

A Work Project, presented as part of the requirements for the Award of Master's degree in
Economics from the Nova School of Business and Economics.

VULNERABILITY TO ENERGY POVERTY IN PORTUGAL: DETERMINANTS AND
REGIONAL DIFFERENCES

TIL DIETRICH

Work project carried out under the supervision of:

Prof. Antonieta Cunha e Sá

04-01-2021

VULNERABILITY TO ENERGY POVERTY IN PORTUGAL: DETERMINANTS AND REGIONAL DIFFERENCES

Abstract

Despite having a mild climate, Portugal leads the list on excess winter mortality and is among the EU countries with the highest energy poverty level. This study explores which households' characteristics influence their vulnerability to being energy-poor and checks for regional differences. It employs 2015 HBS data in a logistic regression model, with the expenditure-based 2M indicator being the dependent variable. The results suggest that, on an aggregate level, retired, single individuals without children living in a large, rented detached house in a rural area are most vulnerable. Furthermore, there is regional variation along the tested dimensions.

Keywords: energy poverty, fuel poverty, vulnerable households, regional differentiation

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

*I would like to thank Professora Antonieta Cunha e Sá for her constant support and guidance.

1 Introduction

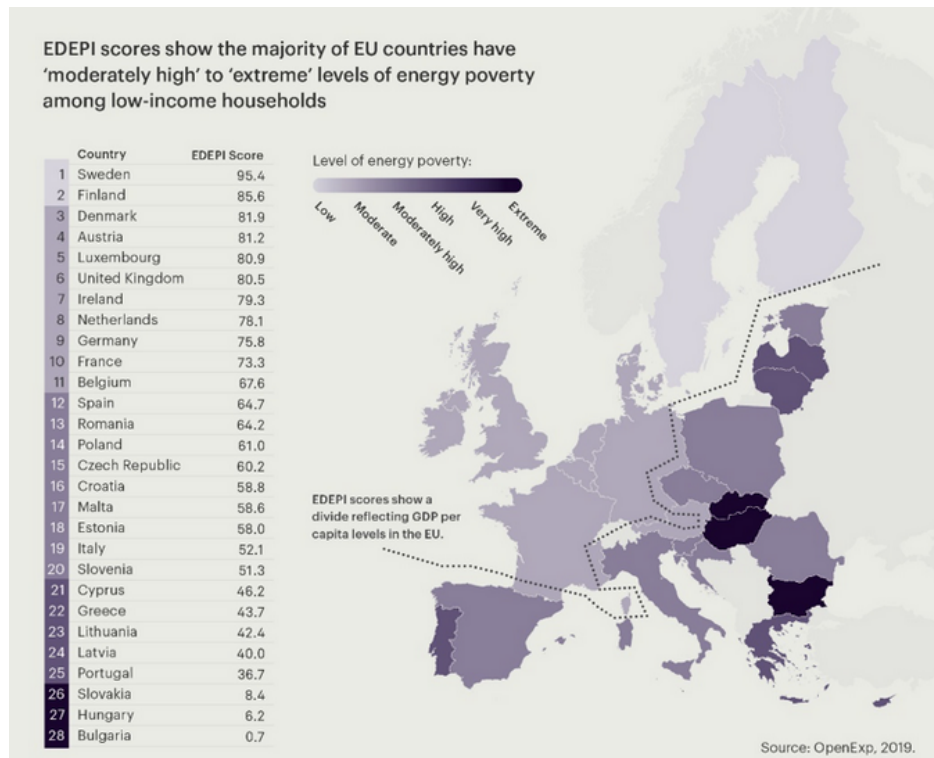
Since the release of “Council Directive 2009/72/EC” and “Council Directive 2009/73/EC” in 2009, in which the European Commission instructs Member States to address energy poverty (EP) in a national action plan, the issue has been widely recognized among European Union (EU) policymakers but also in the scientific community. However, a majority of Member States does not define EP, nor do they measure it or analyze its causes systematically (Bouzarovski and Thomson 2019). Research on the topic documents various adverse effects of EP on physical health such as circulatory diseases, respiratory and cardiovascular problems for various age groups, and arthritis and rheumatism among the elderly (Sovacool 2015; Liddell and Morris 2010; Shortt and Rugkåsa 2007; Evans et al. 2000). Moreover, evidence shows that psychological health is also at risk, resulting from various factors such as not being able to participate in cultural life or having to cut back on consumption expenditures (Dear and McMichael 2011).

In the face of global warming and the deterioration of the planetary environmental system, the transition towards a more sustainable energy system based on renewables is among the primary policy goals of the EU. However, not tackling the EP problem before or during that transition could have severe adverse social consequences when implementing, for example, renewable energy subvention schemes (Hiteva 2013) or a carbon tax (Berry 2019). Therefore, overcoming EP is essential from a poverty alleviation standpoint and for social equity in the energy sector’s on-going socio-technical transition. Addressing and fighting it contributes to achieving the United Nations Sustainable Development Goals (SDGs) (GA 2015). Most notably, number 1 (No Poverty), 3 (Good Health and Well-Being), and 7 (Affordable and Clean Energy). For example, the increasing electrification of heating and cooling systems, taking advantage of a higher share of renewable energies, could accommodate decentralized solutions to fix EP at the local level. This would also indirectly contribute to SDGs 10 (Reduced Inequalities) and 13 (Climate Action).

Despite having a relatively mild climate, Portugal is among the countries with the highest excess winter mortality rates in the EU (Healy 2003). This pattern has been related primarily to respiratory and circulatory diseases (Almendra et al. 2016), suggesting that cold indoor climate in winter plays a role. Figure 1 shows that Portugal ranked fourth on the level of energy poverty

among low-income households in 2019. Being ranked among and only topped by countries with considerably harsher climates. For many households, that means either feel cold in the winter months or save up a considerable amount of money throughout the year for heating.

Figure 1: Map of energy poverty levels in Europe, based on the Domestic Energy Poverty sub-Index (EDEPI), retrieved from EnAct (2020)



Even though there is evidence on this issue from various European countries, current research shows that the factors contributing to a likelihood of EP can differ substantially depending on the context. In Portugal, little information about which households are specifically vulnerable to being energy-poor exist. Structural factors such as low incomes, high electricity prices, and poor dwelling quality do not allow for an explicit targeting of households but only explain why EP is a general problem in Portugal. However, this information is essential for efficient policy interventions and the need for in-depth knowledge on the topic is recognized in Portugal's "Plano Nacional Energia E Clima 2021 – 2030 (PNEC 2030)" (Conselho de Ministros n.º 53/2020: Aprova o Plano Nacional Energia e Clima 2030 (PNEC 2030) 2020).

Therefore, this study investigates which factors are related to a household's higher EP vulnerability in Portugal. Moreover, it assesses whether there are systematic differences between various regions at the NUTS II level. It employs micro-level household data obtained from the

most recent household budget survey (HBS) for Portugal. The variable used to measure EP is an expenditure-based binary indicator that takes the value one if the percentage of energy expenditure in disposable income is at least twice as high as the national median. In line with a binary choice model, a logistic regression estimation is adopted to model the relationship between the indicator and the characteristics of interest.

This study suggests that roughly 16% of Portuguese households were energy poor in 2015, considering the expenditure-based indicator. The percentage of households in the autonomous regions of Açores and Madeira is statistically identical with the mainland. However, the share is below average in the densely populated (14.28%) and high GDP per capita (9.79%) regions. Correspondingly, they are higher in less densely populated and low GDP per capita regions.

The results for the whole country are broadly in line with the literature. They show that factors related to lower vulnerability are educational attainment, completing the secondary education cycle, which reduces the probability of being energy-poor by 2.9%, living with a partner, which leads to a lower probability by 2.4%, and having more rather than fewer kids, decreasing it by 3.7%. Furthermore, being retired, low dwelling quality, non-ownership of the accommodation, and a large detached or semi-detached house relate to higher vulnerability. However, some results are surprising and might be specific to the Portuguese context. For example, the analysis reveals no significant impact of unemployment or disability, on the aggregate level. Moreover, having a heating system seems to be connected to a higher vulnerability to be energy poor. This findings reveals a serious limitation of the employed 2M indicator.

Concerning regional differences, the study reveals that along the three analyzed regional dimensions of islands versus mainland, high population density versus low population density, and high GDP per capita versus low GDP per capita, some factors are more important in some regions than in others. For example, unemployed and disabled individuals are more vulnerable on the islands but not on the mainland. Additionally, the effects of being retired, living without a partner, in a flat, or having high renovation expenses are more substantial in the densely populated regions. The study observes similar patterns for low-income regions. Moreover, educational attainment returns seem to be related to average education levels in a given region. Returns to education are lower in high-income areas, where average educational levels are high

and higher on the islands, where the opposite occurs. These findings imply that the vulnerability profile differs between regions. This knowledge is vital for public policy purposes because it indicates clearly that policies should be region-specific and not spatially uniform.

