Norwegian School of Economics

Bergen, spring 2023

Energy poverty in the European Union

Cross-country patterns and vulnerability

Anna Chekalyuk & Veronika Priakhina

Supervisor: Johannes Mauritzen

Master Thesis, Economics and Business Administration
Majors: Business Analytics & Energy, Natural Resources and the Environment

NORWEGIAN SCHOOL OF ECONOMICS

Executive Summary

This thesis is a quantitative study based on the data gathered from Eurostat. The thesis investigates energy poverty by observing several sides of the problem: geographical distribution in the European Union, cross-country pattern similarities in the EU, and vulnerability of European households to energy poverty, especially when energy prices are unprecedentedly high.

The analysis is performed with the help of such statistical methods as Principal Component Analysis (PCA) and Hierarchical Clustering (HC). According to PCA, the first four Principal Components out of fourteen are sufficient for the analysis since they explain 79% of the variance in the data. Later, HC is applied to those four identified Principal Components, showing that it is optimal to divide the EU countries into seven categories by their predisposition and susceptibility to risks associated with energy poverty. Further, the translog regression approach, along with the HC, is adopted to make a model with an interaction term comprised of the cluster and household electricity price variables to assess the electricity price elasticity of household energy consumption.

This thesis is inspired by similar studies conducted by Recalde et al. (2019) and Chai et al. (2021). However, the paper proposes a different way of tracking energy poverty across Europe, based on social, economic, environmental and energy indicators. The findings of this thesis suggest that the neighboring counties' sensitivity to energy poverty tends to be similar, and southern European states are noticeably more vulnerable to the severe effects of energy poverty.

Keywords: cluster, electricity, energy, energy consumption, energy poverty, Europe, European Union, Hierarchical Clustering, household, price elasticity, Principal Component Analysis

Contents

List of Abbreviations

Abbreviation	Explanation
EP	Energy Poverty
EPAH	Energy Poverty Advisory Hub
EU	European Union
HC	Hierarchical Clustering
IEWB	Index of Economic Well-Being
MEPI	Multidimensional Energy Poverty Index
PCA	Principal Component Analysis
PVE	Proportion of Variance Explained
SEPV	Structural Energy Poverty Vulnerability

List of Figures

List of Tables

List of Appendices

1. Introduction

Energy availability is a prerequisite for achieving a comfortable standard of living. However, in the global reality, when many routine things are digitized and electrified, more than 750 million people worldwide did not have access to electricity in 2021. And this number is only expected to increase (Cozzi et al., 2022).

Energy poverty arises when energy costs take up a large portion of a consumer's income or when people are forced to cut back on energy use to the point that it severely affects their health and well-being (European Commission, n.d., a). This fact illuminates the complexity of the issue. For, in addition to having a private nature and affecting households, energy poverty is distinguished by being at the intersection between multiple dimensions, including but not limited to household income, energy expenditure and housing conditions. Therefore, energy poverty persists as a massive concern to be tackled in the EU (European Commission, n.d., a; Faiella & Lavecchia, 2019).

Energy poverty manifests itself in different ways. For example, almost 35 million EU people, or around 8% of the EU's total population, struggled to keep their houses adequately warm in 2020. A precarious situation for vulnerable social strata in Europe is likely to have become direr with the current energy crisis, which was amplified by geopolitical unrest and the COVID-19 crisis fallout (European Commission, n.d., a).

Energy poverty is notably uneven, with the Central and Eastern European (CEE) region and the former Soviet Republic countries (CIS region) exhibiting the highest incidence of energy poverty among the population. While scientists and decision-makers are starting to pay greater attention to the substantial variances across European regions, the problem of energy poverty is still poorly understood (Turai et al., 2021). That is why in light of the present-day situation, in our master's thesis we decided to discern the cross-country energy poverty pattern similarities in the European Union and analyze the effect of soaring power prices on households across Europe.

Structurally, the answer to the above question is divided into two parts. First, we distinguish geographical energy poverty vulnerability patterns and similarities in energy poverty trends in the EU. In order to do so, we explore a range of energy poverty indicators proposed by the Energy Poverty Advisory Hub (European Commission, n.d., b). We then conduct a Principal

Component Analysis (PCA) along with Hierarchical Clustering (HC), which ultimately allows us to group European countries according to their energy poverty susceptibility. Second, we examine the price elasticity of household energy consumption across the clusters we have identified, thus focusing on energy consumption trends in the EU. We use changes in consumption as a proxy for energy poverty vulnerability, and we consider changes in electricity prices as one of the main external factors that shape households' behavior. We also include both internal and external factors in the regression run on our observed clusters. Finally, all the performed analyses allow us to describe similarities in how European countries are exposed to fluctuations in energy prices.

The theoretical Section 2 precedes the practice-oriented Sections 3-5 aimed at quantitatively answering the research question, with Section 6 showing the main findings, limitations, and suggestions for further research.

Section 2 covers the main background aspects of this thesis. The section explores the ambiguities around the definition of energy poverty, the societal implications of the issue at hand, and the difficulty in gauging it. Moreover, the section touches upon the geographical disparities in the issue's appearance and how they are tied to the economic security of European households.

Section 3 begins with a demonstration of the data chosen for the study. We utilize energy poverty indicators, economic factors, and household data (Eurostat, n.d., a; European Commission, n.d., b) to group countries based on their energy poverty predisposition. Afterwards, we use the selected indicators in Principal Component Analysis (PCA) and identify four Principal Components for further analysis.

Section 4 follows up with the implementation of the above-mentioned components in Hierarchical Clustering (HC). As a result of the HC model, the EU countries are divided into seven different clusters with similar energy poverty trends.

In Section 5, we define household electricity price elasticity models and inspect varying household consumption behavior in different parts of EU during the times of changing electricity prices. Even though the number of clusters in the models is limited to two, the outcomes still indicate clear cross-country energy poverty patterns in Europe.

In Section 6, the key discoveries of the thesis are briefly summed up. In addition, we list some limiting aspects of the paper and suggest possible directions for future studies that may stem from our research.

In summary, PCA and HC proclaim that among the most reasonable ways of grouping European countries is dividing them into seven clusters, notably close to each other from numerous aspects: economic, social, environmental, and geographical. Western European countries constitute the largest cluster, and energy poverty susceptibility ranges across “old” and “new” EU states. From a household electricity price elasticity model, energy poverty manifestation varies considerably across various clusters, making Southern European countries (Portugal, Spain, Slovenia, Italy, Malta, Bulgaria, Greece, and Cyprus) more vulnerable to its dramatic aftermaths.

Knowing the geographic features of energy poverty in Europe can positively influence how this problem is tackled in the EU, as it will facilitate the creation of tailor-made programs and projects and channelling financial support. For instance, it can have implications for the further advancement of The LIFE Clean Energy Transition Programme, which has a budget of over EUR 1 billion for the years 2021–2027 and supports requests for initiatives that explore strategies for combating energy poverty. Consequently, it would be possible to allocate funds primarily to the development of projects in countries especially prone to energy poverty. Alternatively, the theoretical understanding of the issue might be of great value within the scope of the Energy Poverty Advisory Hub (EPAH), which provides a number of materials to assist stakeholders in the implementation of practical measures to address energy poverty (European Commission, n.d., a).

Overall, energy poverty is an important problem because it affects people's health, education, economic development, and the environment. It is a fundamental obstacle to achieving sustainable development goals and improving people's lives worldwide. That is why deeper knowledge about energy poverty is beneficial to politicians, policymakers, industry specialists, scientists, etc.

2. Background

2.1 Fuel and energy poverty uncertainty

Despite the occurrence of energy poverty all around the globe and ever-increasing interest in this issue, there is still no consensus in the scientific community regarding a single and universal definition of energy poverty. Moreover, the terms “energy poverty” and “fuel poverty” are often used interchangeably or as synonyms (Thomson et al., 2016), causing even more confusion.

First measured in solely monetary terms, “poverty” has evolved into a multi-dimensional concept that encompasses multiple aspects of well-being. There is now almost unanimous recognition of the inextricable link between poverty and deprivation. Herewith deprivation can be perceived in view of “constraints on people’s choices to access certain material goods, assets, capabilities, freedoms and opportunities” (Pachauri et al., 2004).

The term fuel poverty was coined in the 1980s to describe the consequences of the 1973 oil crisis. The crisis led to a great number of UK households, especially low-income ones, starting to struggle to keep their homes warm due to increased fuel prices (Bradshaw & Hutton, 1983). In 1991, Boardman characterized fuel-poor households as those that allocate more than 10% of their income to fuel expenditure on all energy services (Boardman, 1991). Since the 1990s fuel poverty has been included in policy agendas in the UK and Ireland (Thomson & Snell, 2013).

Fuel poverty is viewed as a form of deprivation and disadvantage (Boardman, 2009). The idea of deprivation is connected to Healy’s notion of fuel poverty as of “inability to heat the home adequately” (Healy, 2004). From this notion stems a more elaborated Buzar’s definition of an “energy poor household” – a household with the amount of warmth in its home that does not allow for participating in the “lifestyles, customs and activities which define membership of society” (Buzar, 2007). Nevertheless, the definitions above barely revolve around one side of energy poverty – heating, not taking into consideration its other various facets and not making a distinction between fuel and energy poverty.

As per later Boardman’s definition, fuel-poor households do not have sufficient funds to afford the most basic levels of energy to provide them with heating, lighting, cooking, and appliance use. Although the complexity of the issue is highlighted, energy and fuel poverty are used as synonyms (Boardman, 2009).

Regardless of its gravity for academia and policymaking, the conceptual relationship between fuel poverty and energy poverty remains ambiguous, “possibly to the detriment of future progress in the field and ultimately sustainable lifestyles of the population” (Primc et al., 2021). In some sources, “energy poverty” would traditionally be used when discussing developing countries and their “problems of inadequate access to energy in developing countries, involving a host of economic, infrastructural, social equity, education and health concerns” (Bouzarovski et al., 2014). Alternatively, “fuel poverty” would be utilized in industrialized countries to capture households having insufficient financial resources to pay for their primary energy needs (Castaño-Rosa et al., 2019).

However, “fuel poverty” and “energy poverty” are used interchangeably even in the legislative resolutions of the European Parliament (Thomson et al., 2016). In support of such an approach and in order to point out similar results of any kind of household-scale energy deprivation, Bouzarovski and Petrova refer to the aforementioned Buzar’s (Buzar, 2007) definition and consider fuel poverty and energy poverty “under the same conceptual umbrella: as a set of domestic energy circumstances that do not allow for participating in the lifestyles, customs and activities that define membership of society” (Bouzarovski & Petrova, 2015). In our thesis, we will adopt the latter standpoint.

2.2 Effects of energy poverty on society

Although only affecting a small portion of the population, energy poverty poses a serious challenge that is tricky to assess and control, making it hard to effectively address through appropriate policy measures. It is commonly acknowledged that households in developed countries encounter energy poverty primarily as a result of a combination of scarce income, high energy costs, and poor energy efficiency of buildings. However, a growing body of research alludes to the potentially significant role of other drivers, both socioeconomic and spatial (Dalla Longa et al., 2021). Therefore, given the complex and multi-faceted nature of energy poverty, the gravity of the effects that it has on society varies significantly from regions and areas under examination. This chapter provides a review of literature that addresses the issue and brings information about the prominent and hidden consequences of energy poverty, along with the social problems related to it.

There is a broad set of prerequisites for human well-being, and some of them can be linked to having stable access to the needed amount of energy. For example, Healy’s (Healy, 2003) research report explores the question of increased winter mortality in several European countries. He points out that the excess number of deaths during cold days in primarily Southern European countries (namely, Portugal and Spain) correlates with high rates of fuel poverty, thus leading to the conclusion that improvement of socioeconomic circumstances will have a substantial effect on the reduction of seasonal mortality.

As per Frank et al. (2006), energy poverty is clearly associated with threats to infants' health. The scientists assess the results of participation in the Low Income Home Energy Assistance Program, an initiative with the aim of helping American families with rising energy costs, and find a noteworthy relationship. They draw a connection between receiving support, and consequently, being less likely to be exposed to energy poverty, and reporting improvements in children's health. Among those changes for the better are "less anthropometric evidence of undernutrition" and "lower odds of acute hospitalization from an emergency department visit".

Similar findings are presented by Cook & Frank (2008). Since children grow actively, they need to have high-calorific food to support physical development and maintain normal weight-for-age scores. Nevertheless, the scientists confirmed that energy-poor households are often at risk of food insecurity, so the infants from those households that also do not receive any governmental support consume less nutritious food than subsidized children.

Energy poverty affects adults' physical state as well. The econometric analysis of Lacroix & Chaton (2015) is focused on vulnerable households in France, and as a proxy for fuel poverty, the authors select self-reported perception of thermal discomfort. They find an increased probability of reporting poor health among people who struggle with cold housing compared to those who live in adequately heated homes. They judge that despite the measurement of energy poverty being subjective, the negative impact of fuel poverty on the physical state is pronounced and significant.

Not having access to sufficient energy for everyday activities is a key driver for people to opt for traditional solid fuels (coal, wood, crop wastes, biomass, etc.) or kerosene for cooking and heating. Burnt dangerously and inefficiently in outdated stoves or lamps, these types of fuels cause many accidents with scalds, burns and combustion product poisoning. The not immediately obvious outcome of the above-mentioned practices is a higher susceptibility to respiratory and cardiovascular diseases, and cancer, along with the growing occurrence of premature deaths (WHO, 2014).

The health drawbacks of fuel poverty can also influence the quality of education. Indoor air pollution, the consequence of burning traditional fuels in inefficient stoves, with emissions such as carbon monoxide accumulating in unventilated space, is one of the critical reasons for acute respiratory infections found in children and chronic obstructive pulmonary diseases – in adults. Respiratory diseases among children are said to be strongly connected to school absenteeism, principally in developing countries. Lengthy episodes of these infections, which can recur several times a year, prevent students from attending classes (Gaye, 2008) and can worsen their school performance.

Living in energy poverty is deemed to prompt poorer mental health state, and nowadays fuel poverty is listed among the key risk factor associated with mental health in Europe. Low temperatures and dampness are common in energy-poor households, hence creating depressing conditions for mental well-being. Moreover, empirical studies confirm the existence of links between poor thermal comfort, perceived affordability of heat and psychosocial stress (Liddell & Guiney, 2015). Therefore, the ramifications of energy poverty related to mental health are manifold.

Fuel poverty exacerbates social inequality, especially if it is accompanied by policymaking that does not fully take into account the interests of vulnerable groups of the population. In the article by Snell et al. (2015), statistical analysis of the English Housing Survey let the scientists infer several statements. Firstly, there is a higher instance of energy poverty among households with disabled people. And secondly, single disabled people of working age are more frequently living in fuel-poor conditions. These facts illustrate how energy poverty targets and further marginalizes vulnerable social strata.

As mentioned in the article by Sovacool (2012), poverty and energy deprivation are almost inseparable. Not having access to sufficient energy prevents people from engaging in a variety of income-generating activities. For households this would mean refraining from, for instance, using mechanical power for milling grain, illumination for factories and shops, heat for processing crops, and refrigeration for preserving products. In this case, fuel poverty plays a crucial role in reducing productivity and welfare and slowing economic development in rural areas.

The objective of the article by Hassan et al. (2022) is to make an estimation of the impact of energy poverty on the population of BRICS (Brazil, Russia, India, China, and South Africa) countries from the perspectives of health, education, environment, and other socio-economic aspects. The researchers choose carbon emissions as a proxy for environmental quality because these pollutants remarkably contribute to the deteriorating ecological situation. The empirical results of the study demonstrate a positive association between carbon emissions and energy poverty and confirmed the hypothesis that energy poverty escalates pollution. This statement is true for both developing and developed countries (WHO, 2014).

2.3 Energy vulnerability

The incidence of energy poverty is marked by uneven distribution across Europe. The spatial dissemination of this phenomenon is often said to stem from the poor quality of existing infrastructure and residential dwellings, high levels of income inequality, and expensive energy (Boardman, 2009; Recalde et al., 2019). It is seen in the prevailing way of analyzing energy poverty drivers in the form of a “triad”: poor and not energy-efficient dwelling conditions, low household income and high energy prices (Boardman, 2009). However, growing research and deepening knowledge about the problem reveal how complex and multifaceted energy poverty is, so a more comprehensive approach is required.

The energy vulnerability framework attempts to shift from solely looking at energy efficiency and costs to understanding energy poverty in a comprehensive way by examining a variety of factors, including geography, cultural norms, social behaviors, and differences in energy needs (Bouzarovski & Petrova, 2015; Simcock et al., 2017). Developed by Bouzarovski and Petrova (Bouzarovski & Petrova, 2015), this framework stresses the need to incorporate the household’s internal and external factors.

Age, health, mobility, and physical condition are among the internal factors that can be named to determine the time spent at home and, consequently, energy consumed. The most salient external factor is energy prices. By including inner and outer circumstances in the context, the energy vulnerability framework reflects the spatial distribution of the problem and makes it possible to consider the households’ ability to enter and leave the state of energy poverty (Simcock et al., 2017). Hence, this framework embraces the circumstantial nature of exposure to energy poverty risks. This characteristic makes the framework suitable for our all-European study and tracking countries’ vulnerability to changes in energy prices – the main external circumstance. Table 2.1 shows the key energy vulnerability factors.

Table 2.1: *The factors of energy vulnerability (adapted from Bouzarovski & Petrova, 2015)*

Factor	Driving force	Category
Access	Inadequate availability of energy carriers suitable for addressing household needs	External/Internal
Affordability	High cost of fuels relative to household incomes, including the impact of tax or assistance programs. Inability to invest in the new energy infrastructures development	External/Internal
Flexibility	Inability to switch to a form of energy service provision fitting for household demands	Internal/External
Energy efficiency	Disproportionately large waste of usable energy during energy conversions in the household	Internal
Needs	Household energy needs not being met by current energy services due to social, cultural, economic, or health concerns	Internal/External
Practices	Lack of understanding regarding support programs or energy efficiency in households	Internal/External

One can infer that in the current global energy crisis, the households' ability to withstand problems related to soaring energy prices is crucial for avoiding energy poverty vulnerability. Therefore, the capability of dodging price peaks may also be interpreted as one of the "long-term forms of flexibility". Yet flexibility should also include the capacity to adjust energy service needs. A household's ability to avoid both short- and long-term price peaks and, as a result, lower its energy expenses, can be significantly impacted by altering behavior or investing in technical equipment. All these actions affect energy poverty vulnerability (von Platten, 2022).

