Kaunas University of Technology

Faculty of Mathematics and Natural Sciences

Green Energy Poverty Index: the Case of European Union

Master's Final Degree Project

Radvilė Bankauskaitė

Project author

Prof. Dr. Evaldas Vaičiukynas
Supervisor

Assoc. Prof. Meda Andrijauskienė

Supervisor

Kaunas, 2024

Kaunas University of Technology

Faculty of Mathematics and Natural Sciences

Green Energy Poverty Index: the Case of European Union

Master's Final Degree Project

Business Big Data Analytics (6213AX001)

Radvilė Bankauskaitė

Project author

Prof. Dr. Evaldas Vaičiukynas
Supervisor

Assoc. Prof. Dr. Meda Andrijauskienė
Supervisor

Prof. Dr. Mayur Pal
Reviewer

Doc. Dr. Ramina Vaitiekūnienė
Reviewer

Kaunas, 2024

Kaunas University of Technology

Faculty of Mathematics and Natural Sciences

Radvilė Bankauskaitė

Green Energy Poverty Index: the Case of European Union

Declaration of Academic Integrity

I confirm the following:

1. I have prepared the final degree project independently and honestly without any violations of the copyrights or other rights of others, following the provisions of the Law on Copyrights and Related Rights of the Republic of Lithuania, the Regulations on the Management and Transfer of Intellectual Property of Kaunas University of Technology (hereinafter – University) and the ethical requirements stipulated by the Code of Academic Ethics of the University;

2. All the data and research results provided in the final degree project are correct and obtained legally; none of the parts of this project are plagiarised from any printed or electronic sources; all the quotations and references provided in the text of the final degree project are indicated in the list of references;

3. I have not paid anyone any monetary funds for the final degree project or the parts thereof unless required by the law;

4. I understand that in the case of any discovery of the fact of dishonesty or violation of any rights of others, the academic penalties will be imposed on me under the procedure applied at the University; I will be expelled from the University and my final degree project can be submitted to the Office of the Ombudsperson for Academic Ethics and Procedures in the examination of a possible violation of academic ethics.

Radvilė Bankauskaitė

Confirmed electronically

Bankauskaitė, Radvilė. Green Energy Poverty Index: the Case of European Union. Master's Final Degree Project / supervisor prof. dr. Evaldas Vaičiukynas, assoc. prof. dr. Meda Andrijauskienė; Faculty of Mathematics and Natural Sciences, Kaunas University of Technology.

Study field and area (study field group): Applied Mathematics (Mathematical Sciences).

Keywords: Energy poverty, principal component analysis, data envelopment analysis.

Kaunas, 2024, 76p.

Summary

With current conflicts, upcoming recession, and green transition there are some disruptions in the European energy market. This brings a lot of uncertainty and may negatively affect individuals, especially those from the most socioeconomically vulnerable groups. Hence, understanding, measuring and tackling energy poverty is becoming more and more important. While a lot of complex indicators already exist, they often fail to take into account the changes in energy market, specifically the shift from traditional energy resources towards renewable ones. The goal of this study is to create a green energy poverty index that considers renewable energy and green transition in the assessment of energy poverty. Several already existing energy poverty indices were analysed to better understand what are the most suitable indicators and methods for computing the new index. The literature analysis also revealed that green transition can reduce energy poverty if the energy transition actions are implemented with a focus on improving affordability and efficiency of energy. However, if the green transition is premature, it may exacerbate energy poverty. Two methods are selected. Robust Principal Component Analysis is used to create a green energy poverty index that can be easily computed when the values of the selected indicators are available. Data envelopment analysis is used to evaluate efficiency of the countries. Data envelopment analysis evaluates whether the country is progressing in green transition with a focus on inclusivity and energy poverty reduction and provide the target values for energy poverty with the current use of renewable energy resources if they are used efficiency with such focus. The analysis covers 27 EU Member States. Five variables were selected for the index – inability to keep adequately warm, arrears on utility bills, use of renewables for electricity, use of renewables for heating and cooling, share of energy from renewable resources. Robust PCA is used to derive the weights for each of the variables in the index. After that, the green energy poverty index value for each country is calculated. DEA is used to determine which countries are most efficient, meaning they have lowest values of energy poverty indicators associated with the amount of renewable energy resources it has. The newly computed indices are validated using correlation analysis with relevant indicators. The study provides a closer look to the relation between green transition and energy poverty. The index constructed can be used to inform policy decision-making process.