The remainder of this paper is organized as follows. Section 2 defines EP, discusses its measurement and previous findings in a comprehensive literature review. Section 3 describes the data used. Section 4 develops the models employed in this study. Section 5 reports the empirical results, and section 6 discusses them, provides some concluding remarks, and gives policy recommendations. The appendix shows graphs and tables.

2 Literature review

2.1 What is Energy Poverty?

There are diverging meanings assigned to the notion of energy poverty (EP) in the literature. Traditionally, energy poverty (EP) refers to problems resulting from a household's inadequate access to energy in the realm of development economics, covering infrastructural, socio-economic, health, educational and environmental matters (Bouzarovski and Petrova 2015; Bazilian, Nakhooda, and Van de Graaf 2014). The first acknowledgements of the problem date back to the 1970s (Johnson 1974). More recently, EP increasingly contains issues that were formerly under the domain of fuel poverty (FP) not only in the scientific literature but also in policy-making. FP, first acknowledged in the 1980s (Osbaldeston 1984, Lewis 1982), was recognized in the scientific mainstream and among policy-makers, especially in the UK and Ireland, with the release of Boardman's (1991) seminal book on the topic. The concept is predominantly used in the developed world to discuss insufficient space heating problems due to the relative unaffordability of energy stemming from low household incomes or energy inefficient dwellings (Bouzarovski and Petrova 2015). However, in the recent scientific literature and the policy-making domain, the terms are used interchangeably. The caption EP gained traction in the European Union, which set up the EU Energy Poverty Observatory (EPOV) to tackle the problem. Energy vulnerability, another term related to the problem, is defined as "the propensity of an individual to become incapable of securing a materially and socially needed

level of energy service in the home" (Bouzarovski and Thomson 2019). This idea acknowledges that it is not enough to focus on supply-side issues but that it is also essential to consider characteristics that are internal to the households, such as energy demand, efficiency, and economic poverty. Also, it treats EP as a transitional state that is the result of those characteristics (Castaño-Rosa et al. 2019; Hall, Hards, and Bulkeley 2013). There is not one single definition of EP. However, following EPOV ("EU Energy Poverty Observatory: What is energy poverty?" 2018) and Bouzarovski and Petrova (2015), EP is here defined as household energy deprivation that not only leads to adverse effects on health and the standard of living but also enhances social exclusion and does not allow citizens to fulfill their full cultural, social as well as economic potential. Moreover, this paper considers the demand side and above-mentioned transitional state by focusing on household-specific characteristics.

2.2 Measuring Energy Poverty

Given its multidimensionality, there is no single indicator that measures all aspects of EP (Simcock, Walker, and Day 2016). The main approaches to define EP are expenditure-based and the consensual-based. Outcome-based approaches that directly measure an outcome of interest also exist but have significant drawbacks concerning comparability and are often too narrow (Rademaekers et al. 2016).

Expenditure-based approaches rely on a measure of energy expenditure relative to a household's income. The first expenditure-based metric is the 10% threshold established by Boardman (1991). It identifies a household as energy poor if it has an energy expenditure above 10% of disposable income. However, such absolute thresholds are highly sensitive to energy price changes (Hills 2012) and contain a high degree of arbitrariness (Liddell et al. 2012). Therefore, relative thresholds that take into account income and energy expenditure distributions are preferable. The 2M indicator is such a measure and classifies a household as energy poor if its share of energy expenditure in income exceeds twice the national median (Rademaekers et al. 2016). In addition to indicators that identify excessively high shares of energy costs, the concept of Hidden Energy Poverty (HEP) developed. It captures households whose energy expenditure is so low that it is insufficient to meet basic needs (KBF (King Baudouin Foundation)

2015). The M/2 indicator is most commonly used for this type of energy poverty. It identifies a household as energy poor if the total energy expenditure is below half the national median. The EPOV selects both the 2M and the M/2 indicator as the two expenditure-based metrics to arise at a common understanding of EP on the European level and increase comparability between scientific research (Thema and Vondung 2020). Both can be derived from the Household Budget Surveys (HBS), which are conducted by every Member State.

Consensual-based approaches are based on surveys asking citizens directly. The European Survey on Income and Living Conditions (EU-SILC), for example, contains questions on whether a household can keep the home adequately warm, whether it manages to pay its utility bills, or asks about particular housing conditions. These indicators have the advantage of being less complicated and measure the perception of households as well as their experience (Rade-maekers et al. 2016). However, they often deviate from the more objective expenditure-based measures. Price, Brazier, and Wang (2012) identify that consensual-based measures involve problems of self inclusion and -exclusion. They find that only about 60% of households with electricity bills above the threshold report that they cannot keep the house adequately warm. Additionally, of those that reported feeling energy-poor, less than half were captured by the expenditure-based measure. However, the strengths of consensual-based indicators to capture a household's needs more subjectively has been recognized by EPOV. It includes the inability to keep the house adequately warm as well as arrears on utility bills as the two recommended consensual-based metrics. The literature shows that the indicators sometimes deviate substantially in their measurement. This study employs the 2M indicator because it best captures the increased economic stress that households face based on their energy expenditure and contributes towards a common understanding of EP as proposed by EPOV (Thema and Vondung 2020). The M/2 indicator measures a considerably different aspect of EP and is, therefore, not included in the analysis.

2.3 Factors affecting Energy Poverty exposure

Belaïd (2018) uses a logistic regression model to examine the salient determinants for EP in France. He uses the French micro data set on energy consumption PHEBUS. His results suggest

that the most critical drivers are family size, property ownership, employment status, nationality, dwelling size, heating system, and residence renovation. Two striking results are that, according to the author, the odds of being energy-poor are twice as high for families with five or more occupants than for a single-person household and that vulnerability increases with declining dwelling size. Those findings contradict other conclusions in the literature and are, according to the author, linked to working with energy costs per m² in the indicator.

Likewise, Thomson and Snell (2013) assess factors leading to EP at the EU level, also using logistic regression. Their analysis adopts three consensual-based indicators obtained from EU-SILC. They find that residing in a rural area has the largest impact on predicted EP for all indicators. According to the authors, this links to a higher number of poor neighborhoods in rural regions. Moreover, their findings suggest that living in a smaller, detached, and rented accommodation is related to a higher probability of being energy-poor.

Similarly, Legendre and Ricci (2015) explore EP determinants in France using three different indicators. They employ a logit, a complementary log-log, and a mixed effect logit model. Based on these models, they identify that living alone, particularly when being retired, having more kids, using canned gas for cooking and poor insulation make a household more likely to be energy-poor. However, being highly educated, a home-owner, and having a collective heating system decreases vulnerability.

Consortium (2008) conduct a study for the determinants of EP in five EU countries. They identify three main drivers that the countries have in common: low income, dwellings, and energy prices. Within those categories, households, receive social security payments, part-time workers, and highly indebted individuals are energy poor more often. Moreover, the absence of a central heating system, defective insulation, and high humidity in the dwellings heighten EP.

Snell, Bevan, and Thomson (2015) examine the relationship between disability and EP in England using cross-tabulation and logistic regression. They identify that households with disabled individuals usually have a higher EP propensity and that rates are exceptionally high for single working-age people with disabilities.

Recalde et al. (2019) analyze the underlying structural political and socio-economic conditions that determine the degree of EP in EU countries by constructing a structural energy

poverty vulnerability index (SEPV). They identify that south-eastern European countries have a high structural vulnerability, and north-western have a low structural vulnerability for EP. According to the authors, the main structural determinants are income levels and inequality, unemployment, and low levels of government expenditure on social protection and public interventions.

Mashhoodi, Stead, and Timmeren (2019) explore the question of spatial homogeneity and heterogeneity of the EP determinants. They do so by employing a geographically-weighted regression on 2474 neighborhoods in the Netherlands. Their results suggest that the percentage of low-income households and pensioners' share in a neighborhood are spatially homogeneous drivers. However, they find that household size, unemployment, building age, rented properties, number of summer days, and frost days are spatially heterogeneous. Their research suggests that an analysis of EP in a country should allow for heterogeneity of determinants across regions.

This study adds to the literature on three dimensions. First, it assesses the determinants of EP vulnerability in Portugal on the primary expenditure-based indicator as developed by EPOV. Second, it examines spatial variation by checking whether determinants differ across various meaningful regions. Third, by using the proposed indicator, it contributes towards a standard methodology for understanding EP.