Within the framework, the ramifications of energy poverty extend past the elevated energy prices of today and let us foresee potential future vulnerabilities as the energy transition becomes more and more dependent on the so-called "demand-side flexibility". Admitting that energy poverty vulnerability has several sides allows to develop a more sophisticated and holistic picture of the various kinds of risk that can arise among households (von Platten, 2022).

2.4 Measuring energy poverty

Seeing how many factors condition people's access to modern energy services, it becomes clear how challenging it is to account for all the determinants when measuring energy poverty. The dearth of research on this topic in the European Union further exacerbates the problem, as the knowledge about energy poverty is concentrated in few countries. The UK and Ireland, which have a long history of scholarly study, practice-based approaches, and governmental frameworks to address the issue, are the countries where energy poverty and related concepts are most broadly examined (Thomson et al., 2017).

Pye et al.'s (2015) report, published by INSIGHT_E – an energy think tank informing the European Commission, shows a detailed analysis of energy poverty and vulnerable consumers in the energy sector across the EU. The report processes a rather extensive and occasionally disparate body of research to offer a list of indicators that can be used to determine the magnitude of the energy poverty phenomena in European countries. The indicators are aggregated in the following way:

- Income (including the share of the population at risk of poverty);
- Energy consumption (taking into consideration residential energy consumption by fuel type);
- Energy prices (as a crucial factor affecting affordability of energy);
- Tenure status and other housing characteristic influences (such as type of dwelling and heating system);
- Well-being and material deprivation (dwelling conditions and people's ability to pay energy bills).

While not directly pinpointing the phenomena of energy poverty, these metrics do reflect parts of its entire scope. That is why scientists started coming up with the idea of incorporating all these parameters to measure and compare energy poverty in a consistent manner (Maxim et al., 2016).

One of the ways of creating a composite energy poverty index is proposed by Nussbaumer et al. (2012). The scholars develop the Multidimensional Energy Poverty Index (MEPI) by classifying the array of proxy indicators to be utilized and then determining a relative weight and deprivation threshold for each measurement. If the set of deprivations experienced by an individual exceeds the established threshold, the person is considered energy poor (Nussbaumer et al., 2012).

When contrasted with aggregating macro indicators, this composite index gives a more objective picture of energy poverty at the national level, but it is difficult to use at a cross-national level. It necessitates granular and uniform data on the living conditions of the population across several countries, which makes it hard to measure MEPI all across the EU (Maxim et al., 2016).

The approach of Recalde et al. (2019) integrates the idea of a unified index into energy poverty vulnerability framework. The scientists' study demonstrates how the configuration of structural determinants in European states affects their ability to safeguard people from both internal and external influences that can push households into energy poverty. This leads them to label the phenomenon structural energy poverty vulnerability (SEPV) (Recalde et al., 2019).

In order to create a structural energy poverty vulnerability index and investigate the relationship between SEPV and the prevalence of energy poverty, the scholars analyze the broad political and socioeconomic circumstances in each of the EU's 27 member states. The measure, which includes 13 different factors, reveals that SEPV is distributed unevenly across the EU, with the most vulnerable countries exhibiting statistically greater rates of energy poverty. Moreover, SEPV helped detect spatial patterns of energy poverty across Europe (Recalde et al., 2019).

2.5 Socio-economic and demographic differences in European countries

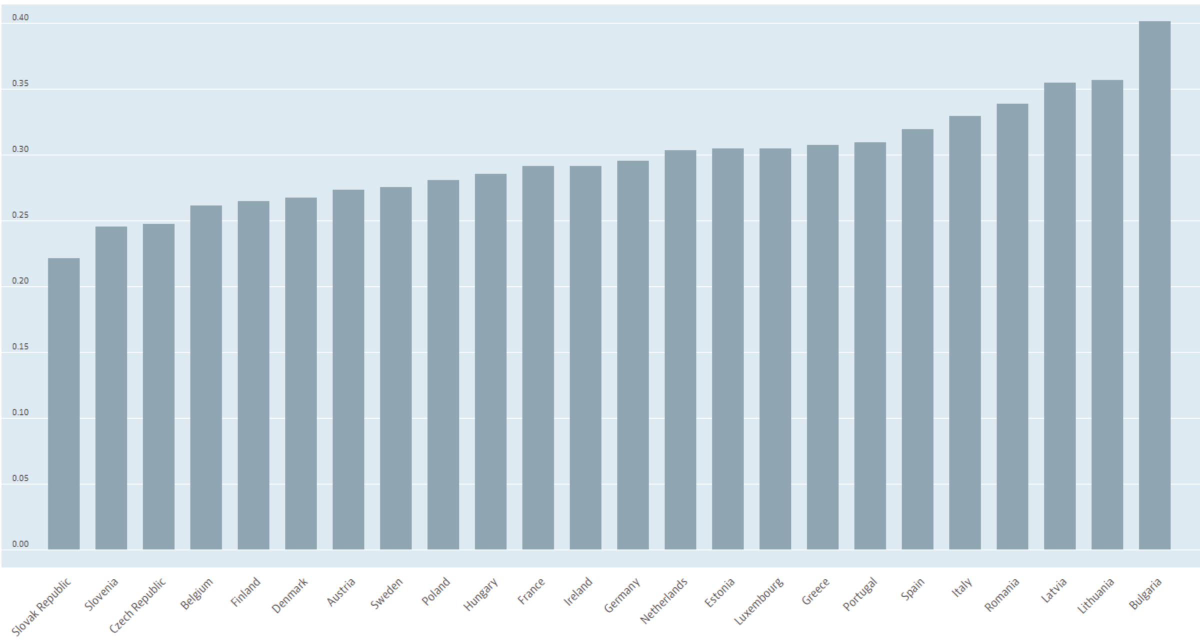
As scientific literature demonstrates, many factors dictate the geography of energy poverty in Europe. For this reason, it is necessary to discuss in more detail the key social, economic, and demographic aspects that determine such heterogeneity in the EU.

The first, and probably the most prominent markers of the socioeconomic disparity in Europe are income and wealth. In addition to growing disparities between the lowest and top income earners within the EU, the "Great Divide" also refers to divergent tendencies between European states. It is challenging to conceptualize a "single European model" since the degrees and “trends” of inequality that exist in Europe are so diverse. The EU’s income inequality has expanded during the past three decades for a wide range of causes, most notably those related to changes in the labor market and redistribution (OECD, 2017).

Moreover, the profile of people at the bottom of the income distribution has altered as well: currently, young people and families with children are more vulnerable than the elderly categories of population. A stabilization, if not a rise, in income disparity has been linked to the severe economic crisis and the typically modest recovery in several European countries. Tax and transfer policies play a substantial role in lowering market income inequality across all of Europe, although certain welfare systems are more effective at managing this redistribution (OECD, 2017).

The figure below shows the values of Gini coefficient in the EU countries. The Gini coefficient is meant to reflect income inequality within the states. It ranges from 0 in the event of perfect equality to 1 and is based on the comparison of cumulative population proportions versus cumulative income proportions that individuals receive (OECD, n.d., b).

Figure 2.1: *Gini coefficient of disposable income inequality in 2018 (or latest year) in the OECD-EU countries (OECD, n.d., b)*



As per figure 2.1, Central Eastern Europe (Slovenia, the Slovak and Czech Republics) is the region with the lowest disposable income inequality. The indicator is even lower than that in the traditionally "wealthier" Nordic countries (Finland, Denmark, and Sweden) and Benelux (Belgium, the Netherlands, and Luxembourg). Among the most unequal states are those from the Baltic region (particularly Latvia and Lithuania) and Bulgaria, the countries allocated to the category of the "new EU".

In general, the EU's population constitutes a steadily declining share of the global population (figure 2.2). The EU population is ageing rapidly, with the median age having increased up to 43.9 years in 2020 compared to 38.4 years two decades ago. The global population is, on the other hand, growing consistently and getting proportionately younger. This can be attributed to surged life expectancy (81.3 years in 2019 against 69.86 in the 1960s), as well as to a declining birth rate (4.05 million live births in 2019 versus 6.79 million in the 1960s). The COVID-19 pandemic caused a spike in "excess mortality" in 2021 (figure 2.3), with the two highest peaks being concurrent with two waves of infection in April and November (Kiss et al., 2022).

Figure 2.2: *EU and world population (Kiss et al., 2022)*

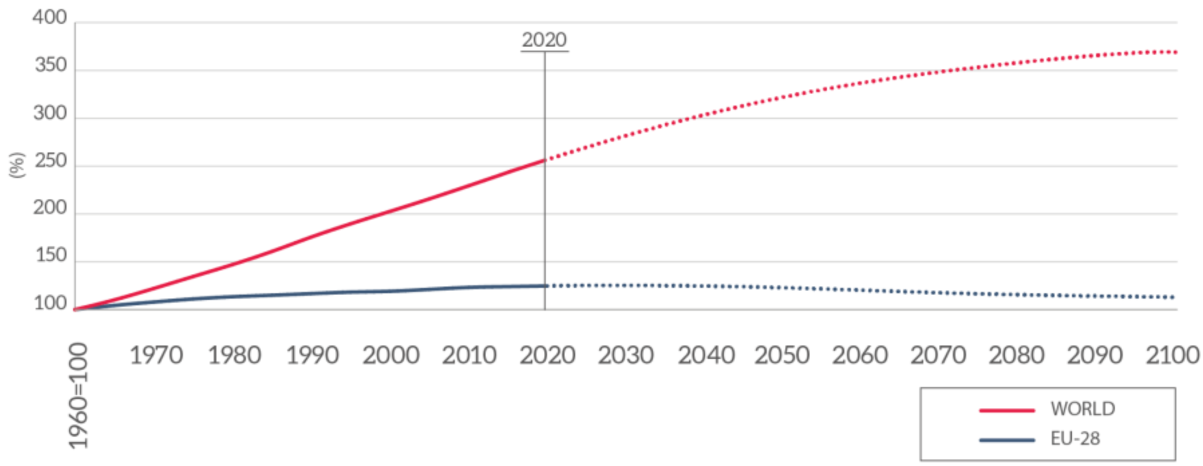
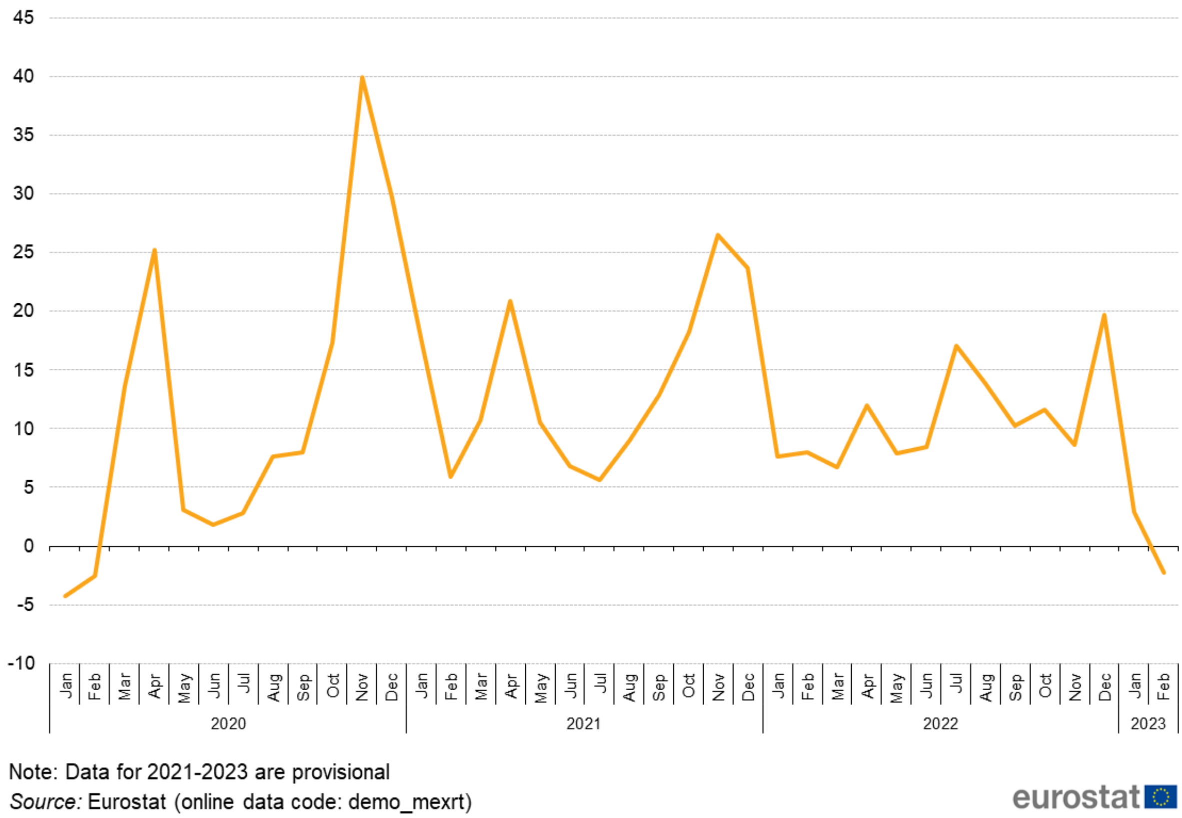


Figure 2.3: *Monthly excess mortality in the EU (Eurostat, 2023a)*



The EU states are affected by demographic developments in different ways. For instance, a significant population drop is seen in several European countries, particularly in rural and isolated areas. The economic downturn that certain regions are now experiencing might be aggravated by this circumstance, hence extending the gap between wealthy and poor social strata. The social, economic, and geographical integrity of the EU is therefore also greatly influenced by demographics. On the contrary, the high population density in metropolitan areas also has certain unfavorable effects, such as pollution and a shortage of affordable housing. All these characteristics confirm that even though the demographic balance in many EU areas has improved as a result of recent migration patterns, migration has an unequal effect on EU regions (Margaras, 2019).

Migration processes are strongly related to employment situations in countries. According to statistics, EU-mobile employees travelling home accounted for over 40% of all migratory movements in the EU between 2015 and 2020. In 2019-2020, some Member States experienced a discernible rise in the number of home comers opposed to emigrants. A prime example is Bulgaria, where six times more citizens returned to their homeland in comparison to the number of people who left their homeland in 2020 (European Commission, 2023).

The migration from non-EU countries does not compensate for the population decline, what forces some European countries to expect a reduction in the number of their citizens in the coming years. Until 2030, this problem is predicted to have an impact on Romania, Bulgaria, Greece, Croatia, Italy, Latvia, Lithuania, and Hungary. Nevertheless, the opposite trend is anticipated in other EU states over the same time frame. Denmark, Ireland, Cyprus, Luxembourg, Malta, and Sweden are among the countries with a positive demographic forecast (European Commission, 2023).

Thus, the heterogeneity of the EU countries is obvious in the light of their exposure to social, economic, and demographic changes. As a corollary, the causal effect of energy price fluctuations may potentially take different forms within Europe.

2.6 Households' economic well-being

It has been mentioned that economic well-being is one of the factors strongly linked to the spread of energy poverty in the EU countries. Thereby, in this chapter we will delve deeper into the aspects taken into account when calculating the level of households' economic well-being.

A household's control over its financial resources and eventual capacity to uphold a minimal standard of living are key factors in determining how well off economically this household. This is due to the mechanism by which households can finance their consumption of goods and services being dictated by their economic resources. Economic resources such as income and wealth, as well as their availability to fund the purchase of products and services utilized for the direct fulfilment of individual or collective needs, may be used to gauge people's economic well-being (Australian Bureau of Statistics, 2010).

The study by Meyer & Sullivan (2011) analyzes the relative pros and cons of applying consumption and income metrics to estimate the well-being of low-income households in the United States. The scholars contend that consumption, rather than yearly income, is a better indicator of the family's long-term resources (or their permanent income) given that the disparities in the accumulation of assets or access to credit between households and over time are not adequately captured by income statistics. The value of government programs is also more likely to be reflected in consumption, making it a more helpful metric for policymakers (Meyer & Sullivan, 2011).

The Federal Reserve of the United States annually fields the Survey of Household Economics and Decisionmaking. The survey is conducted by staff in the Board's Division of Consumer and Community Affairs each autumn since 2013, and it covers many sides of the economic well-being of respondents: living arrangements, employment, mortgage situation, education, retirement, health, insurance, financial literacy, income, consumption, etc. (The Federal Reserve System, 2022).

In Europe, the OECD dashboard considers a whole range of indicators that mould the economic well-being of households. Among those are real household disposable income, net cash transfers to households, real household consumption expenditure, consumer confidence, households' savings rate, households' indebtedness, financial net worth, unemployment rate, and labor underutilization rate. Albeit with limitations, these indicators allow a broader look at this sphere of people's lives (OECD, n.d., a).

Attempts are also being made to combine many heterogeneous indicators into one composite index of economic well-being. Trying to find an answer to the question of how the economic component of social welfare should be measured, Osberg & Sharpe (2005) construct the Index of Economic Well-Being (IEWB). The scientists divide metrics into four equally weighted categories: consumption flows, stocks of wealth, equality, and security. They conclude that index composition should be easily justifiable and not overcomplicated. So, they appeal to the intuition behind focusing on the abovementioned categories as they believe it is simple to argue that citizens are "better off" economically when consumption is sustainable, when total income is distributed fairly, and when people have greater economic security (Osberg & Sharpe, 2005).

Following the Osberg & Sharpe's approach of "keeping it simple", we find it reasonable to limit the number of variables included in the regression and concentrate on electricity consumption, households' characteristics (income, number of households, persons per household, electricity price), and unemployment.