Bankauskaitė, Radvilė. Žaliasis energetinio skurdo indeksas: Europos Sąjungos atvejis. Magistro baigiamasis projektas / vadovai doc. dr. Evaldas Vaičiukynas, asoc. prof. dr. Meda Andrijauskienė; Kauno technologijos universitetas, Matematikos ir gamtos mokslų fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Taikomoji matematika (Matematikos mokslai).

Reikšminiai žodžiai: Energetinis skurdas, principinių komponenčių analizė, duomenų gaubtinė analizė

Kaunas, 2024, 76 p.

Santrauka

Energetikos rinkoje šiuo metu yra daug neužtikrintumo ir pokyčių dėl dabartinių konfliktų, recesijos ir didelio susirūpinimo ekologija. Tai galimai neigiamai veikia Europos gyventojus, o ypač pažeidžiamiausias socialines grupes. Dėl šios priežasties energetinis skurdas ir jo matavimo metodai tampa vis svarbesni. Nors jau egzistuoja daug indeksų, kurie siekia išmatuoti energetinį skurdą, juose dažnai neatsižvelgiama į besikeičiančią energijos rinką, ypač į perėjimą prie didesnio atsinaujinančių energijos išteklių naudojimo. Šio tyrimo tikslas – sukurti žaliąjį energetinio skurdo indeksą, kuriame atsižvelgiama į atsinaujinančios energijos naudojimą vertinant energetinį nepriteklių. Siekiant geriau suprasti, kaip turėtų būti sudarytas naujas žaliosios energijos indeksas, buvo išanalizuoti skirtingi jau esami energetinio skurdo indeksai. Taip pat buvo išnagrinėti įvairūs energetinio skurdo apibrėžimai, kurie yra pateikti akademinėje literatūroje ar politikos dokumentuose. Naujajam indeksui sukurti naudojama atspari pagrindinių komponenčių analizė. Duomenų gaubtinė analizė naudojama šalių efektyvumui įvertinti ir nustatyti energetinio skurdo rodiklių tikslus, kuriuos galima pasiekti su naudpjamu atsinaujinančių energijos šaltinių kiekiu. Analizė apima 27 ES valstybes nares. Indeksui pasirinkti penki kintamieji – gyventojų dalis, kuri neišgali palaikyti deramos šilumos namuose; gyventojų dalis, kuri turi skolų už komunalines paslauga; atsinaujinančių energijos šaltinių kiekis naudojamas elektros energijai; atsinaujinančių energijos šaltinių kiekis naudojamas šildymui ir vėsinimui; atsinaujinančių išteklių naudojimas energijai kaip dalis nuo visos sunaudojamos energijos. Atspari pagrindinių komponenčių analizė naudojamas kiekvieno indekso kintamojo svoriams nustatyti. Po to apskaičiuojama žaliosios energijos skurdo indekso reikšmė kiekvienai šaliai. Indeksas patvirtinamas naudojant koreliaciją su kintamaisiais, kurie yra asocijuojami su energetiniu skurdu. Duomenų gaubtinė analizė naudojama siekiant nustatyti, kurios šalys yra efektyviausios, o tai reiškia, kad jos turi mažiausias energetinio skurdo rodiklių vertes, vertinant turimą atsinaujinančių energijos išteklių kiekį. Rezultatai koreliuoja su naujai sukurtu indeksu ir taip pat yra patvirtinti koreliacine analize. Tyrime atidžiau pažvelgta į perėjimo prie ekologiškos ekonomikos ir energijos nepritekliaus ryšį. Sukurtas indeksas gali būti naudojamas politinių sprendimų priėmimo procesui informuoti.

Table of contents

List of figures

List of tables

List of abbreviations

CEPI – Compound Energy Poverty Indicator

CRS/CCR model – constant returns to scale DEA model

DEA – data envelopment analysis

EPI – Environmental performance index

EU – European Union

GEPI – Green Energy Poverty Index

HEPI – Household Energy Poverty Index

IAEA – International Atomic Energy Agency

IEA – International Energy Agency

MAD – median absolute deviation

MCD estimator – minimum covariance determinants estimator

MEPI – Multidimensional energy poverty index

PC – principal component

PCA – principal component analysis

VRS/BCC – variable returns to scale DEA model

Introduction

Nowadays energy poverty is an important subject among policy makers and academics around the globe, including Europe. Energy prices have been fluctuating in the face of recent crises such as COVID-19 pandemic and the war in Ukraine. Energy market in Europe is also gradually moving towards more sustainable energy sources in the light of climate change and green transition, towards which the EU is moving.