3 Data

This study employs household-level micro-data from the most recent household budget survey Inquérito às Despesas das Famílias (IDEF), conducted by Statistics Portugal (INE) every five years. It is the national survey that underlies the European household budget surveys (HBS). This study's analysis relies on the most recent publication from 2017, which entails data collected between March 2015 and March 2016. The survey aims at providing information on dwelling characteristics, household composition, income, and expenses. Additionally, it contains the household's location at the NUTS II level¹. IDEF 2015 comprises cross-sectional data from 11 398 households, with the central unit of observation being the household while some

1. The NUTS system hierarchically divides the EU territory into regions, mainly for the framing of regional policies and socio-economic analyses.

essential individual characteristics are available. The survey methodology follows a standard HBS procedure, with conducting one interview before and after the recording period to ensure the data's validity. Sampling follows a two-stage randomization procedure. In the first stage, the primary sampling unit (PSU) is drawn at random from a grid of $1km^2$ sectors within the respective NUTS II regions based on the number of accommodations in a section. The secondary unit is randomly selected within the PSU. The econometric model uses a sampling weight constructed from those to adjust for the probability of selection. A limitation of the dataset for this study is that dwelling characteristics such as, for example, age and quality of the building are not available. An ideal dataset to test for determinants of EP would contain those physical building characteristics. This limits the ability to account exactly for dwelling characteristics, but a variable to approximate dwelling quality is constructed and employed in the study.

Table A1 depicts the levels and frequencies of the categorical variables, including the dependent variable. For the analysis, the variables *highest educational attainment* and *number of rooms* are dummy-coded as they consist of more than two categories. The category with more observations must be the baseline for those variables when working with a dummy variable and its interaction terms. Otherwise, there is high structural multicollinearity, which causes inflated standard errors and, therefore, insignificant results where they should not be. Table A2 shows the descriptive statistics for the continuous variables. It reveals that there are no missing values in the sample. The variables for disability, unemployment, retirement, and gender are constructed as shares of total household members (or adults) to decrease the correlation with other variables related to the household's size. As mentioned above, no variable in the dataset captures the dwelling age or quality. Therefore, this study uses the share of equivalized renovation expenditures in disposable income as an approximation². Because the data for disposable household income per person is highly skewed, the current analysis uses that variable's logarithmic transformation. Usually, this is unnecessary because logit models do not assume a normal distribution in the independent variables. However, it improves the fit of the model. Moreover, variables that contain a measure of disposable income or expenditures are equivalized to adjust

2. Equivalization is a procedure to obtain representative household income and expenses, taking into account household composition and size. This study employs the commonly used "modified-OECD" scale (OECD (Organisation for Economic Co-operation and Development) 2013).

for different household sizes. Per person household income is an exception. The following section introduces the econometric model and details the construction of the EP indicator.

4 Methodology

4.1 The model

This study employs a logistic regression model to assess which factors are related to a household's higher vulnerability to be energy-poor. There are various reasons why using logistic regression is best in the context at hand. Firstly, the dependent variable is a binary indicator, taking the value one if a household is energy-poor and zero if it is not. Logistic regression models restrict the predictions to that range. Secondly, through the derivation of marginal effects, it is possible to check how a change in the independent variables affects the dependent variable's probability. In this context, this suits nicely as an increase in the probability of being energy-poor can be interpreted as a rise in a household's vulnerability. Thirdly, introducing a regional dummy with its interaction terms to allow for varying slopes and intercepts in different areas is reasonably straightforward. The logit model takes the following general form:

$$P[y_i = 1|x_i] = \frac{e^{\alpha + x_i\beta}}{1 + e^{\alpha + x_i\beta}},$$

where y_i is the energy poverty indicator, x_i is a vector of regressors, β is a vector of coefficients and α is the constant. The set of regressors differs depending on the varying model specifications and includes a set of control variables. The coefficients in a logistic regression denote the change in the log-odds of the event $Y = 1$, here energy poverty, for a change in a respective regressor. Because the interpretation log-odds is not straightforward, this study reports the average marginal effect, which shows the change in the relative probability of the event $Y = 1$ for a small change in the independent variable, assuming that all other variables are at their average. Marginal effects are essentially partial derivatives with respect to the regressor in question and

take the following general form:

$$ME_j = \frac{\partial P[y_i = 1|x_i]}{\partial x_j}.$$

It is impossible to compute marginal effects for interaction terms because there is no way to change the interaction term directly, only its elements. In other words, it is impossible to partially derive with respect to an interaction term, but solely with respect to its parts. Therefore, this study's primary tool of analysis is the average marginal effect (AME) at the two respective values that a regional dummy can take (0 or 1). The AME describes the effect that a marginal change in x_i has on Y , given that all other variables are at their respective average.

An assessment of the Variance Inflation Factor (VIF) to check for multicollinearity reveals that there is some degree of multicollinearity. However, it is structural multicollinearity, meaning that it results from the inclusion of the regional dummies and their interaction terms. This problem would cause the affected variables' standard errors to be inflated and wrongly reject some coefficients as insignificant. Nevertheless, an alternative model specification with all regressors centered at the mean removes the multicollinearity problem and reveals that the variables in question have identical coefficients and significance levels. The reason for not using this model as the main specification is that the non-interaction coefficients' interpretation is not straightforward. Nonetheless, multicollinearity is not a problem here. The model is estimated using Maximum Likelihood (ML). Therefore, the critical assumption is that the distribution of Y conditional on all x follows a logit distribution. If that is the case, the estimators are consistent and efficient. Another decisive factor is that the data come from a survey. Accordingly, the probability of being selected differs between households based on the above-mentioned two-stage sampling strategy. The dataset contains sampling weights that reflect the inverse probability of being selected under this procedure, which reflect how many households a respective observation represents in the finite population. The sampling weights are assigned to each household in the estimation procedure to adjust for this.

4.2 Dependent variable

The dependent variable in this study is the 2M indicator for EP. It considers a household to be EP if its percentage of energy expenditure in disposable income is higher than twice the national median. This study follows Thema and Vondung (2020) to calculate the indicator based on IDEF2015 data. First, household income and energy expenditure have to be equivalized using the OECD scale to adjust for household size and composition. Subsequently, the equivalized share of energy expenditure in disposable income is obtained. Then, the sampling-weighted median is calculated, and each household that exceeds twice that value is considered energy-poor. The indicator, being a binary dependent variable, is defined as:

$$y_i = \begin{cases} 1, & \text{if the household is energy poor} \\ 0, & \text{if not} \end{cases}$$

4.3 Independent variables

This study uses 13 independent variables, and three different dummies for the respective regional aspect analyzed in the model estimation. The interaction terms of those dummies and the independent variables are also part of the respective models. The individual characteristics are *highest educational attainment*, *share of disabled individuals* and *share of unemployed adults*. *Highest educational attainment* is a categorical variable with four categories (see Table A1) that shows the highest educational level that any of the household members has completed. The category basic serves as the baseline when estimating the model. *Share of disabled individuals* measures what percentage of household members has a disability which entitles them to receive a disability pension, *share of unemployed adults* the percentage of adults that are currently unemployed and *share of retired individuals* the proportion of retirees. The reason to use shares for those variables is to remove correlation with other household size variables.

This study considers various household characteristics. One is *number of children*, a continuous variable measuring how many children under the age of 15 live in the household. Another one is *partner*, a dummy variable that takes the value one if two of the household members are either married or live in a registered partnership. Rented housing is also considered a household

characteristic. It is a dummy variable that takes the value one if the accommodation is rented instead of owned by its inhabitants. A direct measure of occupants' number is not included in the analysis because the indicator is equivalized. It is already adjusted for household size, and obtained results would be biased.

The dwelling characteristics include *heating system*, *flat* and *piped gas*, which are all dummy variables that take the value one if central heating exists, if the accommodation is a flat or has a piped gas connection, respectively. Size of housing is measured by the categorical variable *number of rooms*, containing the three categories: less than three rooms, 3 – 4 rooms, and more than four rooms. The baseline level in the logistic regression is 3 – 4 rooms as it has most observations and can be considered standard given the frequency of 66.82%. As previously mentioned, a drawback of the dataset is that no variables measuring the building's quality are available. Therefore, this study uses the constructed variable *share of renovation expenditures in disposable income* to proxy it.

To also incorporate whether the household is in an urban or rural area, the dummy variable *rural area* is included, taking the value one if it is located in a rural region. *Log of disposable income per person* and *share of female adults*, as a measure of household gender, serve as control variables.

4.4 Regions

This study explores regional differences along three dimensions with different model specifications. As the dataset contains information about in which NUTS II region a household is located, it is straightforward to construct dummy variables to differentiate between the regions of interest. The study examines the following three differences: it explores a significant difference in determinants between the Portuguese mainland and the two autonomous island regions Açores and Madeira exists, assesses whether the most densely populated regions, Norte and the Lisbon metropolitan area, are different from the rest, and compares determinants along the dimension of GDP per capita. The two regions considered having an exceptionally high GDP per capita are the Lisbon metropolitan area and Algarve as they are both above the Portuguese average (“PORDATA: GDP per capita 2018” 2020). The variables *islands*, *densely pop.* and *high*

GDP take the value one if the household is located in the respective regions and zero otherwise.