2.7 Price elasticity in energy context

In energy economics and policy, it is crucial to comprehend how changes in the price of an energy input, like electricity or natural gas, affect the consumption of that input. Price elasticity, or the percentage change in demand when the price increases by 1%, is a widely used way to conveniently summarise this data. This measure enables regulators to accurately plan infrastructure and grid investments and quantify the welfare effects experienced by consumers as the regulatory environment is modified or as utilities enter or leave a market (Labandeira et al., 2012; Miller & Alberini, 2016).

The issue of energy price elasticity is being raised in various countries and regions. For example, Sun & Ouyang (2016) conduct a residential sector analysis to investigate the price and expenditure elasticities of household energy demand during urbanization in China. The scientists discover that the own-price elasticities of electricity, natural gas, and transport energy are all negative, suggesting that in the case of other energies staying constant, a price hike would result in a decrease in demand. Even having considered the different income groups, the scholars do not find evidence for the energy demand being elastic (Sun & Ouyang, 2016).

In the Portuguese case study by Silva et al. (2017), which uses microdata extracted from five independent surveys with five-year intervals between each of them, the conclusion is opposite. The scientists find a rather high medium-/long-run electricity price elasticity, which shows that as electricity costs increase, consumption declines significantly. Furthermore, the researchers perform an income quintiles analysis to demonstrate substantial variations in the elasticities based on income group. The calculated own price elasticities for electricity appear to decrease at higher use quantiles. It implies that households with greater use levels are less sensitive to price changes than those with lower use levels, probably because these households are highly dependent on electricity and are unable to switch to another source when the cost rises (Silva et al., 2017).

The research by Schulte & Heindl (2017) is focused on residential price and expenditure energy elasticity in Germany. By adopting a quadratic expenditure system and scrutinizing expenditure data from 1993 to 2008, the scientists come to noteworthy results. Their study's findings indicate that household spending has a significant impact on how people respond to changes in energy prices. Compared to households in the lowest expenditure quartile, households in the highest expenditure quartile are considerably more sensitive to both heating and electricity price fluctuations. Hence, they claim that households' relative burdens are heavily influenced by overall spending (Schulte & Heindl, 2017).

The study by Csereklyei (2020) shifts the view from country to the whole European Union level. The paper looks at the short- and long-term price and income elasticities of domestic and industrial electricity demand in the EU between 1996 and 2016. Both dynamic panel and instrumental variable models with the between estimator are used to offer reliable estimates on the indicated elasticities, which seem to be more elastic among European businesses rather than households (Csereklyei, 2020).

The cross-country approach is adopted by Liddle et al. (2020) as well. However, the geography of the study expands beyond Europe and covers the period from 1996 to 2014. Time-varying income and energy price elasticities are calculated for a twenty-six-country, middle-income balanced panel, which includes mostly non-OECD states. The economists discover that the price elasticity of the energy demand is either negligible or low and positive (Liddle et al., 2020).

This finding is in line with the conclusion of an empirical studies meta-analysis on the energy price elasticity, conducted by Labandeira et al. (2017). The demand for the most significant energy products, including electricity, natural gas, petrol, diesel, and heating oil, is taken into account in this study together with total energy consumption. The scientists determine that consumers' responses to price fluctuations are comparable across different energy goods and are stronger over the long run than they are in the short term (Labandeira et al., 2017).

In this thesis, we include several features of the described articles and estimate electricity price elasticity of households in different parts of the EU.

3. Principal Component Analysis

As the first step of our analysis, we aim to identify similarities between the EU countries based on the mixture of energy poverty indicators. The methods chosen for identifying similarities in energy poverty indicators are Principal Component Analysis (PCA) followed by Hierarchical Clustering (HC). In Subsection 3.1, we describe energy poverty indicators data. Subsection 3.2 covers PCA methodology. In Subsection 3.3, the results of applying PCA analysis to the dataset are discussed. The consequent usage of HC method and its outcomes is further covered in Section 4.

3.1 Data description

In our research, we analyze energy poverty indicators of 27 European Union countries as of May 2023: Austria (AT), Belgium (BE), Bulgaria (BG), Croatia (HR), Cyprus (CY), the Czech Republic (CZ), Denmark (DK), Estonia (EE), Finland (FI), France (FR), Germany (DE), Greece (EL), Hungary (HU), Ireland (IE), Italy (IT), Latvia (LV), Lithuania (LT), Luxembourg (LU), Malta (MT), the Netherlands (NL), Poland (PL), Portugal (PT), Romania (RO), Slovakia (SK), Slovenia (SI), Spain (ES), Sweden (SE).

For the first part of the analysis, described in Sections 3 and 4, we utilize energy poverty indicators' data by European Commission (n.d., b). The data includes twenty-one indicators listed in Table 3.1. The data covers the period from 2003 to 2022 and includes national-level energy poverty indicators for forty-two countries in Europe.

Table 3.1: *Energy poverty indicators data (European Commission, n.d., b)*

Indicator	Period covered
Arrears on utility bills – No disaggregation – Country average	2003-2022
At Risk of Poverty or Social Exclusion	2003-2020
Biomass prices	2005-2015
Coal prices	2003-2008, 2014, 2015
District heating prices	2003-2015
Dwellings in populated areas – Dwellings in intermediately populated areas	2003-2014
Dwellings with energy label A	2007-2015
Energy expenses by income quintile – Energy expenses, income quintile 1	2005, 2010, 2015
Excess winter mortality/deaths	2005-2014
Fuel oil prices	2003-2015
High share of energy expenditure in income (2M) – No disaggregation – Country average	2010, 2015
Household electricity prices	2007-2021
Household natural gas prices	2007-2021
Inability to keep home adequately warm – No disaggregation – Country average	2004-2021
Low absolute energy expenditure (M/2) – No disaggregation – Country average	2010, 2015

Number of rooms per person by ownership status – Total	2003-2022
Pop. Liv. dwelling equipped with air conditioning	2007
Pop. Liv. dwelling equipped with heating facilities	2007, 2012
Pop. Liv. dwelling with presence of leak, damp and rot	2003-2020
Pop. Liv. dwellings comfortably cool in summer time	2007, 2012
Pop. Liv. dwellings comfortably warm in winter time	2007, 2012

For our analysis, we narrow down the list of indicators best fitting to the purpose of our research and the availability of data for twenty-seven EU countries during a period of time that is of interest for this research. A more detailed description of the chosen indicators is presented in Subsection 3.2.

In the second part of the analysis, discussed in Section 5, we utilized Population and Social Conditions, Economy and Finance, Environment and Energy and Macroeconomic imbalance procedure indicators retrieved from the database of Eurostat (n.d., a) and Household electricity prices indicator by European Commission (n.d., b). The list of Eurostat (n.d., a) indicators utilized in the research is shown in Table 3.2.

Table 3.2: *Eurostat indicators data (Eurostat, n.d., a)*

Indicator	Section	Description of data utilized	Period covered
Number of households by household composition, number of children and working status within households	Population and Social Conditions	Annual number of households (all types) measured in thousands of households for EU countries	2006-2021
Average number of persons per household-by-household composition, number of children and working status within households	Population and Social Conditions	Annual average number of persons per household for all types of households for EU countries	2006-2021
Disaggregated final energy consumption in households – quantities	Environment and Energy	Annual household electricity consumption in Gigawatt-hour for EU countries	2010-2020
Unemployment rate – annual data	Macroeconomic imbalance procedure indicators	Annual unemployment rate for age group 15-74 for EU countries	2003-2022
Distribution of income by quantiles – EU-SILC and ECHP surveys	Population and Social Conditions	Annual figure for top income cut-off point for quantile 1 in EUR for EU countries	1995-2022
GDP and main components (output, expenditure, and income)	Economy and Finance	Annual GDP measure in current prices, million EUR for EU countries	1975-2022

In our research, we filter and limit the data in terms of the time period, countries list, and specific indicator features to answer the research question. The choice of factors for filtering and limiting data is further described in Subsection 3.2 and Subsection 5.1.

Data gathering and transformation, as well as all the consequent data analysis, is performed using R programming language (R Core Team, 2022) in RStudio software, version

2022.12.0.353 (Posit team, 2022). The R packages used in the analysis are described in Appendix (see Table A1).

3.2 Principal Component Analysis of energy poverty indicators

In the first part of the analysis, we aim to identify a mix of energy poverty indicators to be used as a measure for assessing energy poverty pattern similarities across countries in the EU. Principal Component Analysis (PCA) is chosen as a method to determine the most optimal combination of energy poverty indicators. PCA analysis utilizes the concept of combining variables into principal components, which allows to limit the number of variables used to explain the variance in data. (James et al., 2021, p. 498-510). In the first part of the analysis, PCA is applied to data with the list of energy poverty indicators to identify the principal components that most optimally explain the variance between variables across different countries in the EU.

Recalde et al. (2019) perform PCA analysis as one of the steps to develop the Structural Energy Poverty Vulnerability (SEPV) composite index, previously discussed in Subsection 2.4. The indicators used by Recalde et al. (2019) include economic, financial, social, environmental, and energy factors, such as unemployment rate, electricity consumption per capita, etc. The SEPV index, which combines a range of indices in various proportions and is used as a measure of clustering EU countries based on patterns of energy poverty, is produced as a consequence of the application of PCA.

In this research, a similar approach is applied. However, to determine a composite measure for energy poverty, a list of energy poverty indicators already identified by the European Commission (n.d., b) is utilized. The energy poverty indicators listed in Table 3.1 are assessed in terms of the period coverage and availability of data for all EU countries.

Due to a lack of observations for the majority of EU countries (data missing for three or more EU countries), the following indicators are excluded from the analysis: Biomass prices, Coal prices, District heating prices, Dwellings with energy label A, Fuel oil prices, Household natural gas prices. The household electricity prices indicator is excluded as it is used in the second part of the analysis in the electricity price elasticity model as a dependent variable.

In order to perform PCA, for each of the chosen energy poverty indicators, we determine the most recent year when observations for the majority of the EU countries (with the maximum of two countries missing an observation in the selected year) are available. Table 3.3 summarizes the chosen fourteen energy poverty indicators and the year for which the most recent observations are retrieved.

Table 3.3: *Energy poverty indicators utilized in PCA*

Indicator	Description	Year
Arrears on utility bills – No disaggregation – Country average	An indicator measuring the percentage of households which have not been able to pay utility bills on time in the last twelve months due to financial difficulties. (European Commission, n.d., b).	2020
At Risk of Poverty or Social Exclusion	A general poverty indicator measuring the percentage of the population who are at risk of poverty or social exclusion. In order to capture the energy side of poverty, the measure should be used together with other energy poverty indicators. (European Commission, n.d., b).	2020
Dwellings in populated areas – Dwellings in intermediately populated areas	An indicator presenting a % of dwellings, situated in intermediately populates areas (100-499 inhabitants per square meter). (European Commission, n.d., b).	2014
Energy expenses by income quintile – Energy expenses, income quintile 1	The data represents % of income that households in income Quantile 1 spend on energy (gas, electricity, etc.). Income quantiles are a division of households into five groups with Quantile 1 being the group of households with the lowest income (Eurostat, n.d., b).	2010
Excess winter mortality/deaths	An indicator measuring the ratio of winter mortality cases versus average death cases in non-winter months each year (European Commission, n.d., b).	2014
High share of energy expenditure in income (2M) – No disaggregation – Country average	A measure presenting a percentage of households with energy expenditures as a proportion of household budget being more than twice a country's median. The indicator does not account for underconsumption of certain households. (European Commission, n.d., b).	2015
Inability to keep home adequately warm – No disaggregation – Country average	Measures a percentage of households which were not able to be adequately heated. The limitations of the indicator are the difference of "adequate heat" perception for different households in different countries, no details on the reasons of why the homes were not adequately heated and no information on summer energy poverty (covering the percentage of households which were able to be adequately cooled down). (European Commission, n.d., b).	2021
Low absolute energy expenditure (M/2) – No disaggregation – Country average	A measure which indicates a percentage of households with absolute energy expenditure being below half a country's median. This indicator does not account for energy efficiency and conditions of households as we as differences in climate of a country's regions. (European Commission, n.d., b).	2015
Number of rooms per person by ownership status – Total	An indicator presenting an average number of rooms per person by ownership status (European Commission, n.d., b).	2020
Pop. Liv. dwelling equipped with air conditioning	A measure indicating percentage of population living in dwellings equipped with air conditioning (European Commission, n.d., b).	2007
Pop. Liv. dwelling equipped with heating facilities	A measure indicating percentage of population living in dwellings equipped with heating facilities (European Commission, n.d., b).	2012
Pop. Liv. dwelling with presence of leak, damp, and rot	An indicator of the percentage of population with leak, damp, and rot present in their dwelling. The indicator could refer to various reasons for the dwelling issues, including issues with housing construction, poor house heating, climate, etc. (European Commission, n.d., b).	2020

Pop. Liv. dwellings comfortably cool in summer time	An indicator showcasing the share of population living in dwellings which inhabitants believe to have efficient cooling conditions and sufficient insulation against warmth in summer season (European Commission, n.d., b).	2012
Pop. Liv. dwellings comfortably warm in winter time	An indicator showcasing the share of population living in dwellings which inhabitants believe to have efficient heating conditions and sufficient insulation against cold in winter season (European Commission, n.d., b).	2012

Principal Component Analysis is a method which is applied to data that has no missing values (James et al., 2021, p. 510). In total, our chosen dataset has eight missing observations. To perform PCA on the dataset, the missing values are simulated using the *softImpute* package and specifically the *softImpute* method (Hastie & Mazumder, 2021). The method utilizes nuclear-norm regularization in order to impute the missing values of a matrix. At each of the iterations, the algorithm estimates a missing observation and proceeds with an optimization solution on the complete matrix. In our analysis, the energy poverty indicators dataset is transformed into a matrix, and missing values are simulated based on all the non-missing values for energy poverty indicators for different countries. For our case, we use the default approach of alternating least squares ("als"). (Hastie & Mazumder, 2021).

The observations which resulted from applying the *softImpute* method are listed in Table 3.4 with their corresponding energy poverty indicator details.

Table 3.4: *Energy poverty indicator values resulted from the softImpute method application*

Country	Abbreviation	Indicator	Year	Imputed value
Italy	IT	Energy expenses by income quintile – Energy expenses, income quintile 1	2010	9.021113
Luxembourg	LU	Energy expenses by income quintile – Energy expenses, income quintile 1	2010	11.472391
Denmark	DK	High share of energy expenditure in income (2M) – No disaggregation – Country average	2015	15.396091
Italy	IT	High share of energy expenditure in income (2M) – No disaggregation – Country average	2015	16.071200
Denmark	DK	Low absolute energy expenditure (M/2) – No disaggregation – Country average	2015	14.088818
Croatia	HR	Pop. Liv. dwelling equipped with air conditioning	2007	9.472659
Malta	MT	Pop. Liv. dwelling equipped with air conditioning	2007	48.806529
Poland	PL	Pop. Liv. dwelling equipped with heating facilities	2012	92.363891

In Table 3.4, we notice that the imputed value for the population living in dwellings equipped with air conditioning in Croatia is abnormally low, taking into account the location of the country in the Southern part of Europe and the higher requirement for air conditioning facilities in the region. Nevertheless, we make a decision to use the results of the *softImpute* method of matrix completion without changes. Our goal is, through further steps of the analysis, to scale the dataset and apply the PCA method, which shall reduce the dimension of the dataset and smoothen possible discrepancies in the imputed values. The impute of missing values is to be considered as a limitation and in case of abnormal results in PCA and HC analysis, this limitation is to be revisited.

In order to identify the number of Principal Components to use in further analysis, the proportion of variance explained (PVE) measure is utilized for each of the components (James et al., 2021, p. 505-507). The number of Principal Components that reaches approximately 80% in the cumulative proportion of explained variance is selected.

3.3 Principal Component Analysis results discussion

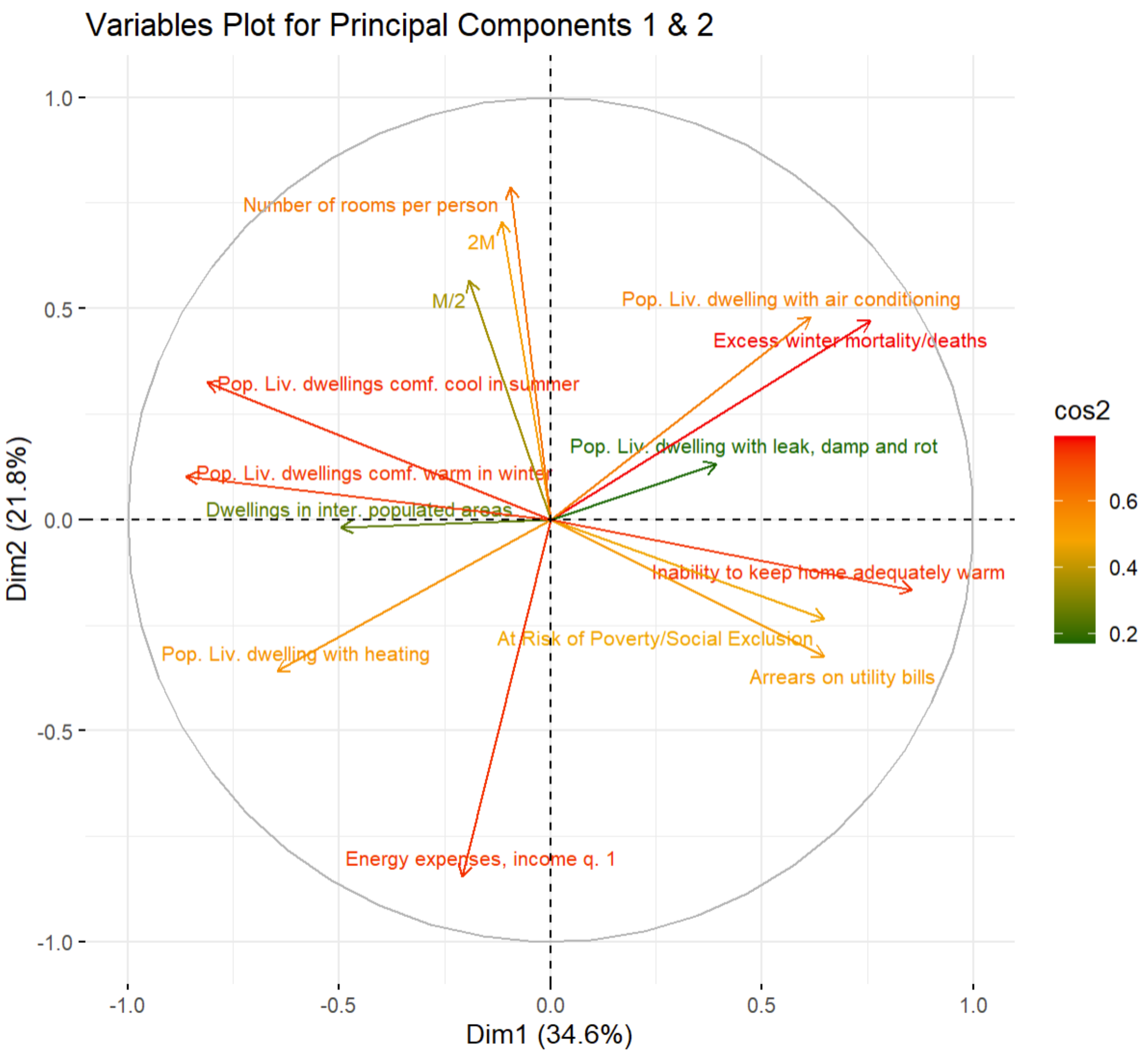
In the following subsection, the results of applying PCA to the transformed energy poverty indicators data are discussed. Principal Components loadings and score vectors are showcased, and the importance of Principal Components in explaining variance in data is analyzed. In conclusion, the number of Principal Components to be used in further analysis is determined.