This brings a lot of uncertainty, especially for the most socioeconomically vulnerable groups in European society. Hence, energy poverty and the methods to measure it are becoming more and more important. However, there is no consensus on the definition of energy poverty and its most suitable indicator. While a lot of indicators and indices already exist, only a few of the existing energy poverty indices take into account sustainability and consider renewable energy sources. This study is aiming to fill this gap by presenting two new indices to evaluate energy poverty. These new indices consider the transition towards greener energy market when assessing energy poverty.

The study proposes two alternative indices – a Green Energy Poverty Index based on principal component analysis, and efficiency score based on data envelopment analysis. They both examine energy poverty and its connection to green transition. Green Energy Poverty Index computed using PCA evaluates how energy poor and lagging behind in green transition the country is. Efficiency index examines whether the progress towards green transition is just, inclusive, and contributes to reduction of energy poverty.

The study focuses on 27 Member States of the European Union (EU). The dataset used in the study covers the time period between 2010 and 2022.

The goal of this study is the energy poverty index that considers renewable energy and green transition in the assessment of energy poverty.

The tasks of the study are as follows:

1. To highlight the existing gap in the literature on energy poverty and green transition nexus;
2. To identify different indicators related to energy consumption and sustainability that are most suitable for the energy poverty index in the EU that taking into account green transition;
3. To construct and index using principal component analysis;
4. To construct and index using data envelopment analysis;
5. To validate the newly created indices;
6. To map energy poverty in the EU using the newly constructed index;
7. To highlight potential issues to research further on the topic.

1. Literature review

This part of the thesis presents relevant scientific literature analysing energy poverty and its measurements. This literature review presents the research done so far and highlights the need for more comprehensive approach in assessing energy poverty in the light of green transition. The literature review is structured as follows. It firstly presents the main aspects of existing definitions of energy poverty and analyses energy poverty in the light of green transition. Later on, energy poverty in the EU is presented and analysed. Finally, the literature review presents different methods to measure energy poverty, including those that already take into account transition towards renewable energy. While some indices already exist, the literature review highlights the need for a more comprehensive approach specifically in the EU.

1.1. Energy poverty

1.1.1. Definition of energy poverty

Energy poverty is a well-researched topic. In the current context, energy poverty is a complex phenomenon that generally relates to an inability of a household to meet their energy needs. The term is coined around by academics and policy makers to highlight the existing challenges of some individuals and households to meet their energy needs. However, there is still a lot of research questions left unanswered and scientific community is not in consensus on the definition of energy poverty or its most suitable measurement (Widuto, 2023).

Varying definitions of 'energy poverty' can be found in the EU policy documents. They all are related to inability of individuals or households to meet their basic energy needs. According to European Commission's proposal for a directive on energy efficiency in 2021, energy poverty is defined as '*a household's lack of access to essential energy services that underpin a decent standard of living and health, including warmth, cooling lighting, and energy to power appliances, in the relevant national context, existing social policy and other relevant policies*'(European Commission, 2021). According to other recent policy documents, energy poverty occurs when a household is not able to access essential energy services (European Commission, 2020), and is linked to a combination of high expenditure on energy, low income, and low energy efficiency (Widuto, 2023).

All countries in Europe, except for Sweden, consider energy poverty to be an important issue in both political and academic debates. While definitions of energy poverty among the EU member states vary, there are some common elements. In general, energy services are understood in a broader context, including heating services, cooling services, lighting, and ability to use different appliances necessary for everyday life. Most of the member states also consider energy services as basic need for all citizens. They also recognise the close links between energy poverty, social exclusion and health issues. Finally, the EU member states generally identify three main causes of energy poverty – high energy prices, low energy efficiency, and low household income (Sokołowski et al., 2019).