5 Results

The data suggest that 16.07% of Portuguese households are energy-poor under the 2M indicator definition. In the autonomous regions of Açores and Madeira, this share is slightly lower, with 15.71%. With 14.28%, the densely populated areas show an even lower level, and in the high GDP per capita regions, the share is at only 9.79%. Tables A4, A5 and A6 show the weighted relative frequencies of energy-poor households within the respective regions.

5.1 Baseline – Portugal

Table A7 depicts the two baseline logistic regression models' estimation results, which consider the whole country. It is necessary to mention that pseudo- R^2 's in the context of non-linear models cannot be compared with their counterparts in linear models as the variance is fixed. Therefore, in contrast to its linear counterpart, it does not depict the dependent variable's explained variance. However, it is a widely-used goodness-of-fit measure. McFadden (1977), the creator of the measure, argues that any value between 0.2 and 0.4 represents an "excellent model fit". Model 1 and 2 both fulfill that criterium. The percentage of correctly predicted observations serves as another goodness-of-fit measure.

Model 1 reveals a puzzling result. It suggests that the existence of a heating system is related to a 4.9% higher probability of being EP and is highly robust to changes in the model specification. This finding is firmly at odds with the previous literature on the topic. A possible explanation is the choice of the indicator as results can vary. Antepara et al. (2020) show that the 2M indicator systematically fails to identify households without a heating system as energy-poor in Southern European countries. This relates to the finding of model 1. In this case, instead of failing to identify households as energy-poor, this study finds a positive relation. That is an undoubtedly interesting finding and adds to the discussion on EP indicators. However, this study excludes *heating* from the analysis because the chosen indicator seems unfit to measure the relationship.

Accordingly, model 2 is estimated without it. McFadden R^2 is now 0.2203, which is slightly lower but still in the acceptable range. Moreover, 85.84% of cases are correctly predicted by model 2 in comparison to 85.67% by model 1. This represents a negligible decrease. Moreover, the significance and magnitude of most effects is unaltered. Therefore, excluding *heating* does not weaken the model's explanatory power and interpretation despite typically being considered relevant.

Column 5 of Table A7 reports the relevant AMEs. The results suggest that the individual characteristics *highest educational attainment* and *share of retired* are statistically significant at a reasonable significance level. Compared to the completion of the primary education cycle, completing the secondary education cycle is, on average, related to a 4.3% lower probability of being energy-poor, and completing higher education leads to a 2.6% lower probability. These AMEs are significant at 1% and 10%, respectively. There is no difference between no education and finishing the primary education cycle. Moreover, a higher share of retired individuals is, on average, related to a 3.1% higher EP probability, being significant at 1%. Unemployment and disability show positive AMEs but are not statistically significant.

In contrast, all household characteristics, are statistically significant at 1% or 5%. On average, renting versus owning the accommodation relates to a 1.9% higher probability of being energy-poor, and living together with a partner decreases it by 3.6%. Both findings are firmly in line with the literature as rented accommodations and living alone are typically related to higher EP vulnerability. A small change in the number of children in the household decreases the probability on average by 5.4%. However, the interpretation of this result should be taken cautiously. Figure A1 shows the marginal effects at each possible value for the variable and reveals that the effect is non-linear. It gradually declines with the number of children, meaning that the effect is largest (-5.9%) for the first child and smallest for the 11th (-0.1%), which is also only significant at 10%. Also, due to the reasons mentioned in the data section, a variable for household size cannot be included, and the number of children is capturing that fact. These findings together provide suggestive evidence that, foremost, the increase in family size from the first child leads to lower EP vulnerability. A household that is located in a rural area, on average, has a 2.2% higher probability of being EP and the effect is significant at 5%. This re-

sult finds strong support in the literature, as most studies show higher EP vulnerability in rural regions.

Similarly, the AMEs for dwelling characteristics are statistically significant at 1%, except for the first category of *number of rooms* and *share of repair exp.*, both being significant at 5%. Accordingly, compared to a flat with 3 – 4 rooms, small accommodations with less than three rooms are related to a 2.6% lower and large residence with more than four rooms to a 4.2% higher probability of EP, on average. Furthermore, having a gas pipeline connection, on average, corresponds to a 5.3% lower probability. Moreover, a marginally higher share of renovation expenditure in disposable income relates, on average, to a 12% higher probability. This result suggests that dwelling quality is a significant and meaningful factor for EP vulnerability. If the accommodation is a flat as opposed to a house, the probability is 3.5% lower.

5.2 Islands

Model 3 allows for effects to differ between the Portuguese mainland and the islands. Column 3 and 4 of Table A8 show the AMEs in the respective regions. McFadden's R^2 is 0.2219, and the model correctly predicts 85.6% of observations.

The results suggest that a household living on the islands is, on average, 3.8% less likely to be energy-poor. Furthermore, some factors seem to be more pronounced in one region than in the other. Unemployment and disability significantly affect the EP vulnerability of households on the Açores and Madeira, but not on the mainland. A marginally higher share of unemployed individuals is, on average related to a 5% higher probability of EP. For the share of disabled individuals in the household, it is 6.8%. Moreover, the effect of higher education is more extensive and more significant on the islands. It is, on average, linked to a 6.4% lower probability. Additionally, living in a flat is not related to a lower probability of EP on the islands while it is on the mainland. There it corresponds to a 3.7% lower relative probability. Moreover, the smaller size of accommodations is only significant on the mainland. Compared to an average-sized (3 – 4 rooms) accommodation, a dwelling with less than three rooms corresponds, on average, to a 2.6% lower and large housing with more than four rooms to a 4.3% higher probability of being energy-poor. On the islands, the effect for larger accommodations is smaller with

2.8% and less significant. Furthermore, having a gas pipeline connection is, on average, related to a 5.3% lower probability on the mainland but has no impact on the islands even though they exist in some parts. Finally, compared to an urban region, living in a rural area leads to a, on average, 2.1% higher probability for EP on the mainland. All other AMEs imply that there is no significant difference between the respective determinants in the two regions. These results suggest that, on average, individual characteristics, like education, unemployment, and disability, play a larger role on the islands than they do on the Portuguese mainland. Whether it is a flat or has a pipeline connection, dwelling characteristics, including size, are more important on the mainland and do not seem to matter on the islands.

The reason for this is likely the difference in the way of living. A comparison of the characteristics in question explains the divergence. Average household income is slightly lower on the islands, but the difference is not statistically significant and is unlikely to be the critical determinant for the spatial variation. The relative lack of importance of residing in a rural area could be explained by the fact that it is much more common on the islands, with 33% as opposed to 27%. Furthermore, the education level is generally lower on the islands, making higher education more outstanding. Another striking difference is that living in a detached house is quite common on the islands as 75% of households live in one, while it is only 50% on the mainland. Additionally, not having a gas pipeline is standard on the islands as only 18% of households are connected as opposed to 35% on the mainland. Moreover, it is much more common on the islands to live in larger dwellings. Therefore, those characteristics are less distinguishing factors in describing a household on the islands.

5.3 Population density

Model 4 assesses regional differences between the two high population density regions Norte and Metropolitan Region of Lisbon and the others. Columns 3 and 4 of Table A9 present the AMEs for the two regions. McFadden's R^2 is 0.2228, and the model correctly predicts 85.61% of observations.

On average, the probability of being energy-poor is not different between the two regions. However, some determinants vary in significance, which suggests they may differ. In the high-

density regions, living together with a partner, on average, relates to a 5% lower relative probability of being energy-poor. Similarly, renting an accommodation corresponds, on average, to a 2.9% higher and living in a flat to a 5.2% lower probability in densely populated regions. Moreover, a marginally higher share of retired individuals is, on average, related to a 4% higher EP probability, but only in the densely populated regions. Additionally, dwelling quality, measured as the share of equivalized renovation expenditures in disposable income, is linked to a 16.8% higher probability in high-density regions. Living in a rural region has no significant effect in the densely populated areas, while it relates to a 3% higher probability in the others. This finding links to wealthier neighborhoods being located in the suburban and rural regions of densely populated areas. Likewise, the effect of living in a small accommodation is insignificant in high-density regions but corresponds to a 3.6% lower probability in low-density areas. The significance of the other factor's AMEs does not differ between regions and are not statistically significantly different from each other, suggesting that their influence is similar. Factors related to the housing situation, such as living in a flat as opposed to a house, whether the property is rented, if the partner also lives there and a higher share of retired individuals seem to only have an effect in the high-density regions. Similarly, a bad dwelling quality is more problematic in those areas, as the abovementioned difference in significance shows.