Figure 3.1 showcases the relationship of variables identified as a result of applying PCA on energy poverty indicators data. We observe fourteen chosen indicators presented in the form of vectors. X- and Y-scales stand for the first and second Principal Components (PC1 and PC2) respectively.

The square cosine value or *cos2* on the right-hand side of Figure 3.1 is the measure of the variables' representation quality (Kassambara & Mundt, 2020). The higher the *cos2* value and the closer it is to 1, the better the quality of variable representation. High *cos2* values for variables in regard to PC1 and PC2 are reflected in the red color of an arrow, showcasing the vector of the variable. Medium *cos2* values are reflected by the orange color of vector arrows, and the lowest *cos2* values are painting the vector arrows in green color.

The importance of the variables to PC1 and PC2 is also presented by the length of the vectors and how close the vectors are to reaching the correlation circle on the plot. The shorter the arrow of the vector, the less important is the variable in regard to PC1 and PC2. (Abdi & Williams, 2010).

Figure 3.1: *PCA variables plot*



The vectors of variables which are closest to each other and moving in the same direction indicate relatively high positively correlated variables. The vectors of variables which are moving in opposite directions are considered negatively correlated. (Jolliffe, 2002).

Figure 3.1 indicates the following relations between energy poverty indicators for PC1 and PC2:

- Population living in dwellings equipped with air conditioning, Population living in dwellings with presence of leak, damp and rot and Excess winter mortality/deaths variables are positively correlated to each other (vectors of the variables are moving in the same direction) and negatively correlated to Population living in dwellings equipped with heating facilities variable (the vectors of the variables are pointing to the opposite directions). The finding is in line with the conclusions of several studies about reasons for seasonal excess mortality, as heat and cold are recognized environmental risk factors for human health. Moreover, this risks predisposition and vulnerability vary across Europe significantly, with both cold and heat having the greatest effects in Eastern Europe and having a growing impact with increasing age (Masselot et al., 2023). People who live in the regions with milder winters and in low efficient buildings feel more reluctant to pay for space heating as well. This cultural behavior also exposes people to the temperature-related health risks (Oliveira Panão, 2021);

- Population living in dwellings comfortably cool in summer time, Population living in dwellings comfortably warm in winter time are positively correlated to each other and negatively correlated to Inability to keep home adequately warm, Arrears on utility bills and At Risk of Poverty or Social Exclusion variables. Given that the inability to pay utility bills and energy inefficient dwellings are two of the most often cited reasons for energy poverty (Pillai et al., 2023), this correlation is contextually related to the one described above. There is a higher chance that poor or low-income households would report having holes in their roofs or walls. Hence, worse housing might cause variations in indoor temperatures even in the absence of apparent differences in consumption: on a hot day, a less efficient home will retain more heat than a more efficient home (Doremus et al., 2022);

- High share of energy expenditure in income (2M), Low absolute energy expenditure (M/2) and Number of rooms per person by ownership status variables are positively correlated to each other. This relationship is explained by the fact that not only the conditions, but also the size of the dwelling affect energy consumption and, hence, energy expenditure. The total of rooms and energy consumption in a household are positively related, so the volume of energy required in a building increases with each extra room (Salari & Javid, 2017). Among the main drivers of the growth are lighting, major appliances (Parker et al., 2011) and heating (Curtis & Pentecost, 2015). Therefore, some scientists even suggest downsizing as an option with manifold economic and societal benefits, namely, lower household expenditure on electricity bills and/or rent, release of capital and reduction in energy consumption (Huebner & Shipworth, 2017).

- The red color of the arrow and the proximity of the vector to the correlation circle for Excess winter mortality/deaths variable point at high representation quality of the variable in PC1 and PC2. Variables Dwellings in intermediately populated areas and Population living in dwellings with presence of leak, damp and rot have the least quality of representation in PC1 and PC2, which is showcased by the green color of the arrows and the short length of the variables' vectors.

In total, PCA analysis generated fourteen Principal Components. Each of the fourteen Principal Components is presented by a loading vector, which consists in turn of PCA loadings. PCA vector loadings are the elements of a Principal Component which give weight to a variable, in our case an energy poverty indicator. The loadings together present a loading vector for the Principal Component. The loadings are constructed in such a way that the sum of the squared values of the loadings for each Principal Component is equal to one. (James et al., 2021, p. 499-503).

Table 3.5 presents the loadings of vectors for the first four Principal Components: PC1, PC2, PC3, and PC4. The table showcases an example of an intermediary output for the PCA method. The loadings for all the Principal Components are included in Appendix (see Table A2).

Table 3.5: *PCA Vector Loadings for the first four Principal Components*

Energy Poverty Indicator	PC1	PC2	PC3	PC4
Arrears on utility bills – No disaggregation – Country average	0.294158	-0.18505	0.256791	-0.35681
At Risk of Poverty or Social Exclusion	0.294132	-0.13461	0.276449	-0.31158
Dwellings in populated areas – Dwellings in intermediately populated areas	-0.22509	-0.01034	-0.21392	-0.27419
Energy expenses by income quintile – Energy expenses, income quintile 1	-0.09476	-0.4836	-0.05974	0.278381
Excess winter mortality/deaths	0.343671	0.270039	-0.19208	0.118336
High share of energy expenditure in income (2M) – No disaggregation – Country average	-0.05263	0.403219	0.437498	0.06263
Inability to keep home adequately warm – No disaggregation – Country average	0.38876	-0.09519	0.020365	-0.14837
Low absolute energy expenditure (M/2) – No disaggregation – Country average	-0.08755	0.323794	0.501843	-0.01705
Number of rooms per person by ownership status – Total	-0.04287	0.450659	-0.29325	0.01258
Pop. Liv. dwelling equipped with air conditioning	0.279425	0.274433	-0.07206	-0.27584
Pop. Liv. dwelling equipped with heating facilities	-0.29395	-0.20549	0.122814	-0.50658
Pop. Liv. dwelling with presence of leak, damp, and rot	0.178411	0.07522	-0.4585	-0.37158
Pop. Liv. dwellings comfortably cool in summer time	-0.3693	0.18579	-0.10605	-0.1677
Pop. Liv. dwellings comfortably warm in winter time	-0.39127	0.058614	-0.00178	-0.2892

The weighted variable values resulted from applying PCA loadings weights then, when summarized, give a Principal Component score for each observation (James et al., 2021, p. 499-503). In other words, the summarized weighted energy poverty indicator values present a single score for every country and each Principal Component. As PCA resulted in fourteen Principal Components, every country is given fourteen scores, one per each Principal Component.

The scores for each of the EU countries for the first four Principal Components are presented in Table 3.6 to provide a visual example of a PCA output. The scores of all the Principal Components are included in Appendix (see Table A3).

Table 3.6: *PCA scores for the first four Principal Components*

Country	Abbreviation	PC1	PC2	PC3	PC4
Austria	AT	-2.04065	0.46722	0.377279	0.014997
Belgium	BE	-1.70247	0.399186	-1.77031	-0.68595
Bulgaria	BG	4.66741	-2.8597	1.219623	-0.10728
Cyprus	CY	3.734148	2.027294	-2.38831	-2.17949
The Czech Republic	CZ	-2.31056	-1.96056	-0.90194	1.438921
Germany	DE	-2.11389	0.740253	0.330494	-0.66472
Denmark	DK	-1.62448	0.633797	-0.47976	0.077593
Estonia	EE	-0.64668	-0.00923	1.404375	0.888883
Greece	EL	3.752826	-0.73114	1.905958	-1.86365
Spain	ES	2.211454	1.272867	-0.53739	-0.63568
Finland	FI	-1.90782	2.374815	2.449898	-0.16952
France	FR	-0.49002	1.044804	0.219411	-0.13569
Croatia	HR	0.131624	-1.7721	0.295817	-0.26257
Hungary	HU	-0.11213	-2.23049	-1.27128	0.134934
Ireland	IE	-0.95283	1.667126	-0.33125	-0.37973
Italy	IT	0.361835	-0.06566	-0.1378	-1.03873
Lithuania	LT	0.945211	-1.12366	1.034895	0.139921
Luxembourg	LU	-1.78878	-0.16007	-1.57798	-0.45943
Latvia	LV	0.643111	-2.27801	0.647673	0.446527
Malta	MT	3.491855	3.979162	0.024844	2.820414
The Netherlands	NL	-1.96213	-0.11005	-2.18013	-0.35489
Poland	PL	-1.04493	-0.76781	1.441594	1.356116
Portugal	PT	3.623786	0.572784	-2.09461	2.355207
Romania	RO	0.66962	-1.18935	2.045818	-0.6489
Sweden	SE	-2.47143	3.066402	2.133758	-0.25984
Slovenia	SI	-1.12615	-0.15714	-1.34545	-0.71561
Slovakia	SK	-1.93794	-2.83074	-0.51523	0.88818

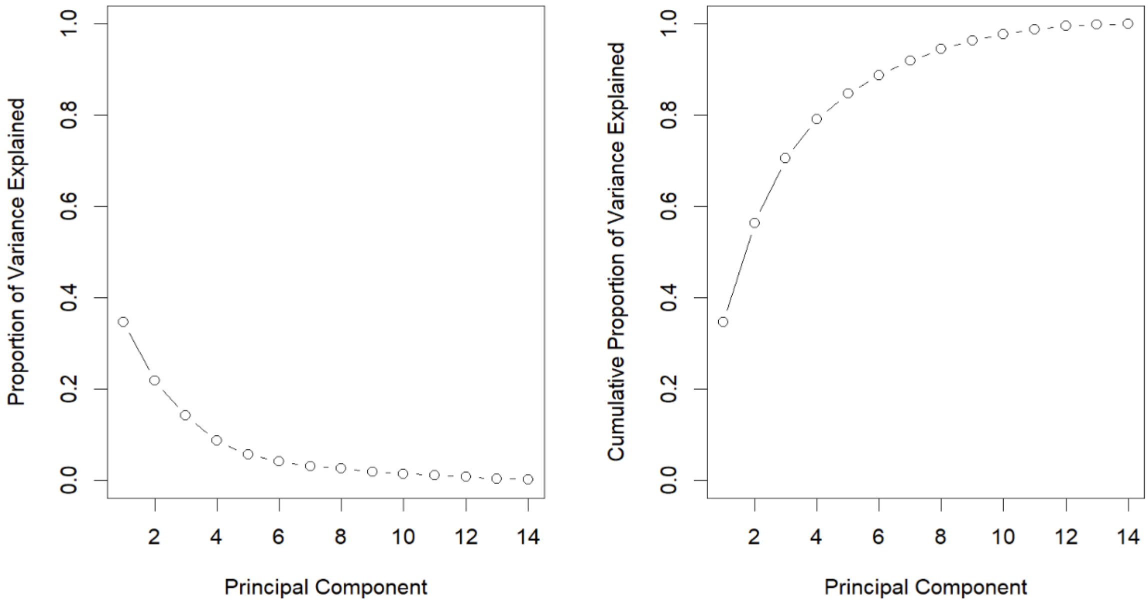
When choosing the number of Principal Components to further utilize in description of data, the proportion of variance in data explained by the components should be taken into consideration (James et al., 2021, p. 505-507). Table 3.7 presents a summary of Proportion of Variance Explained (PVE) and Cumulative PVE for fourteen Principal Components resulted from applying PCA to energy poverty dataset. We see that PC1 has the highest PVE measure – 35% of variance in data explained. Cumulatively, all Principal Components explain 100% of variance in data and the first four Principal Components reach 79% Cumulative PVE.

Table 3.7: *PCA Importance of Components Summary*

Components	Component Importance	
	Proportion of Variance	Cumulative Proportion
PC1	0.346	0.346
PC2	0.218	0.564
PC3	0.141	0.705
PC4	0.086	0.792
PC5	0.056	0.847
PC6	0.041	0.888
PC7	0.031	0.919
PC8	0.026	0.945
PC9	0.018	0.963
PC10	0.014	0.978
PC11	0.011	0.988
PC12	0.007	0.995
PC13	0.003	0.998
PC14	0.002	1.000

Figure 3.2 illustrates the scree plot of the PVE measure for each of the fourteen Principal Components and the plot for the Cumulative PVE measure curve. We observe that after the fourth Principal Component, the proportion of variance explained by each of the subsequent Principal Components decreases significantly, which can be traced on the PVE scree plot in Figure 3.2.

Figure 3.2: *PVE scree plot and Cumulative PVE plot*



The first four Principal Components, explaining 79% of the variance in data, were chosen for further analysis based on the reasoning that further addition of Principal Components is adding complexity to the interpretation of the results, and the PVE value for consequent Principal Components diminishes significantly after the fourth Principal Component.

4. Hierarchical Clustering

In the following section, we discuss the application of the Hierarchical Clustering method for grouping the EU countries based on the energy poverty trends similarities. Subsection 4.1 describes HC methodology and the type of HC model chosen for clustering the energy poverty indicators dataset. Subsection 4.2 reviews the results of HC application to four Principal Components identified in Section 3.

4.1 Hierarchical Clustering methodology

In Subsection 3.3, as the result of conducting PCA on the energy poverty indicators dataset, we have identified four Principal Components. The goal of the next step of our analysis is to identify the groups of EU countries with similar energy poverty indicators. To perform the grouping of EU countries, we partially follow the methodology described by Recalde et al. (2019). Recalde et al. (2019) use PCA method that resulted in the Structural Energy Poverty Vulnerability (SEPV) composite index, which is comprised of Principal Components, to perform the Hierarchical Clustering method and determine groups of EU countries based on their vulnerability to structural energy poverty.

Hierarchical Clustering (HC) belongs to a group of clustering methods. The purpose of clustering is to organize data into groups of observations that are most similar to each other and most distanced from observations in other clusters. The result of applying HC to a dataset is a dendrogram, which is visualised as an upside-down tree-looking plot. The upper part of the dendrogram represents one cluster that is then separated into branches, grouping observations in distinct groups (clusters) based on dissimilarity measure. The grouping continues until the observations are separated into singular groups consisting of one observation at the bottom of the dendrogram. Therefore, the observations that were separated into branches closer to the top of the dendrogram will tend to have the most dissimilarities, while the values that were grouped closer to the bottom of the dendrogram will have the most similarities. (James et al., 2021, p. 521-525).

When grouping data into clusters through Hierarchical Clustering, the following should be taken into consideration (James et al., 2021, p. 525-531):

- Dissimilarity measure – a measure that will evaluate the distance between observations. Examples of dissimilarity measures include Euclidean distance and correlation-based distance. For our analysis, Euclidean distance is chosen as a dissimilarity measure since in our analysis, we aim to group countries with overall similar values across energy poverty indicators. In case the goal of the analysis was to cluster countries based on the correlation of energy poverty indicators trends without regard to differences in energy poverty indicator values, then the correlation-based distance would have been chosen;

- Type of linkage – a measure that identifies dissimilarities between groups. Common types of linkage are Complete, Single, Average and Centroid. In order to identify the optimal type of linkage, we utilize the Agnes method of HC. Agnes is a bottom-up approach which starts with treating single observations as separate clusters and through a series of iterations combines the nearest clusters until there is one cluster consisting of all the observations. The benefit of the approach is that it provides an agglomerative coefficient measure of comparing clustering structures of different HC models. The agglomerative coefficient is an average of all $1 - m_i$ values, where i represents a single observation and mi is a dissimilarity measure for i in the first cluster merger divided by the last cluster merger (Maechler et al., 2022). In the analysis, we utilize the *agnes* function under the *cluster* package (Maechler et al., 2022) to evaluate a strong clustering structure by comparing agglomerative coefficients for models using different linkage types. We choose the type of linkage which showcases the highest agglomerative coefficient;

- The height of the dendrogram at which the clusters are to be cut – this will determine the number of clusters that will be received as an output of Hierarchical Clustering. The optimal number of clusters and, therefore, the height of the dendrogram cut is determined by running the Silhouette test;

- The requirement to scale the dataset to have a standard deviation one. As we proceed with Hierarchical Clustering after performing PCA on scaled data, this step is omitted.

The common methods for grouping data in homogenous clusters are K-means Clustering and Hierarchical Clustering. For our research, we choose the HC method as it allows to group data without having to pre-define the number of clusters prior to applying the method. In our analysis, we aim to identify the optimal number of clusters after applying HC by running the

Silhouette test. For this purpose, the *fviz_nbclust* function is utilized under the *factoextra* package (Kassambara & Mundt, 2020). The function allows to identify and visualize the optimal number of clusters through various methods. For our analysis, we choose the Silhouette method, which measures the average silhouette width of clusters and provides a visualization of the optimal number of clusters based on this measure (Kassambara & Mundt, 2020).