Scientific literature also does not offer a single definition of energy poverty. It is analysed taking into account various dimensions, such as social factors, health, geographical factors, economic factors, and political initiatives (Siksnelyte-Butkiene et al., 2021). Considering the energy poverty definitions in scientific literature more broadly, they can be divided into two main categories (Streimikiene et al., 2020):

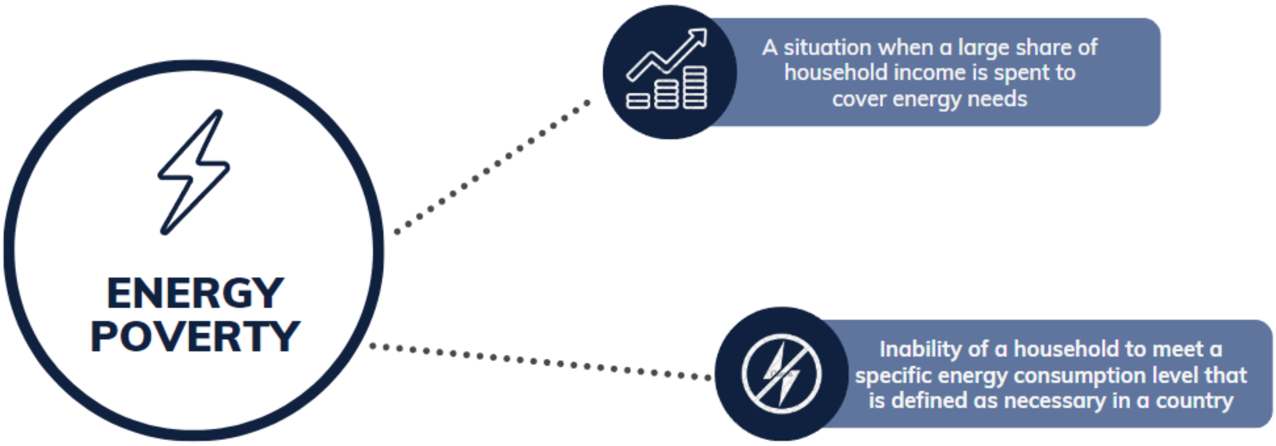

Figure 1. Broad definitions of energy poverty

Until last decade, energy poverty was a rather neglected topic in the scientific community and energy policy. Energy policy discussions and publications on energy and technology used to focus on modern technologies and innovations, while everyday technologies that impact most of the people were often overlooked. In the first decade of the 21st century there were only 8 per cent of papers in the top scientific energy journals that at least to some extent discussed energy poverty. This challenge was also not adequately addressed in the policy agenda (Sovacool, 2014). For example, UN Millenium Development Goals (*United Nations Millennium Development Goals*, n.d.), presented in 2000, did not tackle energy poverty or any related challenges, even though energy poverty and energy deprivation are seen as some of the main challenges to development. The situation changed in the second decade of the 21st century. In 2011, the International Energy Agency and some of the UN institutions included a discussion on energy poverty and its challenges in the World Energy Outlook (IEA, 2012). In 2012, the UN declared that year 'International Year of Sustainable Energy for All'. This increased focus on energy poverty and related questions in policy discussions resulted in increased interest in the topic in the scientific community. Since then, numerous organisations and scholars focused on better defining and understanding energy poverty, including its causes, potential measurements, and different ways to tackle it (Sovacool, 2014).

In the EU context, in 2018 European Parliament and Council's regulation on the Governance of the Energy Union and Climate Action highlighted the urgent need to better define and understand the energy poverty (European Parliament & The Council of the European Union, 2018). Hence, while there is no EU-wide consensus on the definition of the energy poverty, several recent policy documents touch upon this issue. Electricity Directive mentions this problem and highlights the need to calculate and measure which households can be considered as vulnerable costumers. It proposes that such calculation could include several criteria, namely '*'low income, high expenditure of disposable income on energy, and poor energy efficiency*' (European Commission, 2019). Commission Recommendation on energy poverty presents specific recommendations for the Member States on how to tackle the causes of the energy poverty. They include development of a systematic approach to the liberalisation of energy markets, use of specific indicators to better understand and assess energy poverty, adoption of complex and comprehensive approach to energy poverty and introduction of social policy measures and energy efficiency improvements (European Commission, 2020). Energy Efficiency Directive, issued by the European Commission in 2021, not only outlines the measures for increasing energy efficiency, but also highlights the incentives to address energy poverty. It stresses the need to address energy poverty when striving for energy efficiency and suggests that increasing energy efficiency is one of the most effective measures to tackle energy poverty (European Commission, 2021).