5.4 GDP per capita

Model 5 distinguishes between high GDP per capita regions (Metropolitan Region of Lisbon and Algarve) and the rest of the country. Columns 3 and 4 of Table A10 depict the AMEs in the two regions. McFadden's R^2 is 0.2228, and the model correctly predicts 85.73% of observations.

In general, the probability of being energy-poor is similar. Interestingly, many factors that are significant in the rest of the country do not relate to EP vulnerability in the high GDP per capita areas. For example, the effect of educational attainment seems to be less pronounced in those areas. While having completed higher education is, on average, related to a 3% lower probability in the rest of the country, it is insignificant in the high-income areas. This links to the fact that education levels are generally greater in high GDP regions. Interestingly, these re-

sults follow the same logic as the observation in the previous section. Lower levels of education on the islands relate to a higher return to education. Similarly, the return is lower in high GDP areas. Moreover, ownership appears to be no factor in high GDP areas, but renting accommodation corresponds to a 2.2% higher probability of being energy-poor in the lower-income regions. Dwelling quality does not seem to affect energy poverty in high-income regions significantly. However, it has a large and significant effect in low-income areas, where it relates to as high as a 15.2% higher probability of EP. A possible explanation for this divergence is that households in high-income regions compensate for bad dwelling quality by heating and can afford it. At the same time, this does not happen in other areas. Compared to a 3 – 4 room housing, smaller accommodations with less than three rooms (4.1% lower), rural areas (1.8% higher), and gas pipeline connection (5.8% lower) are significant in low-income regions and have no statistically significant effect in the high-income regions. The difference in residing in a rural area could be similar to that observed for densely populated regions in that wealthier neighborhoods are located in rural areas within those high-income regions. Moreover, a gas pipeline is more common in high-income regions, which could cause the relative lack of importance. All other AMEs do not statistically significantly differ from each other. Therefore, the factors are considered to have the same influence.

6 Discussion

The performed analysis reveals specific distinct characteristics that increase EP vulnerability for a household in Portugal. In the following, this paper summarizes the most critical characteristics into high vulnerability profiles, compares them to the literature, and highlights important considerations.

The high vulnerability profile for an average Portuguese household points towards a retired, single individual without children. Also, the individual has only completed the basic cycle of education. Moreover, this household would rent a large rather than small detached house of low quality without gas connection in a rural area. The above described AMEs suggest that particularly the type of dwelling is an important driver. A comparison with section 2.3 reveals that this profile matches those of most other studies but is at odds with Belaïd (2018), which can

be explained by the different indicator choice. However, it is interesting to note that this paper does not find any significant effect of unemployment or disability for the average Portuguese household. A possible reason is that the analysis controls for income per person, which is also suggested by findings from Pocinho (2020).

Furthermore, the analysis shows considerable differences in the high vulnerability profiles between the Portuguese mainland and the islands. In the mainland, the profile is similar to the Portuguese average, except that the magnitude of some effects changes. High vulnerability on the islands, however, is characterized differently. Most notably, disability and unemployment strongly contribute to higher vulnerability. This finding can be related to the low number of recipients of social electricity support payments. The data in this study stems mainly from 2015. As of July 2016, a new automatic assignment procedure exists. In the following year, the number of recipients in the Açores increased by 243.6% and on Madeira by 218%. The increase in the mainland was less pronounced, indicating that those households on the islands were underrepresented in that period (Martins et al. 2019). As households with unemployed or disabled individuals are eligible for this support, the observation explains the higher vulnerability of those households on the islands in the observed period. It would be interesting to assess whether the discrepancy was eliminated with the change by employing the 2020 HBS as soon as it gets released. Another factor is that the returns to education may be higher due to the lower average education level. Interestingly, dwelling characteristics like living in a large detached house, having no piped gas connection, and being located in a rural area do not point towards higher vulnerability on the islands. This is likely related to the difference in lifestyles as living in a rural area is more common. For example, in areas like Funchal, there are a large number of detached houses, even in urban regions. Those findings are in line with Mashhoodi, Stead, and Timmeren (2019), who points toward the spatial heterogeneity of, for example, unemployment, dwelling quality, and size. The study shows that they are likely to be connected to different ways of living, like living in urban and non-urban areas.

Additionally, there are some differences in the high vulnerability profiles between low-density and high-density regions. Low education levels and no children characterize low-density region profiles. It is likely that those low-density regions overlap with rural areas,

where those aspects are common. The opposite holds for high-density regions. Also, the effect of dwelling size is very pronounced as small accommodations relate to lower and large accommodations to higher vulnerability. The non-existence of a gas pipeline and being located in a rural area are other characteristics of that area. In the high-density regions, the picture is slightly different. Here, being single, uneducated, and having no children increases vulnerability. Furthermore, the household is characterized by a high share of retired individuals. This links to older people typically living in worse conditions in more densely populated areas. Additionally, non-ownership of the housing and living in a low-quality detached house without a gas pipeline are contributing factors. Those characteristics are typical for relatively poor, urban neighborhoods. Another explanation is that dwellings in less densely populated areas commonly have fireplaces for heating purposes and are, therefore, less reliant on piped gas and can afford to heat these places. Furthermore, retired individuals in those often rural regions survive on a subsistence type economy. Moreover, only larger accommodations are related to high vulnerability, while living in a small residence has no influence. This could be related to disproportionately high renting costs for small housing in densely populated regions. A comparison of the two vulnerability profiles suggests that factors related to the accommodation's size and ownership status are more pronounced in the densely populated regions, and so is being single or retired. Usually, people in sparsely populated areas live in their own houses and are old. Therefore, renting costs do not weigh in so strongly. However, in the densely populated regions, retirees typically live in rented accommodation, and the costs weigh-in highly in lower budgets.

Finally, vulnerability profiles considerably differ between high GPD per capita regions and the rest of the country. For the low-income regions, the profile is in line with the Portuguese average. However, high-income regions differ substantially. The profile is characterized by only three factors: being single, having no children, and living in a detached house. This is an interesting finding as it suggests that the reasons for EP are mostly structural in those regions. Also, dwelling characteristics such as quality, size, and ownership play no role. A possible explanation for this is the lower number of poor neighborhoods in these regions, characterized by low quality, small and rented accommodations. Furthermore, as already mentioned in section

5.4, bad dwelling quality can be compensated by increased heating, which can be affordable in regions where average incomes are higher.

These findings imply that, generally speaking, Portuguese household characteristics that amplify the vulnerability to being energy-poor are similar to those in other countries. In the face of the excessive EP levels, this could either suggest more significant structural problems, such as low incomes and electricity prices, which are not assessed in this household-level analysis. Nonetheless, it could also mean that the identified factors are more pronounced in the Portuguese context and, therefore, the higher EP levels prevail. This could be assessed through a comparative cross-country analysis. Another important implication is that aggregate studies without regional differentiation are incomplete and do not adequately represent underlying geographical disparities. The findings represented in this paper add to a developing body of literature that emphasizes the importance of spatial inequalities in the context of EP alleviation (e.g., Robinson, Bouzarovski, and Lindley 2018; Bouzarovski and Simcock 2017).

Portugal's "Plano Nacional Energia E Clima 2021 – 2030 (PNEC 2030)" explicitly mentions the importance of the alleviation of EP for an inclusive and fair transition towards a decarbonized energy system. It aims at targeting the most vulnerable household and states that "for an effective combat, it is necessary to know the national reality in greater detail in order to target the measures more effectively" (Conselho de Ministros n.º 53/2020: Aprova o Plano Nacional Energia e Clima 2030 (PNEC 2030) 2020, 109). This study contributes to the goal of enhancing knowledge about the Portuguese context to inform policymakers in the process of developing a public policy strategy to combat inequality and promote social inclusion, which is planned to be completed by the end of 2021 (Conselho de Ministros n.º 53/2020: Aprova o Plano Nacional Energia e Clima 2030 (PNEC 2030) 2020). The local governments involved in this process should be aware of particular drivers of EP in their respective regions as they are likely to differ. This study shows that incorporating the spatial dimension into the process of designing EP alleviation programs is essential for effectively targeting the most vulnerable households. Moreover, from a financing perspective, it is necessary to make the program cost-effective and economically efficient, which can only be achieved by targeting those households that are indeed most vulnerable. Additionally, this study suggests that closely monitoring local

circumstances is vital and that large scale aggregate national programs are likely to be inefficient.