4.2 Cross-country energy poverty pattern similarities in the European Union

In Section 3, we identified the first four Principal Components which explain 79% of the variance in energy poverty indicator data. In this part of the analysis, we use four Principal Components as input to the Hierarchical Clustering method, the goal of which is to group EU countries that have the most similarities in energy poverty indicators.

As discussed in Subsection 4.1, we choose Euclidean distance as a dissimilarity measure. To identify a type of linkage to be utilized, the *agnes* function (Maechler et al., 2022) is run. Table 4.1 presents agglomerative coefficients of the HC method performed on our dataset using Average, Single, Complete and Ward types of linkage. The agglomerative coefficient of Ward linkage shows the value closest to one, therefore, the HC method with Ward linkage has the strongest clustering structure and is chosen to be utilized in our analysis.

Table 4.1: *Agglomerative coefficients of HC utilized with different linkage types*

Type of linkage	Average	Single	Complete	Ward
Agglomerative coefficient	0.7379951	0.6294240	0.8167799	0.8813094

The results of applying HC to the first four Principal Components are presented in Figures 4.1 and 4.2. Figure 4.1 showcases countries grouped in four clusters. The number of clusters in this case is chosen arbitrarily to review countries' groupings in a relatively small number of clusters. The initial assumption is made that four clusters could potentially consist of and distinguish the North of Europe and the South of Europe, "old EU" and "new EU" countries – therefore, four clusters are chosen as the first number of clusters to be tested.

Figure 4.1: *Hierarchical Clustering dendrogram with four clusters highlighted*

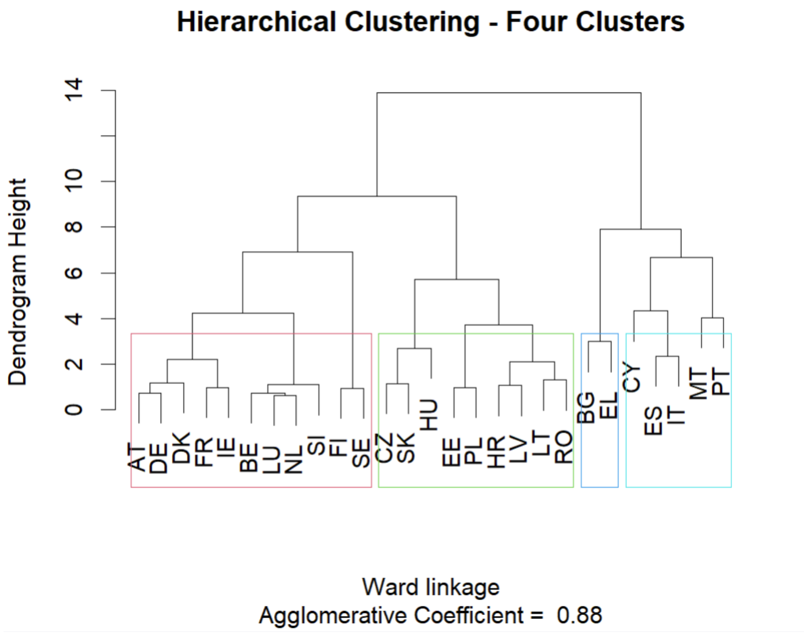
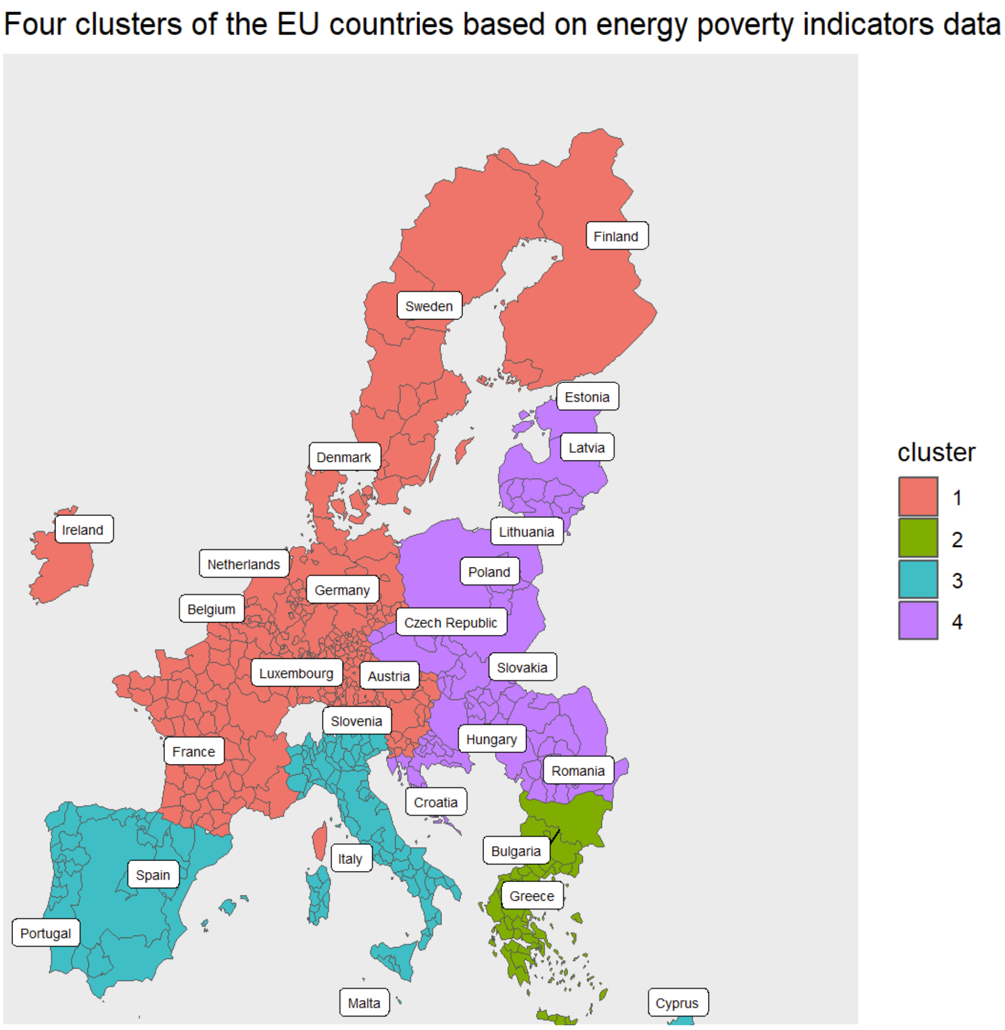


Figure 4.2: *The map of four clusters (based on source code by Povea (2023))*



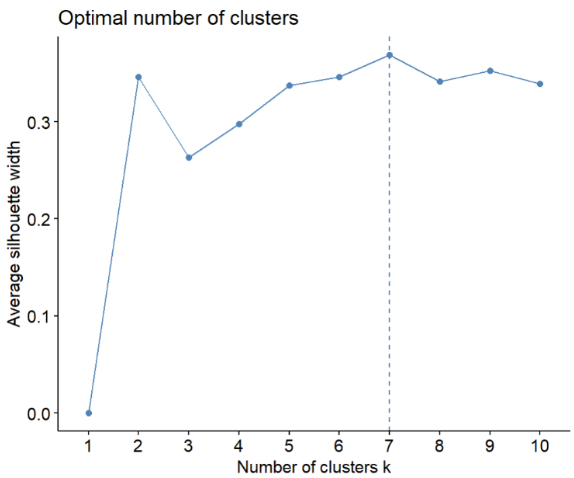
Countries are identified in clusters based on the combination of the indicators described above. However, this grouping also allows us to trace some similarities between countries in terms of their socio-economic and geographical features. For example, the majority of the countries of the “old EU” are assigned to the first, largest cluster. This cluster is distinguished by the concentration of countries with a high level of economic development and an emphasis on the growth of energy generation from renewable sources, which may have a positive relationship with the level of GDP in the state in the long run (Ohler & Fetters, 2014).

The fourth cluster is mostly comprised of post-Soviet and/or the “new EU” countries, which have significant economic and historical connections. The connections as strong can also be observed between the “Mediterranean” (plus Portugal) countries in the third cluster. The second cluster consists of only two countries – Bulgaria and Greece, the states that were dramatically affected by the 2008 economic crisis.

When it comes to geography, clustering also shows certain patterns. Thus, the first cluster includes the countries of Western and Northern Europe, the fourth – the eastern and central part of the European Union, the second and third – the southern states.

In order to identify the optimal number of clusters, the Silhouette method described in Subsection 4.1 is performed. Figure 4.3 presents the result of applying the Silhouette method, which suggests that seven is the optimal number of clusters for our data in comparison to the initially chosen four clusters.

Figure 4.3: *The Silhouette method result for optimal number of clusters identification*



Figures 4.4 and 4.5 present seven clusters identified through the Agnes Hierarchical Clustering method with Euclidean distance as a dissimilarity measure and Ward linkage type.

Figure 4.4: *Hierarchical Clustering dendrogram with seven clusters highlighted*

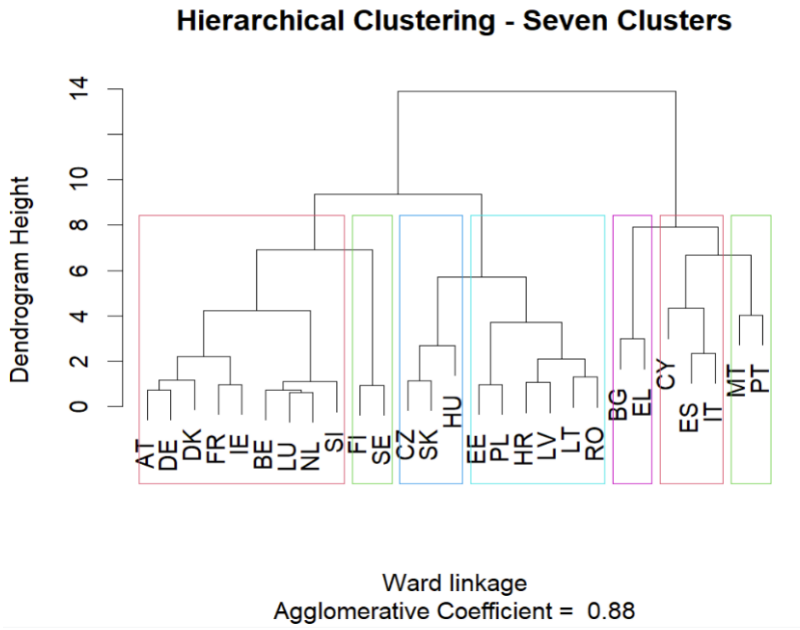
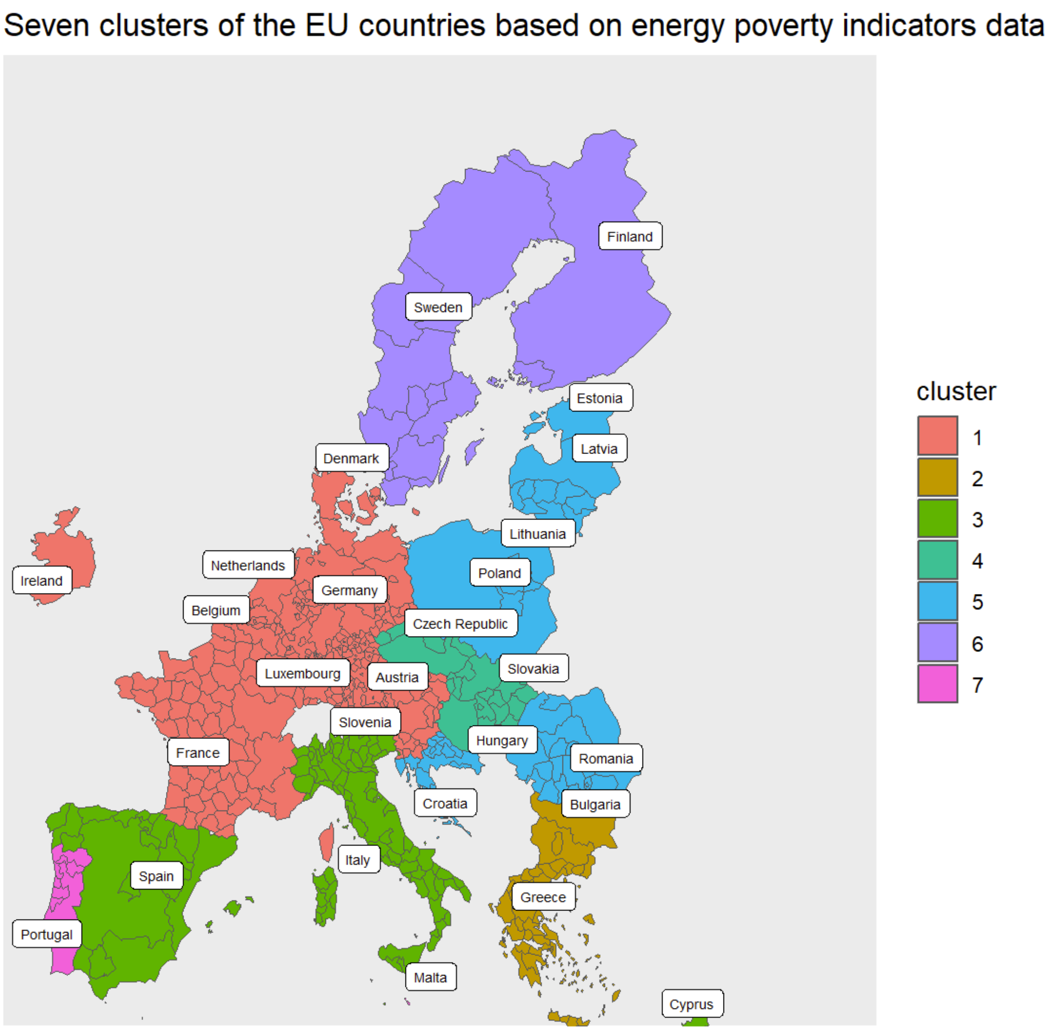


Figure 4.5: *The map of seven clusters (based on source code by Povea (2023))*



The division into seven clusters is much more granular, which allows us to look at the geographic features of the manifestation of energy poverty in Europe somewhat more accurately.

As it was when identifying four clusters, the first, largest cluster is formed by the countries of the “old EU”. However, now two Nordic countries, Sweden, and Finland, have been singled out from this set into a separate sixth cluster (note that Denmark is still included in the first cluster). This observation can be explained by several factors that unite these countries. For example, they are the European leaders in terms of the share of renewables in the energy mix (Eurostat, 2023b), and both continue to have pro-nuclear power policies even after the Fukushima nuclear power accident of 2011 (Litmanen et al., 2017).

Notably, the second cluster remains the same, thus indicating a great socio-economic similarity between Greece and Bulgaria. For instance, it is suggested that the energy poverty situation in these two countries is equally aggravated by energy affordability and energy efficiency difficulties (Dubois & Meier, 2016). Furthermore, reports by The World Bank imply that achieving the goals of energy efficiency, economic growth, and gas emissions reduction presents significant challenges for these two states (Obradović & Lojanica, 2017).

The third cluster is still comprised of the hot Southern countries with a coastline in the Mediterranean (Spain, Italy, and Cyprus), with Portugal and Malta, among the few countries in Europe with increased demand for electricity in 2022 (Jones, 2023), standing apart.

The states that were previously included in one cluster when dividing the EU countries into four groups now belong to two clusters. The Czech Republic, Slovakia and Hungary, the landlocked EU countries with a high dependence on gas supply (Beyer & Molnar, 2022), form a separate cluster, while other countries stay in the same cluster. In terms of geography, the regional pattern remains almost unchanged in comparison to the four-cluster division, meaning that neighboring countries indeed share similarities in the context of energy poverty vulnerability.

5. Household electricity price elasticity

The following section aims to review energy poverty from a different angle – from the perspective of households in different clusters of EU countries and their vulnerability to changes in electricity prices. Subsection 5.1 describes electricity price elasticity models built to identify the effect of changes in electricity prices and other social, economic, environmental, and energy indicators on household electricity consumption. In the analysis, countries are represented by clusters, derived using the HC model from Section 4. Subsection 5.2 discusses the results of applying the price elasticity model and concludes whether the comparison of household electricity consumption in different clusters is relevant based on the model results.

5.1 Price elasticity model

In the following part of the research, we aim to look at the energy poverty phenomenon in the EU from a household perspective. In the analysis, we utilize Eurostat (n.d., a) indicators in Population and Social Conditions, Economy and Finance, Environment and Energy, and Macroeconomic imbalance procedure sections as well as Household electricity prices indicator by European Commission (n.d., b) which are listed in Subsection 3.1. The goal of this part of our analysis is to estimate the price elasticity of energy consumption by households. Specifically, we are to compare the electricity price elasticity for household electricity consumption across clusters, derived from applying the HC model to energy poverty indicator data, to identify the most vulnerable clusters and income groups to changes in electricity price.

We limit the indicators data by the time period from 2015 to 2020 to accommodate for the missing variables for electricity consumption before the year 2015 and for the year 2021 for the majority of EU countries. The data for the selected period has six missing observations for the Disaggregated final energy consumption in households indicator for Belgium, Cyprus and Lithuania. Five observations are missing for the Distribution of income by quantiles indicator for Slovakia. The number of missing observations is relatively small compared to the size of the dataset. Moreover, as the focus of the analysis is to compare electricity price elasticity across clusters, the missing observations are to be mitigated by the present observations for other countries in the same cluster. Therefore, the decision is made to remove rows with missing observations. The final list of variables utilized in the price elasticity model is included in Table 5.1.