While energy poverty generally relates to the inability to meet energy needs, the understanding of this problems differs depending on a country, region or social group in question. Cultural norms and consumer preferences may significantly impact energy needs and the way these needs are satisfied (Nussbaumer et al., 2012). The issue is also understood very differently in developing and developed countries. In developing countries, energy poverty relates to both affordability and structural energy coverage, while in developed countries the energy affordability is considered as the most important challenge and energy infrastructure often has secondary importance (Ruiz-Rivas et al., 2022). Indeed, a study focusing on relationship between low-carbon energy and energy poverty in China and Germany presents energy poverty index (EPI) and indicates that while in China energy poverty is concerned with both affordability and access, in Germany energy affordability is the main concern (Bonatz et al., 2019). The understanding of energy poverty also differs in colder and warmer regions. In colder regions, energy poverty is closely linked to the necessary heating and household's ability to ensure it, while in warmer regions it is less relevant (Ruiz-Rivas et al., 2022).

The definition of energy poverty may also be understood in general or context-specific terms. Generally, it could be understood as deprivation of energy services. This general definition usually considers internationalisation of energy markers and policies aiming to mitigate climate change. Country- or context-specific definition of energy poverty allows to consider various specificities and nuances of energy poverty in a given context. The use of both of these definitions may have its advantages and disadvantages. General definition allows for comparison between countries but may fail to capture energy poverty issue in specific national contexts. Context-specific definition limits overgeneralisation of the energy poverty risk and aids national policy makers in effectively addressing energy poverty in their specific national context, but may complicate international comparison and generalisation (Sokołowski et al., 2019).

1.1.2. Energy poverty in the light of green transition

Energy market around the world is slowly changing due to the changes in the availability of different resources and the climate change concerns. Traditional energy sources such as coal are being used less, while the use of renewable energy resources is becoming more prominent. Already in 2010, United National Development Programme in their Human Development Report urged national governments to assess their energy pricing. They argued that the price should take into account the environmental costs of using conventional energy resources, for example, fossil fuels. The report argued that this pricing would be an incentive for the consumers to change their behaviour and improve their energy efficiency (UNDP, 2010). Consequently, these changes in energy consumption habits and trends may also affect energy poverty. However, the focus on causal link between clean energy and energy poverty is rather recent in academia. This section presents an overview of the main positive and negative effects of green transition on energy poverty. It highlights that while transition towards renewable energy may reduce energy poverty, especially in the long-term, premature transition, which is not accompanied by well-thought-out policies, may bring some negative consequences for the most vulnerable groups in society.

Influence of transition towards more renewable energy on energy poverty can be assessed through analysing the effect of both energy poverty and use of renewable energy on other economic and social aspects. For example, a study on the connection between energy poverty and development outcomes assesses the effect of both energy poverty and renewable energy on development outcomes, such as life expectancy, education, income, and similar. Thee authors of the study found that the transition towards renewable energy to some extent compensates the negative effects of energy poverty on development outcomes (Adom et al., 2021). A study focusing on a causal relationship between transition towards low-carbon energy and energy poverty in China finds significant bidirectional causality between use of low-carbon energy transition and reduction of energy poverty. The study highlights that transition towards low-carbon energy indirectly reduces energy poverty through increased availability of energy services, improved cleanness of energy consumption, strengthened energy management systems, and improved efficiency and affordability of energy (Dong et al., 2021).

While some existing studies point to positive effect of wider use of renewable energy on energy poverty, green transition also may have some negative effects on energy poverty. This negative effect often depends on the policies that are implemented to support green transition. The studies show that the policy makers often have to compromise between energy security, sustainability and affordability. The environmental concerns and climate change motivated the policy makers to focus on a rapid transition towards renewable energy, which, in some cases, may compromise energy security, especially if the demand for energy is increasing (Hussain et al., 2023). Not everyone can afford more efficient unconventional energy sources and reduction of availability of conventional energy resources, such as fossil fuels, or increase in their price may heighten the vulnerabilities of the poor. A study on climate policy in Europe finds that poorly designed climate policy measures and premature green transition in energy sector, especially coupled with increasing energy prices, may exacerbate energy poverty trap (Belaïd, 2022). A study analysing effects of Hungarian National Energy and Climate Plan on energy poverty illustrates how the inability of the Plan to address the drivers of energy poverty and vulnerabilities of the groups heavily relying on solid fuels may result in an increased energy poverty risk (Bajomi et al., 2021). A study on low-carbon development in Germany and China also finds that transition from traditional energy sources towards renewable energy sources may result in higher energy costs, which are especially burdensome for the most vulnerable groups (Bonatz et al., 2019). The already mentioned study on connection between energy poverty and development outcomes also finds that even though most negative effects of the transition towards renewable energy on development indicators are likely to disappear in the long-term, increased energy poverty may be long-lasting (Adom et al., 2021).

