastern Europe. Utilities Policy*, 91, 101841, https://doi.org/10.1016/j.jup.2024.101841.

[19] Statista (2025), *Share of households unable to keep their home adequately warm in the European Union in 2023, by country*. Data available at: https://www.statista.com/statistics/1260733/eu-energy-poverty-by-country/.

[20] Teixeira, I., Ferreira, A.C., Rodrigues, N., Teixeira, S. (2024), *Energy poverty and its indicators: A multidimensional framework from literature. Energies*, 17(14), 3445, https://doi.org/10.3390/en17143445.

[21] Usman, A., Ozturk, I., Hassan, A., Zafar, S.M., Ullah, S. (2021), *The effect of ICT on energy consumption and economic growth in South Asian economies: an empirical analysis. Telematics and Informatics,* 58, 101537.

[22] Wen, Z., Ye, B. (2014), *Analyses of mediating effects: the development of methods and models. Advances in Psychological Science*, 22(5), 731.

[23] Zhang, S.H., Yang, J., Feng, C. (2023), *Can internet development alleviate energy poverty? Evidence from China. Energy Policy*, 173, 113407, https://doi.org/10.1016/j.enpol.2022.113407.