Table 5.1: *Electricity price elasticity model variables (Eurostat, n.d., a)*

Variable	Type of variable	Indicator
Electricity consumption	Continuous	Disaggregated final energy consumption in households – quantities (Eurostat, n.d., a)
Household electricity prices	Continuous	Household electricity prices (European Commission, n.d., b)
Number of households	Continuous	Number of households by household composition, number of children and working status within households (Eurostat, n.d., a.)
Persons per household	Continuous	Average number of persons per household-by-household composition, number of children and working status within households (Eurostat, n.d., a)
Unemployment rate	Continuous	Unemployment rate – annual data (Eurostat, n.d., a)
Household income in 1st quantile	Continuous	Distribution of income by quantiles – EU-SILC and ECHP surveys (Eurostat, n.d., a)
GDP	Continuous	GDP and main components (output, expenditure, and income) (Eurostat, n.d., a)
Cluster	Categorical	Cluster of the EU countries

Chai et al. (2021) build a model to identify the price elasticity of household electricity demand in Queensland regions, Australia. The goal of the model is to estimate the effect of electricity price change on households' electricity consumption. The authors perform translog regression on household energy consumption data for the regions with electricity average price for households as one of the independent variables. Chai et al. (2021) perform log transformation on the dependent variable electricity consumption as well as continuous independent variables electricity average price, housing costs and disposable income. The authors include the squared term of electricity price and a list of categorical independent variables, which describe the type of households and households' geographical location.

In our research, we follow the approach of Chai et al. (2021) and start with building a model as showcased in Figure 5.1. We log transform continuous variables including household electricity consumption, household income for 1st quantile, GDP, number of households, unemployment rate, persons per household and household electricity price. This model does not include cluster variable and intends to review the relations between dependent and independent variables without distinguishing observations for countries in different clusters.

Figure 5.1: *Electricity price elasticity model of household electricity demand without cluster variable*

$$
\begin{aligned}
\log(household\ electricity\ consumption_i) \\
&= \beta_0 + \beta_1 \log(household\ income\ 1st\ quantile_i) + \beta_2 \log(GDP_i) \\
&+ \beta_3 \log(number\ of\ households_i) + \beta_4 \log(unemployment\ rate_i) \\
&+ \beta_5 \log(persons\ per\ household_i) + \beta_6 \log(household\ electricity\ price_i) + u_i
\end{aligned}
$$

Secondly, we change the initial model by adding a cluster variable and an interaction term for cluster and electricity price variables to identify whether trends of household electricity price changes are dependent on a cluster. The transformed model is described in Figure 5.2. Our goal is to verify whether one unit change in the logarithm value of electricity price is as well having an effect on the cluster variable. The assumption is made that in different clusters the increase of household electricity prices on a yearly basis will have different magnitude. As the final step, we compare the results of models described in Figures 5.1 and 5.2 and conclude on the relevance of differentiating household electricity consumption in different clusters.

Figure 5.2: *Electricity price elasticity model of household electricity demand with an interaction term for cluster and log value of electricity price*

$$\begin{aligned}\log(household\ electricity\ consumption_i) \\ &= \beta_0 + \beta_1 \log(household\ income\ 1st\ quantile_i) + \beta_2 \log(GDP_i) \\ &+ \beta_3 \log(number\ of\ households_i) + \beta_4 \log(unemployment\ rate_i) \\ &+ \beta_5 \log(persons\ per\ household_i) + \beta_6 \log(household\ electricity\ price_i) \\ &+ \beta_7 cluster_i + \beta_8 \log(household\ electricity\ price_i)\ cluster_i + u_i\end{aligned}$$

$$cluster_i \begin{cases} 1\ if\ i\ is\ an\ observation\ for\ a\ country\ in\ 2nd\ cluster \\ 0\ if\ i\ is\ an\ observation\ for\ a\ country\ in\ 1st\ cluster \end{cases}$$

In order to simplify the model and the comprehension of interaction terms for clusters, the decision is made to decrease the number of HC clusters, identified in section 4.2, to two clusters as per Figure 5.3 and Figure 5.4. According to Figure 4.3, the average silhouette width of clusters for a model with two clusters is among the highest together with six-, seven- and nine-cluster models. Therefore, the two-cluster model provides a relatively optimal clustering structure and is beneficial for comparing log-log models described in Figure 5.1 and Figure 5.2 due to the lower number of interaction terms to consider.

Figures 5.3 and 5.4 present the division of the EU countries into two clusters. One can notice a prominent geographic trend in such clustering: the second and smallest cluster is comprised of Southern European countries – Portugal, Spain, Slovenia, Italy, Malta, Bulgaria, Greece, and Cyprus. Hence, energy poverty distribution in Europe might indeed be dictated by the geographic and socio-economic peculiarities of specific regions, an assumption further tested in this thesis.

Figure 5.3: *Hierarchical Clustering dendrogram with two clusters highlighted*

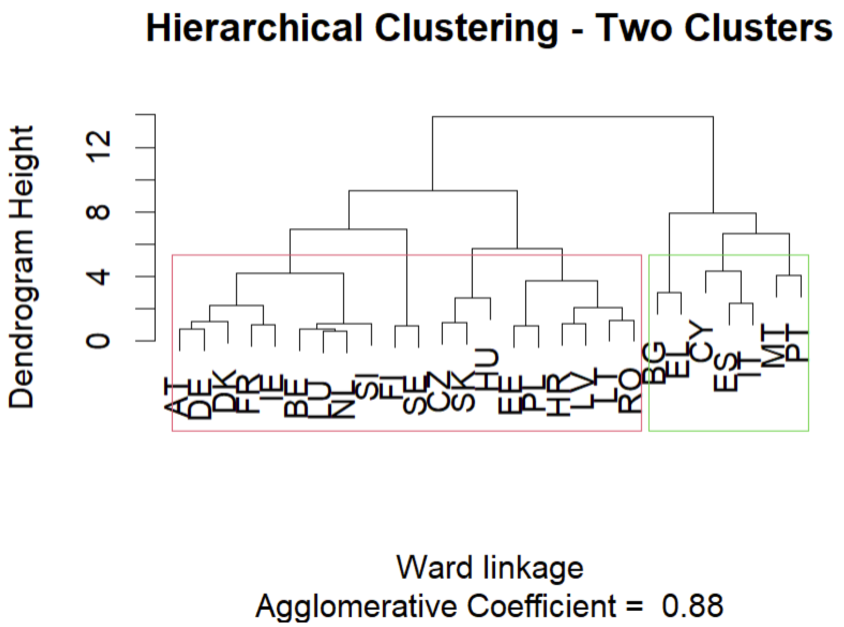
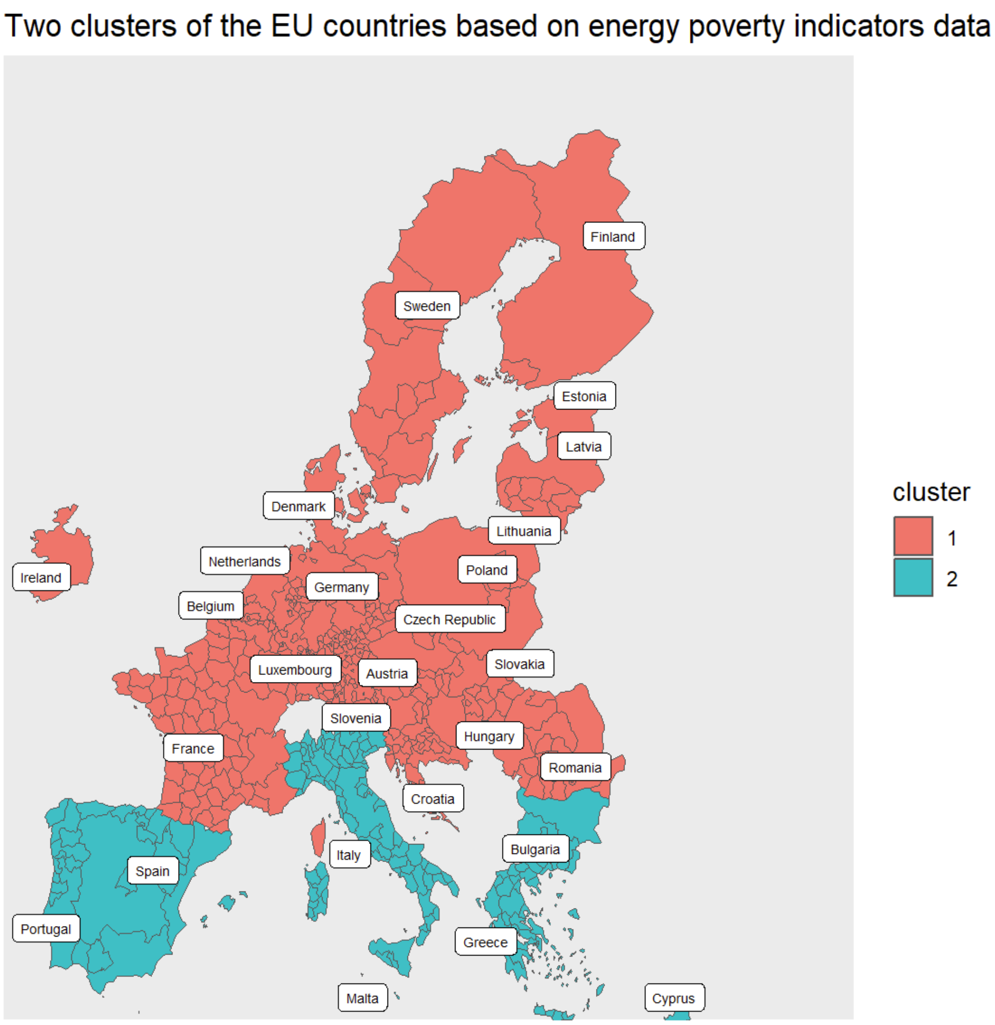


Figure 5.4: *The map of two clusters (based on source code by Povea (2023))*



5.2 Electricity price elasticity of household energy consumption

The following Subsection describes the results of fitting the electricity price elasticity of household energy consumption models presented in Subsection 5.1 to the dataset. Firstly, data is fitted to the electricity price elasticity model without cluster variable as per Figure 5.1. Secondly, a model with an interaction term consisting of the cluster and household electricity prices variables is tested. The summaries for the two models are showcased in Tables 5.2 and 5.3.

Table 5.2: *Summary of the electricity price elasticity model of household electricity consumption without an interaction term and cluster variable*

Coefficients:				
	Estimate	**Std. Error**	**t value**	**P-value (> | t |)**
Intercept	-1.79926	0.62323	-2.887	0.00447**
log(household income 1st quantile)	1.00249	0.15549	6.447	1.49e-09***
log(GDP)	0.24279	0.05867	4.138	5.83e-05***
log(number of households)	0.73818	0.05368	13.752	< 2e-16***
log(unemployment rate)	0.39878	0.05272	7.564	3.74e-12***
log(persons per household)	-1.30806	0.22066	-5.928	2.05e-08***
log(household electricity price)	-0.27582	0.11378	-2.424	0.01655*
Signif. Codes: 0 - *** 0.001 - ** 0.01 - *				
Residual standard error: 0.2605 on 149 degrees of freedom Multiple R-squared: 0.966 Adjusted R-squared: 0.9647 F-statistic: 706 on 6 and 149 DF, p-value: < 2.2e-16				

The summary of the first model, which is presented in Table 5.2, points to the following relations between variables:

- For all the independent variables we observe a p-value < 0.05. All independent variables in the model have a statistically significant effect on the dependent variable – household electricity consumption;

- Household income for the quantile 1 variable has a statistically significant positive effect on household electricity consumption – a 1% increase in household income in quantile 1 is followed by a 1% increase in household electricity consumption;

- There is a statistically significant and positive dependence between GDP and household electricity consumption – a 1% increase in GDP value leads to a 0.24% increase in household electricity consumption;

- For the number of households variable, we observe a statistically significant and positive correlation with household electricity consumption – a 1% increase in the number of households is reflected by a 0.74% increase in household electricity consumption;

- The unemployment rate independent variable has a statistically significant positive correlation with household electricity consumption – a 1% increase in the unemployment rate is followed by a 0.40% increase in household electricity consumption;

- Persons per household has a statistically significant and negative effect on household electricity consumption – a 1% increase in the persons per household indicator is reflected by a 1.31% decrease in household electricity consumption;

- Household electricity price showcases a statistically significant negative effect on household electricity consumption – a 1% increase in household electricity prices leads to a 0.28% decrease in household electricity consumption.

Table 5.3: *Summary of the electricity price elasticity model of household electricity consumption with an interaction term for cluster and log value of electricity price*

Coefficients:	**Estimate**	**Std. Error**	**t value**	**P-value (> | t |)**
Intercept	-2.25723	0.56983	-3.961	0.000116***
log(household income 1st quantile)	1.18416	0.14427	8.208	1.05e-13***
log(GDP)	0.26107	0.05315	4.912	2.37e-06***
log(number of households)	0.73776	0.04866	15.162	< 2e-16***
log(unemployment rate)	0.40339	0.05139	7.850	7.94e-13***
log(persons per household)	-1.30841	0.20385	-6.419	1.78e-09***
Cluster 2	-0.95194	0.28037	-3.395	0.000882***
log(household electricity price)	-0.16492	0.11294	-1.460	0.146376
Cluster 2 : log(household electricity price)	-0.65095	0.15143	-4.299	3.11e-05***

Signif. Codes:
0 - ***
0.001 - **
0.01 - *

Residual standard error: 0.2336 on 147 degrees of freedom
Multiple R-squared: 0.973
Adjusted R-squared: 0.9716
F-statistic: 662.9 on 8 and 147 DF, p-value: < 2.2e-16

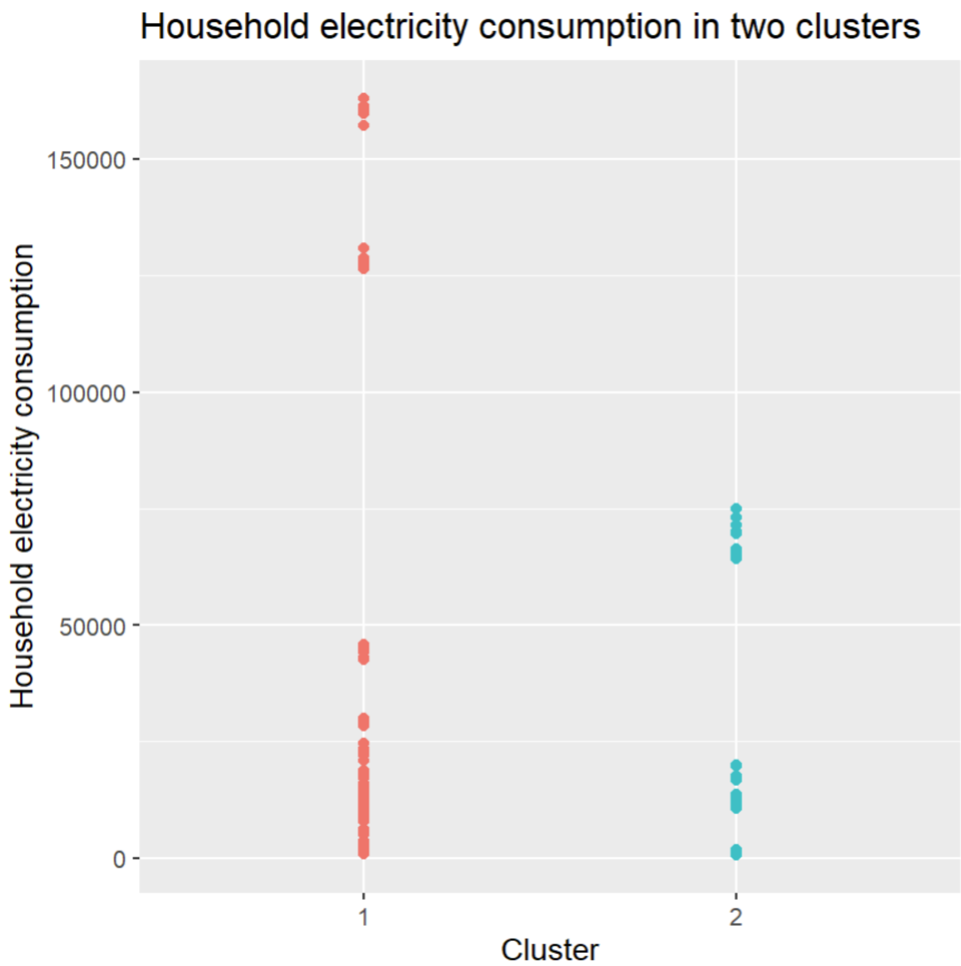
As per the summary, showcased in Table 5.3, the following observations are made in regard to the electricity price elasticity model of household electricity demand with an interaction term for cluster and log value of electricity price:

- For all the independent variables except for household electricity prices, we observe a p-value < 0.05. All independent variables in the model except for household electricity prices have a statistically significant effect on the dependent variable – household electricity consumption;
- Household income for the quantile 1 variable has a statistically significant positive effect on household electricity consumption – a 1% increase in household income in quantile 1 is followed by a 1.18% increase in household electricity consumption;
- There is a statistically significant and positive dependence between GDP and household electricity consumption – a 1% increase in GDP value leads to a 0.26% increase in household electricity consumption;
- For the number of households variable, we observe a statistically significant and positive correlation with household electricity consumption – a 1% increase in the number of households is reflected by a 0.74% increase in household electricity consumption;
- The unemployment rate independent variable has a statistically significant positive correlation with household electricity consumption – a 1% increase in the unemployment rate is followed by a 0.40% increase in household electricity consumption;
- Persons per household has a statistically significant and negative effect on household electricity consumption – a 1% increase in persons per household indicator is reflected by a 1.31% decrease in household electricity consumption;
- Household electricity price showcases no statistically significant correlation with household electricity consumption;
- Based on a p-value < 0.05, the cluster of countries to which households belong has a significant effect on changes in household electricity consumption. Cluster 2 countries compared to countries in Cluster 1 have a significant negative effect on the percent

change of household electricity consumption. On average, household electricity consumption in Cluster 2 is 61.4% less than in Cluster 1. This trend can also be observed in Figure 5.5, showcasing electricity consumption for households in countries in Cluster 1 and Cluster 2;

- The interaction term between cluster and logarithm value of electricity prices shows a p-value < 0.05, which indicates that there is a statistically significant evidence of dependence between cluster and logarithm value of electricity prices, and the addition of the interaction term improved the electricity price elasticity model. Based on the estimates of the model, the household electricity consumption for countries in Cluster 2 decreases by 0.81% when household electricity prices increase by 1%. For Cluster 1, household electricity consumption decreases by 0.16% when household electricity prices increase by 1%. This indicates that households in Cluster 2 are more responsive to the changes in electricity prices compared to households in Cluster 1.