There are some limitations regarding the results in this paper. One is that the study employs a variable that not directly measures the quality or age of a dwelling but approximates it through renovation expenditures. This might misrepresent the effect for two main reasons. First, some households might live in low-quality dwellings but might not be able to afford to fix the building. Second, typically tenants are not responsible for renovating their accommodation, but their landlords are the ones bearing the costs and possibly pass it on through higher rents. Therefore, the magnitude of this effect should be interpreted with caution. Nonetheless, the direction of the effect is unlikely to be altered. Another limitation is the choice of indicator on two dimensions. First, the expenditure-based 2M indicator cannot identify households as energy-poor that cut back on their energy expenditures. This problem is similar to the previously mentioned argument about renovation costs. Second, the indicator employs equivalized expenditure and income data. This means that it is already adjusted for household size. Coefficients from variables, like the number of children, might now be overestimated. That is because an additional household member might not increase energy consumption by the same amount as it decreases equivalized expenditures. Studies testing multiple indicators on a dataset with more differentiated household information are likely to yield more robust results.

7 Conclusion

Portugal's problem with high EP levels and the related excess winter mortality is well-documented. Policymakers, instructed by the EU, have to develop mitigation policies. This is essential for poverty alleviation and a vital component of a fair and just decarbonization transmission of the energy system. Also, it strongly contributes to achieving the UN SDGs. A sound understanding of the characteristics that make a Portuguese household vulnerable is vital for designing such policies. Therefore, this paper analyzed and interpreted which determinants shape a household's vulnerability profile and whether those differ between various regions. The analysis shows that retired, uneducated single individuals without children are particularly vulnerable, on average. Further characteristics are non-ownership of the residence, living in a

larger detached house of low dwelling quality, and not having a gas pipeline connection. At the aggregate Portuguese level, disability and unemployment cannot be linked with EP. However, the study reveals that there are considerable inequalities in vulnerability characteristics between regions. Most notably, unemployed and disabled individuals living on the islands belong to the risk group, while they do not on the mainland. A possible reason for this are differences in the distribution of energy support schemes during the measurement period. Newer data might therefore offset or balance the found differences. Additionally, dwelling-related characteristics are less relevant on the islands, likely due to a different, more rural way of living and milder climate. A comparison of low population density and high population density regions reveals that in the latter, factors related to accommodation ownership, size, and quality are essential whereas they are of minor importance in the former. Usually, those characteristics are prevalent in poor urban neighborhoods. Finally, energy-poor households in high-income regions are characterized by single individuals, those with a small family, living in large dwellings. Other dwelling characteristics are insignificant, which is likely connected to the lower proportion of poor neighborhoods that would contain a high share of dwellings with the mentioned characteristics.

This study contributes to the developing body of literature on energy poverty. It enhances the understanding of characteristics that relate to a Portuguese household's energy poverty vulnerability. Additionally, it shows a regional divergence of those characteristics in the Portuguese context and underlines that a household's location shapes the relevance of certain factors. Therefore, it directly contributes to the Portuguese government's efforts to enhance knowledge about the topic to inform the design of efficient EP alleviation policies.

Given the considerable spatial differences, further research should incorporate this dimension as an aggregate analysis that may smoothen out characteristics that are only relevant in specific underrepresented regions. Furthermore, research on hidden energy poverty in Portugal using the M/2 indicator would add meaningful knowledge. Moreover, it could focus on smaller regions than NUTS II regions as the areas in this study are still relatively broad. An analysis at the municipality level would significantly enhance neighborhood-specific knowledge on energy poverty. This is left for future research.

References

- Almendra, Ricardo, Paula Santana, Elisabete Freire, and João Vasconcelos. 2016. "Seasonal mortality patterns and regional contrasts in Portugal." *Bulletin of Geography. Socio-economic Series* 32 (32): 7–17.
- Antepara, Iñigo, Lefkothea Papada, João Pedro Gouveia, Nikolas Katsoulakos, and Dimitris Kaliampakos. 2020. "Improving Energy Poverty Measurement in Southern European Regions through Equivalization of Modeled Energy Costs." *Sustainability* 12 (14): 5721.
- Bazilian, Morgan, Smita Nakhooda, and Thijs Van de Graaf. 2014. "Energy governance and poverty." *Energy Research & Social Science* 1:217–225.
- Belaïd, Fateh. 2018. "Exposure and risk to fuel poverty in France: Examining the extent of the fuel precariousness and its salient determinants." *Energy Policy* 114:189–200.
- Berry, Audrey. 2019. "The distributional effects of a carbon tax and its impact on fuel poverty: A microsimulation study in the French context." *Energy Policy* 124:81–94.
- Boardman, Brenda. 1991. *Fuel poverty: from cold homes to affordable warmth*. Pinter Pub Limited.
- Bouzarovski, Stefan, and Saska Petrova. 2015. "A global perspective on domestic energy deprivation: Overcoming the energy poverty–fuel poverty binary." *Energy Research & Social Science* 10:31–40.
- Bouzarovski, Stefan, and Neil Simcock. 2017. "Spatializing energy justice." *Energy Policy* 107:640–648.
- Bouzarovski, Stefan, and Harriet Thomson. 2019. "TRANSFORMING ENERGY POVERTY POLICIES IN THE EUROPEAN UNION: SECOND ANNUAL REPORT OF THE EUROPEAN UNION ENERGY POVERTY OBSERVATORY."

- Castaño-Rosa, Raúl, Jaime Soís-Guzmán, Carlos Rubio-Bellido, and Madelyn Marrero. 2019. "Towards a multiple-indicator approach to energy poverty in the European Union: A review." *Energy and Buildings* 193:36–48.
- Conselho de Ministros n.º 53/2020: Aprova o Plano Nacional Energia e Clima 2030 (PNEC 2030), Resolução do. 2020.
- Consortium, EPEE. 2008. "Diagnosis of Causes and consequences of fuel poverty in Belgium, France, Italy, Spain and United Kingdom." *Work Package 2*.
- Dear, Keith BG, and Anthony J McMichael. 2011. *The health impacts of cold homes and fuel poverty*.
- EnAct. 2020. "OVER-EXPOSED: Energy Poverty in Centra and Eastern Europe." Accessed: 2020-12-21. <http://www.coldathome.today/overexposed-energy-poverty-in-central-eastern-europe>.
- "EU Energy Poverty Observatory: What is energy poverty?" 2018. Accessed: 2020-12-07. <https://www.energypoverty.eu/about/what-energy-poverty>.
- European Parliament, "Directive 2009/72/EC of the, of the Council of 13 July 2009 concerning common rules for the internal market in electricity, and repealing Directive 2003/54/EC". 2009.
- European Parliament, "Directive 2009/73/EC of the, of the Council of 13 July 2009 concerning common rules for the internal market in electricity, and repealing Directive 2003/55/EC". 2009.
- Evans, Julie, Sophie Hyndman, Sarah Stewart-Brown, David Smith, and Sophie Petersen. 2000. "An epidemiological study of the relative importance of damp housing in relation to adult health." *Journal of Epidemiology & Community Health* 54 (9): 677–686.
- GA, UN. 2015. "Transforming our world: the 2030 Agenda for Sustainable Development." *Division for Sustainable Development Goals: New York, NY, USA*.

- Hall, Sarah Marie, Sarah Hards, and Harriet Bulkeley. 2013. "New approaches to energy: equity, justice and vulnerability. Introduction to the special issue." *Local Environment* 18 (4): 413–421.
- Healy, John D. 2003. "Excess winter mortality in Europe: a cross country analysis identifying key risk factors." *Journal of Epidemiology & Community Health* 57 (10): 784–789.
- Hills, John. 2012. "Getting the measure of fuel poverty: Final Report of the Fuel Poverty Review." *CASereport* 72.
- Hiteva, Ralitsa Petrova. 2013. "Fuel poverty and vulnerability in the EU low-carbon transition: the case of renewable electricity." *Local Environment* 18 (4): 487–505.
- Johnson, JR. 1974. "Ceramics in advanced energy systems." In *Ceramic Engineering and Science*, 75–91. Springer.
- KBF (King Baudouin Foundation). 2015. *The Energy Poverty Barometer*. Brussels.
- Legendre, Bérangère, and Olivia Ricci. 2015. "Measuring fuel poverty in France: Which households are the most fuel vulnerable?" *Energy Economics* 49:620–628.
- Lewis, Paul. 1982. *Fuel poverty can be stopped*. National Right to Fuel Campaign.
- Liddell, Christine, and Chris Morris. 2010. "Fuel poverty and human health: a review of recent evidence." *Energy policy* 38 (6): 2987–2997.
- Liddell, Christine, Chris Morris, SJP McKenzie, and Gordon Rae. 2012. "Measuring and monitoring fuel poverty in the UK: National and regional perspectives." *Energy policy* 49:27–32.
- Martins, Rita, Patrícia Pereira da Silva, Micaela Antunes, and Adelino Fortunato. 2019. *ES-TUDO SOBRE A APLICAÇÃO DA TARIFA SOCIAL DE ENERGIA EM PORTUGAL*. Coimbra.