Figure 5.5: *Household electricity consumption in two given clusters*



Comparing the two models, we observe that the coefficient values for independent variables household income for quantile 1, GDP, number of households, unemployment rate, and persons per household are similar. However, we see a difference in the interpretation of the household electricity price effect on household electricity consumption when we add control for clusters in the model. The model that adds an interaction term between the logarithm value of household electricity prices and cluster points to differences in household consumption trends when household electricity prices increase. In Cluster 2, household electricity consumption decreases by 0.65% more when household electricity prices increase by 1% than household electricity consumption in Cluster 1.

We conclude that the addition of control for cluster in the model helps to identify different electricity consumption trends in different clusters. Moreover, observation of electricity consumption behavior in households from different clusters is helpful in determining in which countries of the EU households will be more affected by the surging electricity prices.

6. Discussion

6.1 Main findings

Having gathered indicators from Eurostat (Eurostat, n.d., a) and European Commission (European Commission, n.d., b) for our research, we have analyzed the economic, financial, social, environmental, and energy factors of energy poverty. We performed PCA on data, which resulted in four Principal Components. Further on, we have utilized those components in hierarchical clustering. Based on HC results, we have concluded that the most optimal way of grouping is seven categories. We have spotted strong geographical trends within clusters, with energy poverty susceptibility differing between the "old" and "new" EU countries and Western European countries forming the biggest cluster. We have also noticed that in general, there are commonalities between neighboring states' vulnerability to energy poverty.

In the second part of the analysis, we have looked at energy poverty from the household perspective. We have formulated a household electricity price elasticity model to identify households' behavior and their reaction to energy price changes. We have observed that in different clusters, the effect varies notably. This finding confirms the importance of region-specific approach to addressing energy poverty.

We have verified the relevance of differentiating EU countries based on their cluster attribution. The addition of clustering improves our model of electricity price elasticity by identifying the countries that might be more susceptible to the drastic consequences of price fluctuations. The Southern European countries are more prone to reductions in electricity consumption when prices rise, meaning that these states (Portugal, Spain, Slovenia, Italy, Malta, Bulgaria, Greece, and Cyprus) are more vulnerable to energy poverty as well. This is explained by the fact that household electricity consumption in the above-mentioned countries, compared to the countries from the other cluster, drops by 0.65% more when household electricity prices increase by 1%.

6.2 Limitations

While our thesis successfully deals with the question of finding regional energy poverty pattern similarities in the European Union and examining the vulnerability of different income groups of European households to energy poverty, there are some limitations in this study worth considering.

First of all, there is still a certain degree of ambiguity associated with the term “energy poverty” itself, as it is quite broad and often used interchangeably with “fuel poverty”. Consequently, some aspects of energy poverty in the study may be attributed to “regular” poverty as well.

When it comes to the data utilized, the availability and granularity of the information obtained also led to certain limitations. Some indicators are collected once every five years, whereas a major part of them – on an annual basis, with indicators having different years for which the most recent observations are retrieved. Due to this issue, seven out of twenty-one Eurostat’s energy poverty indicators were excluded from PCA, which also required the simulation of missing values in the datasets.

Another limitation is related to the scope and depth of the study. The observation in this thesis is relatively broad and “homogeneous”, without any distinct focus either on the crisis years (2008 and 2020 in particular) or specific countries and regions. Therefore, this thesis does not go deeper into the “local” details and provides a wide yet comprehensive perspective on the energy poverty situation in Europe.

6.3 Suggestions for further research

During the process of writing this thesis, several related issues that are potentially interesting for further research were identified.

To begin with, it might be of great interest to add several income quantiles to the electricity price elasticity model and compare energy poverty indicators based on household income groups, instead of choosing only one income quantile. This way the research on households’ energy poverty vulnerability can get even more economically detailed.

Then, it can be worth increasing the number of clusters in the model. This approach will allow scientists to make more qualitative research with deepening intro national aspects of energy poverty. Such a study can create basis for implication of clustering model to policy-making that will result in more well-thought and region-specific policy recommendations.

Moreover, it is possible to elaborate the electricity price elasticity model by taking development of renewables into consideration. Having included it as a factor of comparison, one can analyze, for instance, correlation between renewable energy adoption and energy poverty alleviation and formulate a number of predictions regarding evolution of energy markets in Europe.

7. Conclusion

This master's thesis sought to identify common patterns of energy poverty throughout the European Union's member states and examine how susceptible various income European households are to energy poverty, especially during periods of high energy costs. In order to do so, a literature and background review, Principal Component Analysis, Hierarchical Clustering, and estimation of electricity price elasticity of household energy consumption were carried out in this study.

Energy poverty, "a set of domestic energy circumstances that do not allow for participating in the lifestyles, customs and activities that define membership of society" (Bouzarovski & Petrova, 2015), is a problem difficult to measure with one or even several indicators. Due to this complexity and a multifaceted nature of the issue, it is important yet challenging to study energy poverty. However, research in this area is useful for many people: politicians, policymakers, industry specialist, scientists, to name but a few. All this motivated our choice of the thesis topic.

In our analysis, we have decided to use approaches of Recalde et al. (2019) and Chai et al. (2021) and examine energy poverty from multiple perspectives. The goal of our research was to find energy poverty similarities for different countries in the EU by conducting PCA and HC. As a result of the first half of the research, we have determined clusters of countries with similar energy poverty indicators.

In the second part of the analysis, our aim was to see household electricity consumption behavior in those clusters. We have created households' electricity price elasticity model and identified higher reaction of households in the second cluster (Portugal, Spain, Slovenia, Italy, Malta, Bulgaria, Greece, and Cyprus) in terms of electricity consumption to changes in electricity prices. Hence, we confirmed how crucial it is to have a region-specific approach when trying to mitigate and combat risks brought up by energy poverty.

We believe that our paper strongly contributes to the ongoing research of such a complex concept of energy poverty and its implications on the welfare of European citizens. Our analysis forms a basis for further research in different energy poverty patterns in the EU and the substantiality of various policymaking efforts in different parts of Europe.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs Computational Statistics, 4*(2), 433-459. https://doi.org/10.1002/wics.101

Beyer, S., & Molnar, G. (2022, September 14). *Accelerating energy diversification in Central and Eastern Europe*. IEA. Retrieved April 21, 2023, from https://www.iea.org/commentaries/accelerating-energy-diversification-in-central-and-eastern-europe

Boardman, B. (1991). Fuel poverty: from cold homes to affordable warmth. Belhaven Press, London

Boardman, B. (2009). Fixing Fuel Poverty: Challenges and Solutions (1st ed.). Routledge. https://doi.org/10.4324/9781849774482

Bouzarovski, S., Petrova, S., & Tirado-Herrero, S. (2014). From fuel poverty to energy vulnerability: The Importance of Services, Needs and Practices. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2743143

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

Bradshaw, J., & Hutton, S. (1983). Social policy options and fuel poverty. *Journal of Economic Psychology*, *3*(3-4), 249–266. https://doi.org/10.1016/0167-4870(83)90005-3

Buzar, S. (2007). Energy Poverty in Eastern Europe: Hidden Geographies of Deprivation (1st ed.). Routledge. https://doi.org/10.4324/9781315256504

Castaño-Rosa, R., Solís-Guzmán, J., Rubio-Bellido, C., & Marrero, M. (2019). Towards a multiple-indicator approach to energy poverty in the European Union: A review. *Energy and Buildings*, *193*, 36–48. https://doi.org/10.1016/j.enbuild.2019.03.039

Chai, A., Ratnasiri, S., & Wagner, L. (2021). The impact of rising energy prices on energy poverty in Queensland: A microsimulation exercise. *Economic Analysis and Policy, 71,* 57-72. https://doi.org/10.1016/j.eap.2021.03.014

Cook, J. T., & Frank, D. A. (2008). Food Security, Poverty, and Human Development in the United States. *Annals of the New York Academy of Sciences*, *1136*(1), 193–209. https://doi.org/10.1196/annals.1425.001

Cozzi, L., Wetzel, D., Tonolo, G., & Hyppolite, J. (2022, November 3). *For the first time in decades, the number of people without access to electricity is set to increase in 2022*. IEA. Retrieved March 30, 2023, from https://www.iea.org/commentaries/for-the-first-time-in-decades-the-number-of-people-without-access-to-electricity-is-set-to-increase-in-2022

Csereklyei, Z. (2020). Price and income elasticities of residential and industrial electricity demand in the European Union. *Energy Policy*, *137*, 111079. https://doi.org/10.1016/j.enpol.2019.111079

Curtis, J., & Pentecost, A. (2015). Household fuel expenditure and residential building energy efficiency ratings in Ireland. *Energy Policy*, *76*, 57–65. https://doi.org/10.1016/j.enpol.2014.10.010

Dalla Longa, F., Sweerts, B., & van der Zwaan, B. (2021). Exploring the complex origins of energy poverty in the Netherlands with machine learning. *Energy Policy*, *156*, 112373. https://doi.org/10.1016/j.enpol.2021.112373

Database. Eurostat. (n.d., a). Retrieved March 20, 2023, from https://ec.europa.eu/eurostat/web/main/data/database

Doremus, J. M., Jacqz, I., & Johnston, S. (2022). Sweating the energy bill: Extreme weather, poor households, and the energy spending gap. *Journal of Environmental Economics and Management*, *112*, 102609. https://doi.org/10.1016/j.jeem.2022.102609

Dowle, M., & Srinivasan, A. (2023). *data.table: Extension of `data.frame`. R package version 1.14.8*. CRAN. https://CRAN.R-project.org/package=data.table

Dubois, U., & Meier, H. (2016). Energy affordability and energy inequality in Europe: Implications for policymaking. *Energy Research & Social Science*, *18*, 21–35. https://doi.org/10.1016/j.erss.2016.04.015

Economic Well-Being of U.S. Households in 2021: Appendixes. The Federal Reserve System. (2022, May). Retrieved March 24, 2023, from https://www.federalreserve.gov/publications/files/2021-supplement-economic-well-being-us-households-202205.pdf

Energy poverty. European Commission. (n.d., a). Retrieved March 20, 2023, from https://energy.ec.europa.eu/topics/markets-and-consumers/energy-consumer-rights/energy-poverty-eu_en

Excess mortality - statistics. Eurostat. (2023a, April 12). https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Excess_mortality_-_statistics

Faiella, I., & Lavecchia, L. (2019). Chapter 6 - Energy poverty indicators. In *Urban Fuel Poverty* (pp. 127–141). Academic Press. https://doi.org/10.1016/B978-0-12-816952-0.00006-5

Firke, S. (2023). *janitor: Simple Tools for Examining and Cleaning Dirty Data*. R package version 2.2.0. CRAN. https://CRAN.R-project.org/package=janitor

Frank, D. A., Neault, N. B., Skalicky, A., Cook, J. T., Wilson, J. D., Levenson, S., Meyers, A. F., Heeren, T., Cutts, D. B., Casey, P. H., Black, M. M., & Berkowitz, C. (2006). Heat or Eat: The Low Income Home Energy Assistance Program and Nutritional and Health Risks Among Children Less Than 3 Years of Age. *Pediatrics*, *118*(5), e1293–e1302. https://doi.org/10.1542/peds.2005-2943

Gaye, A. (2008). Access to Energy and Human Development. Human Development Occasional Papers (1992-2007) HDOCPA-2007-25, Human Development Report Office (HDRO), United Nations Development Programme (UNDP).

Glossary: Income quintile group. Eurostat. (n.d., b). Retrieved April 7, 2023, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Income_quintile_group

Grolemund, G., Wickham, H. (2011). Dates and Times Made Easy with lubridate. *Journal of Statistical Software, 40*(3), 1-25. https://www.jstatsoft.org/v40/i03/

Hassan, S. T., Batool, B., Zhu, B., & Khan, I. (2022). Environmental complexity of globalization, education, and income inequalities: New insights of energy poverty. *Journal of Cleaner Production*, *340*, 130735. https://doi.org/10.1016/j.jclepro.2022.130735

Hastie, T., Mazumder, R. (2021, May 9). *Matrix Completion via Iterative Soft-Thresholded SVD*. CRAN. https://cran.r-project.org/web/packages/softImpute/softImpute.pdf

Healy, J. D. (2003). Excess winter mortality in Europe: A cross country analysis identifying key risk factors. *Journal of Epidemiology & Community Health*, *57*(10), 784–789. https://doi.org/10.1136/jech.57.10.784

Healy, J. D. (2004). Housing, Fuel Poverty and Health: A Pan-European Analysis. Aldershot, Ashgate

Household Economic Wellbeing and Progress. Australian Bureau of Statistics. (2010, September 15). Retrieved March 24, 2023, from https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1370.0~2010~Chapter~Household%20economic%20wellbeing%20and%20progress%20(5.3.1)

Households' economic well-being: the OECD dashboard. OECD. (n.d., a). Retrieved March 24, 2023, from https://www.oecd.org/sdd/na/household-dashboard.htm

Huebner, G. M., & Shipworth, D. (2017). All about size? – The potential of downsizing in reducing energy demand. *Applied Energy*, *186*(2), 226–233. https://doi.org/10.1016/j.apenergy.2016.02.066

Income inequality. OECD. (n.d., b). Retrieved March 24, 2023, from https://doi.org/10.1787/459aa7f1-en

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer.

Jolliffe, I.T. (2002). *Principal Component Analysis* (2nd ed.). Springer-Verlag New York.

Jones, D. (2023, January 31). *European Electricity Review 2023*. Ember. Retrieved April 21, 2023, from https://ember-climate.org/insights/research/european-electricity-review-2023/

Kassambara, A., & Mundt, F. (2020, April 1). *Extract and Visualize the Results of Multivariate Data Analyzes.* CRAN. https://cloud.r-project.org/web/packages/factoextra/factoextra.pdf

Kiss, M., Negreiro, M., Niestadt, M., Nijenhuis, C., Van Lierop, C., & Van der Linde, A. (2022, May). *Demographic outlook for the European Union 2022*. Think Tank | European Parliament. Retrieved March 24, 2023, from https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2022)729461

Kuhn, M., Jackson, S., Cimentada, J. (2022). *corrr: Correlations in R*. R package version 0.4.4. CRAN. https://CRAN.R-project.org/package=corrr

Labandeira, X., Labeaga, J. M., & López-Otero, X. (2012). Estimation of elasticity price of electricity with incomplete information. *Energy Economics*, *34*(3), 627–633. https://doi.org/10.1016/j.eneco.2011.03.008

Labandeira, X., Labeaga, J. M., & López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy Policy*, *102*, 549–568. https://doi.org/10.1016/j.enpol.2017.01.002

Lacroix, E., & Chaton, C. (2015). Fuel poverty as a major determinant of perceived health: the case of France. *Public Health*, *129*(5), 517–524. https://doi.org/10.1016/j.puhe.2015.02.007

Liddell, C., & Guiney, C. (2015). Living in a cold and damp home: frameworks for understanding impacts on mental well-being. *Public Health*, *129*(3), 191–199. https://doi.org/10.1016/j.puhe.2014.11.007

Liddle, B., Smyth, R., & Zhang, X. (2020). Time-varying income and price elasticities for Energy Demand: Evidence from a middle-income panel. *Energy Economics*, *86*, 104681. https://doi.org/10.1016/j.eneco.2020.104681

Litmanen, T., Kari, M., Kojo, M., & Solomon, B. D. (2017). Is there a Nordic model of final disposal of spent nuclear fuel? Governance insights from Finland and Sweden. *Energy Research & Social Science*, *25*, 19–30. https://doi.org/10.1016/j.erss.2016.10.009

Maechler, M., Rousseeuw, P., Struyf, A., & Hubert, M. (2022, August 22). *"Finding Groups in Data'": Cluster Analysis Extended Rousseeuw et al.* CRAN. https://cran.r-project.org/web/packages/cluster/cluster.pdf

Margaras, V. (2019, January). *Demographic trends in EU regions*. European Parliamentary Research Service. Retrieved March 24, 2023, from https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/633160/EPRS_BRI(2019)633160_EN.pdf

Masselot, P., Mistry, M., Vanoli, J., Schneider, R., Iungman, T., Garcia-Leon, D., Ciscar, J.-C., Feyen, L., Orru, H., Urban, A., Breitner, S., Huber, V., Schneider, A., Samoli, E., Stafoggia, M., de'Donato, F., Rao, S., Armstrong, B., Nieuwenhuijsen, M., … Aunan, K. (2023). Excess mortality attributed to heat and cold: a health impact assessment study in 854 cities in Europe. *The Lancet Planetary Health*, *7*(4), e271–e281. https://doi.org/10.1016/s2542-5196(23)00023-2

Massicotte, P., South, A. (2023). *rnaturalearth: World Map Data from Natural Earth.* R package version 0.3.2. CRAN. https://CRAN.R-project.org/package=rnaturalearth

Maxim, A., Mihai, C., Apostoaie, C.-M., Popescu, C., Istrate, C., & Bostan, I. (2016). Implications and measurement of Energy Poverty Across the European Union. *Sustainability*, *8*(5), 483. https://doi.org/10.3390/su8050483

Meyer, B. D., & Sullivan, J. X. (2011). Viewpoint: Further results on measuring the well-being of the poor using income and consumption. *The Canadian Journal of Economics / Revue Canadienne d'Economique*, *44*(1), 52–87. http://www.jstor.org/stable/41336351

Miller, M., & Alberini, A. (2016). Sensitivity of price elasticity of demand to aggregation, unobserved heterogeneity, price trends, and price endogeneity: Evidence from U.S. Data. *Energy Policy*, *97*, 235–249. https://doi.org/10.1016/j.enpol.2016.07.031

National indicators. European Commission. (n.d., b). Retrieved March 20, 2023, from https://energy-poverty.ec.europa.eu/observing-energy-poverty/national-indicators_en

Nussbaumer, P., Bazilian, M., & Modi, V. (2012). Measuring energy poverty: Focusing on what matters. *Renewable and Sustainable Energy Reviews*, *16*(1), 231–243. https://doi.org/10.1016/j.rser.2011.07.150

Obradović, S., & Lojanica, N. (2017). Energy use, CO_2 emissions and economic growth – causality on a sample of SEE countries. *Economic Research-Ekonomska Istraživanja*, *30*(1), 511–526. https://doi.org/10.1080/1331677x.2017.1305785

Ohler, A., & Fetters, I. (2014). The causal relationship between renewable electricity generation and GDP growth: A study of energy sources. *Energy Economics*, *43*, 125–139. https://doi.org/10.1016/j.eneco.2014.02.009

Oliveira Panão, M. J. N. (2021). Lessons learnt from using energy poverty expenditure-based indicators in a mild winter climate. *Energy and Buildings*, *242*, 110936. https://doi.org/10.1016/j.enbuild.2021.110936

Ooms, J. (2023). *writexl: Export Data Frames to Excel 'xlsx' Format.* R package version 1.4.2. CRAN. https://CRAN.R-project.org/package=writexl

Osberg, L., & Sharpe, A. (2005). How should we measure the "economic" aspects of well-being? *Review of Income and Wealth*, *51*(2), 311–336. https://doi.org/10.1111/j.1475-4991.2005.00156.x

Pachauri, S., Mueller, A., Kemmler, A., & Spreng, D. (2004). On Measuring Energy Poverty in Indian Households. *World Development*, *32*(12), 2083–2104. https://doi.org/10.1016/j.worlddev.2004.08.005

Parker, D.S., Fairey, P.W., & Hendron, R.H. (2011). Updated Miscellaneous Electricity Loads and Appliance Energy Usage Profiles for Use in Home Energy Ratings, the Building America Benchmark Procedures and Related Calculations. Revised.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal, 10* (1), 439-446. https://doi.org/10.32614/RJ-2018-009

Pillai, A., Reaños, M. A., & Curtis, J. (2023). Keep out the cold: An analysis of potential gaps in fuel poverty policies in Ireland. *Energy Research & Social Science*, *98*, 103012. https://doi.org/10.1016/j.erss.2023.103012

Posit team. (2022). *RStudio: Integrated Development Environment for R.* Posit Software, PBC, Boston, MA. http://www.posit.co/

Povea, E. (2023). *Lab 3* [Source code]. ECN425 Solving Economic and Social Problems with Big Data, Norwegian School of Economics, Bergen, Norway.

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

Pye, S., Dobbins, A., Baffert, C., Brajković, J., Grgurev, I., De Miglio, R., & Deane, P. (2015). Energy poverty and vulnerable consumers in the energy sector across the EU: analysis of policies and measures. *INSIGHT_E.* Retrieved March 12, 2023, from https://energy.ec.europa.eu/system/files/2015-07/INSIGHT_E_Energy%2520Poverty-Main%2520Report_0.pdf

R Core Team. (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Recalde, M., Peralta, A., Oliveras, L., Tirado-Herrero, S., Borrell, C., Palència, L., Gotsens, M., Artazcoz, L., & Marí-Dell'Olmo, M. (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. https://doi.org/10.1016/j.enpol.2019.07.005

Renewable energy statistics. Eurostat. (2023b, January). Retrieved April 21, 2023, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Renewable_energy_statistics

Salari, M., & Javid, R. J. (2017). Modeling household energy expenditure in the United States. *Renewable and Sustainable Energy Reviews*, *69*, 822–832. https://doi.org/10.1016/j.rser.2016.11.183

Schulte, I., & Heindl, P. (2017). Price and income elasticities of residential energy demand in Germany. *Energy Policy*, *102*, 512–528. https://doi.org/10.1016/j.enpol.2016.12.055

Silva, S., Soares, I., & Pinho, C. (2017). Electricity demand response to price changes: The Portuguese case taking into account income differences. *Energy Economics*, *65*, 335–342. https://doi.org/10.1016/j.eneco.2017.05.018

Simcock, N., Thomson, H., Petrova, S., & Bouzarovski, S. (Eds.). (2017). *Energy poverty and vulnerability: a global perspective*. Routledge.

Slowikowski, K. (2023). *ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'.* R package version 0.9.3. CRAN. https://CRAN.R-project.org/package=ggrepel

Snell, C., Bevan, M., & Thomson, H. (2015). Justice, fuel poverty and disabled people in England. *Energy Research & Social Science*, *10*, 123–132. https://doi.org/10.1016/j.erss.2015.07.012

South, A. (2011). rworldmap: A New R package for Mapping Global Data. *The R Journal, 3*(1), 35-43. http://journal.r-project.org/archive/2011-1/RJournal_2011-1_South.pdf

Sovacool, B. K. (2012). The political economy of energy poverty: A review of key challenges. *Energy for Sustainable Development*, *16*(3), 272–282. https://doi.org/10.1016/j.esd.2012.05.006

Sun, C., & Ouyang, X. (2016). Price and expenditure elasticities of residential energy demand during urbanization: An empirical analysis based on the household-level survey data in China. *Energy Policy*, *88*, 56–63. https://doi.org/10.1016/j.enpol.2015.10.012

Sweden. IEA. (n.d.). Retrieved February 18, 2023, from https://www.iea.org/countries/sweden

The impact of demographic change in Europe. European Commission. (2023, January 17). Retrieved March 24, 2023, from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/new-push-european-democracy/impact-demographic-change-europe_en

Thomson, H., & Snell, C. (2013). Quantifying the prevalence of fuel poverty across the European Union. *Energy Policy*, *52*, 563–572. https://doi.org/10.1016/j.enpol.2012.10.009

Thomson, H., Bouzarovski, S., & Snell, C. (2017). Rethinking the measurement of energy poverty in Europe: A critical analysis of indicators and data. *Indoor and Built Environment*, *26*(7), 879–901. https://doi.org/10.1177/1420326x17699260

Thomson, H., Snell, C., & Liddell, C. (2016). Fuel poverty in the European Union: a concept in need of definition? *People, Place and Policy*, *10*(1), 5–24. https://doi.org/10.3351/ppp.0010.0001.0002

Turai, E., Schmatzberger, S., Broer, R., Gerőházi, É., Somogyi, E., Szemző, H., Slesarenok, S., Alceva, L., Adzieva, V., Klekovska, M., Kupšys, K., & Kybartienė, E. (2021, May 27). *Overview report on the Energy Poverty Concept*. ComAct – Community Tailored Actions for Energy Poverty Mitigation in the CEE and CIS region. Retrieved March 20, 2023, from https://www.bpie.eu/wp-content/uploads/2021/05/ComAct-D1.1_Overview-report-on-the-energy-poverty-concept_Final-version_UPDATED-1.pdf

Understanding the socio-economic divide in Europe. OECD. (2017, January 26). Retrieved March 24, 2023, from http://oe.cd/cope-divide-europe-2017

von Platten, J. (2022). Energy poverty in Sweden: Using flexibility capital to describe household vulnerability to rising energy prices. *Energy Research & Social Science*, *91*, 102746. https://doi.org/10.1016/j.erss.2022.102746

WHO Indoor Air Quality Guidelines: Household Fuel Combustion. (2014). World Health Organization.

Wickham, H. (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software, 21*(12), 1-20. http://www.jstatsoft.org/v21/i12/

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., Hester, J., Chang, W., & Bryan, J. (2022). d*evtools: Tools to Make Developing R Packages Easier.* R package version 2.4.5. CRAN. https://CRAN.R-project.org/package=devtools

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation.* R package version 1.1.0. CRAN. https://CRAN.R-project.org/package=dplyr

Wickham, H., Miller, E., & Smith, D. (2023). *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files.* R package version 2.5.2. CRAN. https://CRAN.R-project.org/package=haven

Wickham H, & Bryan J (2023). *readxl: Read Excel Files.* R package version 1.4.2. CRAN. https://CRAN.R-project.org/package=readxl

Yutani, H. (2023). *ggsflabel: Labels for 'sf' with 'ggplot2'.* R package version 0.0.1. CRAN. https://yutannihilation.github.io/ggsflabel/

Zeileis, A., Grothendieck, G. (2005). zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software, 14*(6), 1-27. https://doi.org/10.18637/jss.v014.i06

Appendix

Table A1: *List of R packages utilized in the data analysis*

R package	Authors
cluster	Maechler et al. (2022)
corrr	Kuhn et al. (2022)
data.table	Dowle & Srinivasan (2023)
devtools	Wickham et al. (2022)
dplyr	Wickham, François et al. (2023)
factoextra	Kassambara & Mundt (2020)
ggplot2	Wickham (2016)
ggrepel	Slowikowski (2023)
ggsflabel	Yutani (2023)
haven	Wickham, Miller, et al. (2023)
janitor	Firke (2023)
lubridate	Grolemund & Wickham (2011)
readxl	Wickham & Bryan J (2023)
reshape2	Wickham (2007)
rnaturalearth	Massicotte & South (2023)
rworldmap	South (2011)
sf	Pebesma (2018)
softImpute	Hastie & Mazumder (2021)
stats	R Core Team (2022)
tidyverse	Wickham et al. (2019)
writexl	Ooms (2023)
zoo	Zeileis & Grothendieck (2005)

Table A2: *PCA Vector Loadings*

Energy Poverty Indicator	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
Arrears on utility bills – No disaggregation – Country average	0.294158	-0.18505	0.256791	-0.35681	0.21425	-0.18578	0.262694	-0.01951	-0.58269	-0.08596	0.30127	0.112051	-0.21173	-0.20274
At Risk of Poverty or Social Exclusion	0.294132	-0.13461	0.276449	-0.31158	-0.02175	0.520911	0.273628	-0.13768	0.205859	0.363371	-0.38856	-0.11127	-0.12238	0.05881
Dwellings in populated areas – Dwellings in intermediately populated areas	-0.22509	-0.01034	-0.21392	-0.27419	0.804335	0.155743	-0.33303	-0.02886	0.083032	0.151643	0.042215	0.062949	-0.0825	0.053753
Energy expenses by income quintile – Energy expenses, income quintile 1	-0.09476	-0.4836	-0.05974	0.278381	-0.12678	-0.1184	-0.14053	-0.2038	0.018148	0.633805	0.310055	-0.12675	-0.16326	-0.20933
Excess winter mortality/deaths	0.343671	0.270039	-0.19208	0.118336	0.08136	0.087332	0.143555	-0.34752	-0.06359	0.279388	0.220164	0.40622	0.549782	0.067811
High share of energy expenditure in income (2M) – No disaggregation – Country average	-0.05263	0.403219	0.437498	0.06263	0.094523	0.159668	-0.16976	-0.28222	0.005809	-0.06514	0.270764	-0.54658	0.167428	-0.30339
Inability to keep home adequately warm – No disaggregation – Country average	0.38876	-0.09519	0.020365	-0.14837	-0.07554	0.159078	-0.16509	0.444017	0.506369	-0.11611	0.510964	0.10155	0.008094	-0.14354
Low absolute energy expenditure (M/2) – No disaggregation – Country average	-0.08755	0.323794	0.501843	-0.01705	-0.15745	-0.09017	-0.41822	0.044245	-0.02662	0.252832	-0.07845	0.538921	-0.2522	0.027131
Number of rooms per person by	-0.04287	0.450659	-0.29325	0.01258	-0.05319	0.154883	0.163286	0.557701	-0.29225	0.414285	-0.00241	-0.1595	-0.10421	-0.22597

ownership status – Total														
Pop. Liv. dwelling equipped with air conditioning	0.279425	0.274433	-0.07206	-0.27584	0.024465	-0.65531	0.028566	-0.07954	0.264124	0.233974	-0.01252	-0.30779	-0.12604	0.304746
Pop. Liv. dwelling equipped with heating facilities	-0.29395	-0.20549	0.122814	-0.50658	-0.25546	-0.0182	-0.1853	0.196726	-0.19131	0.168532	0.077739	-0.11553	0.567787	0.232934
Pop. Liv. dwelling with presence of leak, damp, and rot	0.178411	0.07522	-0.4585	-0.37158	-0.36006	0.117435	-0.42856	-0.34729	-0.13721	-0.13593	-0.11311	-0.00934	-0.15906	-0.3017
Pop. Liv. dwellings comfortably cool in summer time	-0.3693	0.18579	-0.10605	-0.1677	-0.21446	0.257091	0.290143	-0.23049	0.052497	-0.04659	0.497294	0.057664	-0.35326	0.399981
Pop. Liv. dwellings comfortably warm in winter time	-0.39127	0.058614	-0.00178	-0.2892	-0.01491	-0.23593	0.37974	-0.09472	0.369221	0.027085	-0.07081	0.231494	0.116414	-0.58443

Table A3: *PCA Score Vectors*

Abbreviation	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
AT	-2.04065	0.46722	0.377279	0.014997	0.003434	-0.04459	0.343561	0.128433	0.230104	-0.78567	-0.38745	0.177159	0.257144	0.11933
BE	-1.70247	0.399186	-1.77031	-0.68595	1.471478	0.920152	-0.23308	0.531479	-0.26869	0.455954	-0.02375	-0.05249	-0.12908	0.223849
BG	4.66741	-2.8597	1.219623	-0.10728	1.120685	0.802766	-0.38292	0.99001	-0.49551	0.32135	-0.05208	0.087268	0.467375	-0.08108
CY	3.734148	2.027294	-2.38831	-2.17949	-1.38604	-1.36973	-0.92713	-0.12534	0.500744	0.323175	0.050898	0.01314	0.142344	0.046968
CZ	-2.31056	-1.96056	-0.90194	1.438921	0.016055	-1.11381	-0.14633	0.203885	0.174796	0.408806	0.435977	-0.10933	0.120415	0.001415
DE	-2.11389	0.740253	0.330494	-0.66472	0.591652	0.720199	-0.12057	0.055038	0.389949	0.20567	-0.28744	0.118368	-0.22512	-0.19462
DK	-1.62448	0.633797	-0.47976	0.077593	-1.23745	-0.08718	0.204869	0.172934	-0.42892	-0.3372	0.234627	-0.19975	-0.23988	-0.03312
EE	-0.64668	-0.00923	1.404375	0.888883	-1.1647	0.399116	-0.10361	-0.26872	-0.45036	0.866709	-0.22446	-0.39904	0.094377	-0.0333
EL	3.752826	-0.73114	1.905958	-1.86365	0.940428	-0.97723	0.545222	0.132737	-0.56018	-0.4217	0.397327	-0.44783	-0.34588	0.195634
ES	2.211454	1.272867	-0.53739	-0.63568	-0.23734	0.339497	0.771856	0.030621	0.529339	0.117666	-0.16595	0.318303	-0.2112	-0.11581
FI	-1.90782	2.374815	2.449898	-0.16952	0.777323	-0.89423	-1.05918	0.884719	-0.35939	0.067145	-0.43378	0.272989	-0.23375	-0.18495
FR	-0.49002	1.044804	0.219411	-0.13569	-1.15872	0.162867	-0.39775	0.388331	-0.49908	-0.56547	-0.25401	0.592621	0.051127	0.204117
HR	0.131624	-1.7721	0.295817	-0.26257	0.68747	-0.46529	1.122402	-0.49128	0.158714	-0.57809	-0.29893	0.078277	-0.03442	-0.13229
HU	-0.11213	-2.23049	-1.27128	0.134934	-0.22253	-0.45238	-0.67584	-0.21397	-0.84442	0.428385	0.152032	0.146797	-0.24163	-0.10845
IE	-0.95283	1.667126	-0.33125	-0.37973	-0.74063	1.073578	0.799185	-0.20153	-0.97024	0.061294	0.704545	0.240978	0.183071	0.147501
IT	0.361835	-0.06566	-0.1378	-1.03873	0.769876	0.359371	-0.72272	-0.90169	0.718088	0.261622	-0.41967	-0.08631	0.177601	0.144089
LT	0.945211	-1.12366	1.034895	0.139921	-1.65337	0.709374	0.0301	1.517103	1.037683	-0.25784	0.508604	-0.08389	-0.00805	-0.13852
LU	-1.78878	-0.16007	-1.57798	-0.45943	0.013427	0.544813	0.782254	0.170528	0.128674	0.65442	-0.17377	-0.00069	-0.18821	-0.02204
LV	0.643111	-2.27801	0.647673	0.446527	-1.60937	0.135383	0.007807	-0.60861	-0.27695	-0.15933	-0.90415	-0.57087	-0.08229	0.019906
MT	3.491855	3.979162	0.024844	2.820414	0.525583	-0.98224	1.242986	0.091253	0.018233	0.458678	-0.18058	-0.0056	0.091034	0.005414
NL	-1.96213	-0.11005	-2.18013	-0.35489	0.589869	0.044716	0.564883	0.802175	0.084066	-0.41056	-0.47942	-0.44206	0.212921	0.056236
PL	-1.04493	-0.76781	1.441594	1.356116	-0.08922	-0.58269	-0.68254	-0.41691	0.151299	-0.13924	-0.09031	0.397004	0.059818	0.222222
PT	3.623786	0.572784	-2.09461	2.355207	0.670941	1.342535	-1.12943	-0.53412	0.091811	-0.79307	0.214647	-0.08715	-0.26933	-0.0435
RO	0.66962	-1.18935	2.045818	-0.6489	-0.09174	0.726426	0.598324	-1.14057	0.468794	0.351284	0.246379	0.472889	-0.05323	0.000592
SE	-2.47143	3.066402	2.133758	-0.25984	0.523347	0.327334	-0.5687	-0.52553	0.369761	-0.13018	0.649013	-0.68005	0.16774	-0.03236
SI	-1.12615	-0.15714	-1.34545	-0.71561	0.358069	-0.60355	0.035434	-0.76589	-0.55695	-0.47049	0.274088	0.122558	0.331905	-0.40814
SK	-1.93794	-2.83074	-0.51523	0.88818	0.531473	-1.0352	0.10092	0.094897	0.65865	0.066671	0.507597	0.126714	-0.09481	0.140908