- Mashhoodi, Bardia, Dominic Stead, and Arjan van Timmeren. 2019. "Spatial homogeneity and heterogeneity of energy poverty: a neglected dimension." *Annals of GIS* 25 (1): 19–31.
- McFadden, Daniel. 1977. *Quantitative methods for analyzing travel behavior of individuals: some recent developments*. Institute of Transportation Studies, University of California Berkeley.
- OECD (Organisation for Economic Co-operation and Development). 2013. *OECD framework for statistics on the distribution of household income, consumption and wealth*. Chap. 8. OECD Publishing.
- Osbaldeston, Joan. 1984. "Fuel poverty in UK cities." *Cities* 1 (4): 366–373.
- Pocinho, Maria Dias. 2020. "ENERGY POVERTY IN PORTUGAL: WHICH HOUSEHOLDS ARE THE MOST VULNERABLE?" Master's thesis, Master Thesis - Nova SBE.
- "PORDATA: GDP per capita 2018." 2020. Accessed: 2020-12-13. [https://www.pordata.pt/en/Municipalities/GDP+per+capita+\(2016\)-896](https://www.pordata.pt/en/Municipalities/GDP+per+capita+(2016)-896).
- Price, Catherine Waddams, Karl Brazier, and Wenjia Wang. 2012. "Objective and subjective measures of fuel poverty." *Energy Policy* 49:33–39.
- Rademaekers, Koen, Jessica Yearwood, Alipio Ferreira, Steve Pye, Ian Hamilton, Paolo Agnolucci, David Grover, Jiř Karásek, and Nataliya Anisimova. 2016. "Selecting indicators to measure energy poverty." *Trinomics: Rotterdam, The Netherlands*.
- Recalde, Martina, Andrés Peralta, Laura Oliveras, Sergio Tirado-Herrero, Carme Borrell, Laia Palència, Mercè Gotsens, Lucia Artazcoz, and Marc Marí-Dell'Olmo. 2019. "Structural energy poverty vulnerability and excess winter mortality in the European Union: Exploring the association between structural determinants and health." *Energy Policy* 133:110869.
- Robinson, Caitlin, Stefan Bouzarovski, and Sarah Lindley. 2018. "'Getting the measure of fuel poverty': The geography of fuel poverty indicators in England." *Energy Research & Social Science* 36:79–93.

- Shortt, Niamh, and Jorun Rugkåsa. 2007. ““The walls were so damp and cold” fuel poverty and ill health in Northern Ireland: results from a housing intervention.” *Health & place* 13 (1): 99–110.
- Simcock, Neil, Gordon Walker, and Rosie Day. 2016. “Fuel poverty in the UK: beyond heating.” *People, Place and Policy* 10 (1): 25–41.
- Snell, Carolyn, Mark Bevan, and Harriet Thomson. 2015. “Welfare reform, disabled people and fuel poverty.” *Journal of Poverty and Social Justice* 23 (3): 229–244.
- Sovacool, Benjamin K. 2015. “Fuel poverty, affordability, and energy justice in England: Policy insights from the Warm Front Program.” *Energy* 93:361–371.
- Thema, Johannes, and Florian Vondung. 2020. *EPOV Indicator Dashboard: Methodology Guidebook*.
- Thomson, Harriet, and Carolyn Snell. 2013. “Quantifying the prevalence of fuel poverty across the European Union.” *Energy Policy* 52:563–572.

Appendix

Table A1: List and description of categorical variables

Variable	Categories	Frequency (in %)
Energy poor (2M indicator)	No	83.36
	Yes	16.64
Living with partner	No	54.94
	Yes	45.06
Highest educational attainment	None	7.39
	Basic	46.87
	Secondary	21.85
	Higher	23.89
Rented housing	No	77.17
	Yes	22.83
Accommodation is a flat	No	57.92
	Yes	42.08
Number of rooms	< 3	7.12
	3 – 4	66.82
	> 4	26.07
Existence of central heating	No	89.37
	Yes	10.63
Rural area	Urban	68.01
	Rural	31.99

Table A2: Descriptive statistics of continuous variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Share of unemployed adults	11.398	0.1016	0.2252	0	1
Share of disabled individuals	11.398	0.0299	0.1344	0	1
Share of retired individuals	11.398	0.2969	0.4061	0	1
Number of children	11.398	0.382	0.7194	0	11
Share of renovation expenditures in disposable income	11.398	0.0099	0.061	0	2.48
Log of disposable income per person	11.398	0.0146	0.6461	4.94	12.78
Share of female adults	11.398	0.4704	0.324	0	1

Table A3: Regional dummies

Variable	Categories	Frequency (in %)
Islands	Mainland	76.14
	Islands	23.86
High population density	Low density	58.37
	High density	41.63
High GDP per capita	Low GDP per capita	67.79
	High GDP per capita	31.21

Table A4: Weighted cross-tabulation of energy poor households per region

		Region		
		Mainland	Islands	Total
Not energy poor	Count	9,148.38	418.47	9,566.84
	% within region	83.92%	84.29%	83.93%
Energy poor	Count	1,753.15	78.00	1,831.16
	% within region	16.08 %	15.71%	16.07%
Total	Count	10,901.53	496.47	11,398
	% within region	100.00%	100.00%	100.00%

This cross-tabulation was constructed by incorporating the sample weights to adjust for how many households in the finite population are represented by a respective observation. Therefore, the counts include decimals and do not correspond to the number of households that were sampled in a respective region.

Table A5: Weighted cross-tabulation of energy poor households per region

		Population density		
		Low	High	Total
Not energy poor	Count	3,566.22	6,000.62	9,566.84
	% within region	81.09%	85.72%	83.93%
Energy poor	Count	831.83	999.32	1,831.16
	% within region	18.91%	14.28%	16.07%
Total	Count	4,398.05	6,999.948	11,398
	% within region	100.00%	100.00%	100.00%

This cross-tabulation was constructed by incorporating the sample weights to adjust for how many households in the finite population are represented by a respective observation. Therefore, the counts include decimals and do not correspond to the number of households that were sampled in a respective region.

Table A6: Weighted cross-tabulation of energy poor households per region

		GDP per capita		
		Low	High	Total
Not energy poor	Count	6,208.00	3,358.84	9,566.84
	% within region	80.89%	90.21%	83.93%
Energy poor	Count	1,466.58	364.58	1,831.16
	% within region	19.11%	9.79%	16.07%
Total	Count	7,674.58	3,723.42	11,398
	% within region	100.00%	100.00%	100.00%

This cross-tabulation was constructed by incorporating the sample weights to adjust for how many households in the finite population are represented by a respective observation. Therefore, the counts include decimals and do not correspond to the number of households that were sampled in a respective region.

Figure A1: Average marginal effects of the number of children on the probability to being energy-poor

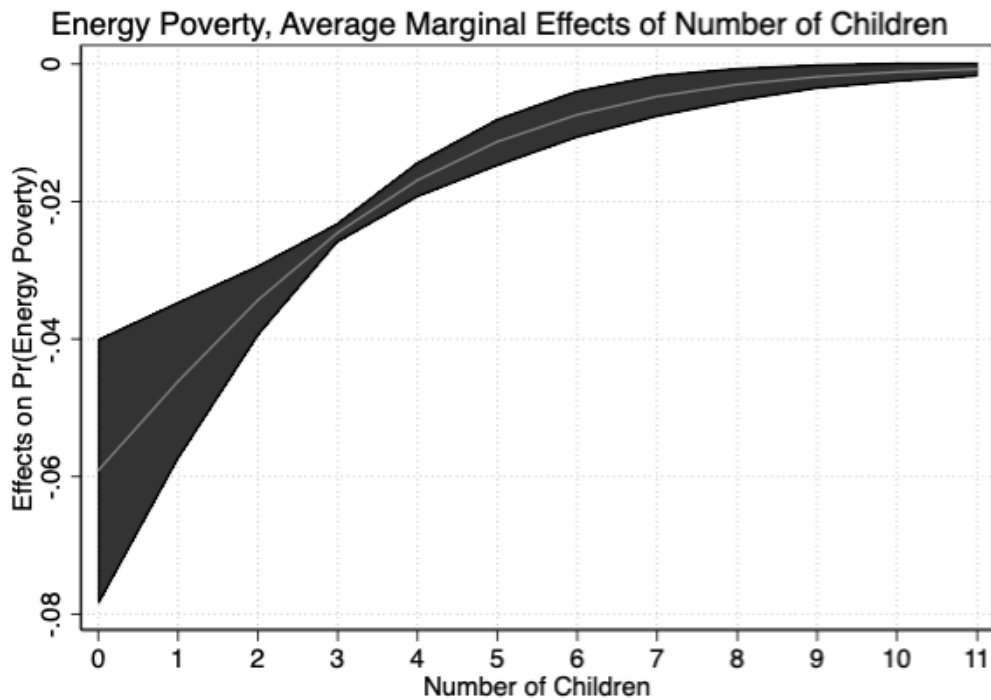


Table A7: Results of logistic regression models

	Model 1		Model 2	
	β (SE)	AME (SE)	β (SE)	AME (SE)
<i>partner</i>	−0.348*** (0.079)	−0.037*** (0.008)	−0.337*** (0.078)	−0.036*** (0.008)
<i>educational attainment</i>				
None	0.125 (0.128)	0.015 (0.015)	0.093 (0.127)	0.011 (0.015)
Secondary	−0.463*** (0.105)	−0.047*** (0.01)	−0.421*** (0.105)	−0.043*** (0.01)
Higher	−0.285** (0.129)	−0.03** (0.013)	−0.238* (0.129)	−0.026* (0.013)
<i>share of unemployed</i>	0.188 (0.169)	0.02 (0.018)	0.129 (0.168)	0.014 (0.018)
<i>share of disabled</i>	0.377 (0.231)	0.04 (0.024)	0.31 (0.232)	0.033 (0.025)
<i>share of retired</i>	0.297*** (0.111)	0.031*** (0.012)	0.295*** (0.11)	0.031*** (0.012)
<i>number of children</i>	−0.521*** (0.081)	−0.055*** (0.008)	−0.51*** (0.079)	−0.054*** (0.008)
<i>rented</i>	0.186** (0.09)	0.02** (0.009)	0.183** (0.09)	0.019** (0.01)
<i>flat</i>	−0.342*** (0.097)	−0.036*** (0.01)	−0.333*** (0.097)	−0.035*** (0.01)
<i>share of repair exp.</i>	1.122** (0.456)	0.118** (0.048)	1.134** (0.458)	0.12** (0.049)
<i>number of rooms</i>				
Less than 3	−0.22 (0.136)	−0.021* (0.012)	−0.271** (0.135)	−0.026** (0.012)
More than 4	0.361*** (0.094)	0.04*** (0.011)	0.377*** (0.094)	0.042*** (0.011)
<i>piped gas</i>	−0.516*** (0.116)	−0.054*** (0.012)	−0.502*** (0.116)	−0.053*** (0.012)
<i>rural</i>	0.198** (0.088)	0.021** (0.009)	0.206** (0.088)	0.022** (0.009)
<i>heating system</i>	0.467*** (0.081)	0.049*** (0.009)		
Intercept	17.228*** (0.826)		16.94*** (0.817)	
McFadden R ²	0.2255		0.2203	
% correctly predicted	85.67%		85.64%	
N	11,398		11,398	

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Baseline categories: Basic education; 3 – 4 rooms

Table A8: Results of Model 3: Islands

	Model 3		
	β	AME	
	(SE)	(SE)	
		Mainland	Islands
<i>partner</i>	−0.342*** (0.082)	−0.036*** (0.009)	−0.032*** (0.012)
<i>educational attainment</i>			
None	0.078 (0.132)	0.009 (0.016)	0.042 (0.029)
Secondary	−0.413*** (0.11)	−0.043*** (0.011)	−0.05*** (0.014)
Higher	−0.22* (0.133)	−0.024* (0.014)	−0.064*** (0.02)
<i>share of unemployed</i>	0.109 (0.176)	0.012 (0.019)	0.05** (0.024)
<i>share of disabled</i>	0.277 (0.246)	0.03 (0.026)	0.068** (0.028)
<i>share of retired</i>	0.293** (0.114)	0.03** (0.012)	−0.002 (0.02)
<i>number of children</i>	−0.505*** (0.083)	−0.054*** (0.009)	−0.046*** (0.009)
<i>rented</i>	0.163* (0.094)	0.018* (0.01)	0.039** (0.016)
<i>flat</i>	−0.353*** (0.1)	−0.037*** (0.01)	−0.019 (0.017)
<i>share of repair exp.</i>	1.092** (0.46)	0.117** (0.049)	0.241*** (0.092)
<i>number of rooms</i>			
Less than 3	−0.266* (0.141)	−0.026** (0.013)	−0.024 (0.019)
More than 4	0.379*** (0.098)	0.043*** (0.011)	0.028* (0.016)
<i>piped gas</i>	−0.526*** (0.119)	−0.053*** (0.011)	−0.008 (0.021)
<i>rural</i>	0.195** (0.092)	0.021** (0.01)	0.02 (0.014)
<i>islands</i>	−0.501** (0.215)	−0.038*** (0.008)	
Intercept	17.048*** (0.824)		
Regional interactions included	Yes		
Controls included	Yes		
McFadden R ²	0.2219		
% correctly predicted	85.6%		
N	11,398		

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Baseline categories: Basic education; 3 – 4 rooms

Control variables: Age, income and gender

Table A9: Results of Model 4: Population density

	Model 4		
	β	AME	
	(SE)	(SE)	
		Low density	High density
<i>partner</i>	−0.172 (0.114)	−0.019 (0.012)	−0.05*** (0.011)
<i>educational attainment</i>			
None	0.053 (0.173)	0.006 (0.021)	0.018 (0.022)
Secondary	−0.515*** (0.153)	−0.054*** (0.015)	−0.034** (0.014)
Higher	−0.285 (0.197)	−0.032 (0.021)	−0.023 (0.017)
<i>share of unemployed</i>	0.037 (0.239)	0.004 (0.026)	0.022 (0.024)
<i>share of disabled</i>	0.253 (0.316)	0.028 (0.034)	0.037 (0.035)
<i>share of retired</i>	0.186 (0.16)	0.02 (0.018)	0.04** (0.016)
<i>number of children</i>	−0.57*** (0.117)	−0.062*** (0.012)	−0.05*** (0.01)
<i>rented</i>	0.054 (0.132)	0.006 (0.015)	0.029** (0.013)
<i>flat</i>	−0.047 (0.154)	−0.005 (0.017)	−0.052*** (0.013)
<i>share of repair exp.</i>	0.445 (0.683)	0.049 (0.074)	0.168*** (0.057)
<i>number of rooms</i>			
Less than 3	−0.375* (0.196)	−0.036** (0.018)	−0.02 (0.017)
More than 4	0.278** (0.127)	0.032** (0.015)	0.052*** (0.016)
<i>piped gas</i>	−0.584*** (0.194)	−0.06*** (0.018)	−0.041*** (0.014)
<i>rural</i>	0.24** (0.118)	0.03** (0.013)	0.02 (0.015)
<i>densely pop.</i>	−0.072 (0.223)	−0.004 (0.008)	
Intercept	17.08*** (0.84)		
Regional interactions included	Yes		
Controls included	Yes		
McFadden R ²	0.2228		
% correctly predicted	85.61%		
N	11,398		

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Baseline categories: Basic education; 3 – 4 rooms

Control variables: Age, income and gender

Table A10: Results of Model 5: GDP per capita

	Model 5		
	β	AME	
	(SE)	Low GDP	High GDP
<i>partner</i>	−0.297*** (0.091)	−0.032*** (0.01)	−0.051*** (0.015)
<i>educational attainment</i>			
None	0.075 (0.147)	0.009 (0.018)	0.024 (0.028)
Secondary	−0.44*** (0.122)	−0.046*** (0.012)	−0.031* (0.018)
Higher	−0.276* (0.155)	−0.03* (0.016)	0.008 (0.021)
<i>share of unemployed</i>	0.103 (0.199)	0.011 (0.022)	0.021 (0.029)
<i>share of disabled</i>	0.201 (0.157)	0.022 (0.027)	0.063 (0.053)
<i>share of retired</i>	0.32** (0.13)	0.035** (0.014)	0.025 (0.019)
<i>number of children</i>	−0.524*** (0.094)	−0.056*** (0.01)	−0.048*** (0.011)
<i>rented</i>	0.2* (0.105)	0.022* (0.012)	0.017 (0.017)
<i>flat</i>	−0.253** (0.117)	−0.027** (0.012)	−0.04** (0.018)
<i>share of repair exp.</i>	1.399** (0.62)	0.152** (0.067)	0.035 (0.069)
<i>number of rooms</i>			
Less than 3	−0.431** (0.177)	−0.041*** (0.015)	0.011 (0.02)
More than 4	0.321*** (0.104)	0.04*** (0.012)	0.069*** (0.023)
<i>piped gas</i>	−0.573*** (0.153)	−0.058*** (0.014)	−0.027 (0.017)
<i>rural</i>	0.165* (0.097)	0.018* (0.011)	0.034 (0.022)
<i>high GDP</i>	−0.328 (0.243)	−0.016 (0.01)	
Intercept	16.972*** (0.831)		
Interactions included	Yes		
Controls included	Yes		
McFadden R ²	0.2228		
% correctly predicted	85.73%		
N	11,398		

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Baseline categories: Basic education; 3 – 4 rooms

Control variables: Age, income and gender