The existing literature shows that transition towards renewable energy may worsen or improve the situation of energy poverty depending on a specific economic and political context. Hence, this connection should be assessed in a case-by-case manner. This study aims to investigate this connection in the EU Member States by presenting green energy poverty indices that evaluate energy poverty while taking into account the shift towards renewable energy consumption.

1.2. Energy poverty in the EU

European Pillar of Social Rights Action Plan considers energy to be one of the essential services to which each person in the EU should be entitled to (European Commission, n.d.-b). Still, energy poverty is a serious challenge. According to the available data, the percentage of Europeans who could not afford keeping their home adequately warm has been gradually decreasing in past few years, after the increase from 2010 to 2012 (Eurostat, n.d.). The change in the inability to keep home adequately warm is seen in the Figure 2 below. However, this does not necessarily mean that the situation is improving. The share of individuals not being able to keep their home adequately warm in 2018, 7.6 % , translated into 34 million EU residents not being able to keep their home adequately warm (European Commission, 2020). In 2020, this number increased to 36 million (Widuto, 2023).

The situation is also significantly different across the Member States. The Figure 2 below presents the percentage of citizens who were not able to keep their home adequately warm in the EU. The graph also presents the best and worst performing countries – Finland and Bulgaria, respectively. In 2021, Finland had the lowest percentage of this indicator and Bulgaria had the highest. In 2010, the difference between these two countries was above 65 percentage points and in 2021 it was above 20 percentage points (Eurostat, n.d.). This indicates that the energy poverty is varying significantly across the Member States and in some countries the situation is even more dire than the EU average indicates.

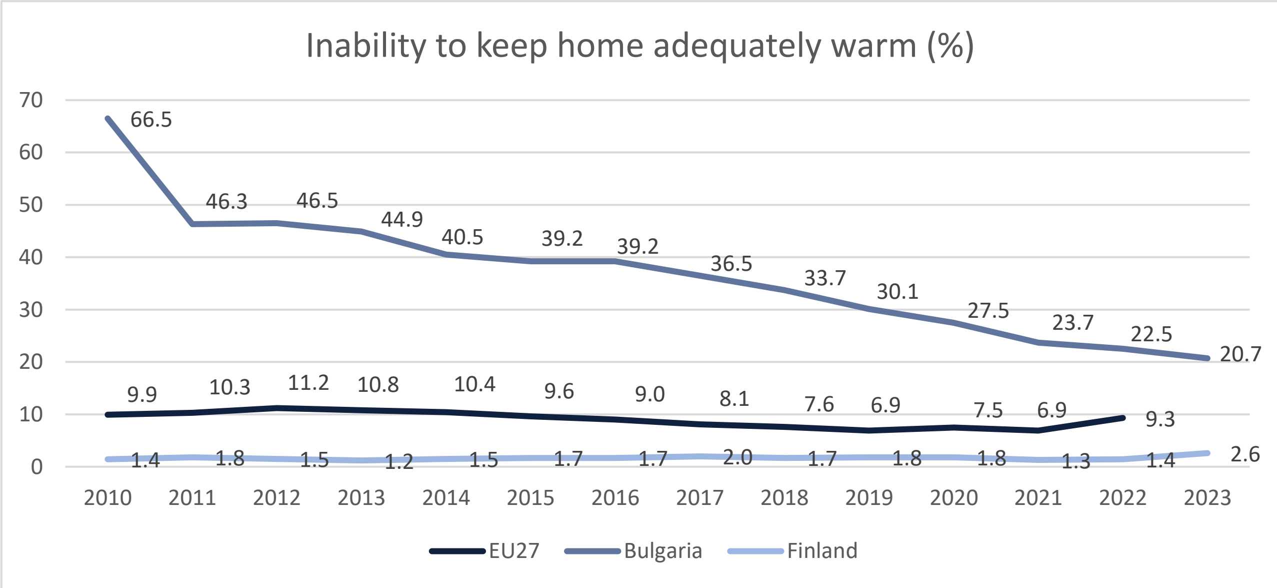

Figure 2. Inability to keep home adequately warm in the EU, Finland and Bulgaria (Eurostat, n.d.)

The energy poverty in the recent years has been fuelled by increasing energy prices and uncertainty in the energy supply. In the past years, the European Union and its Member States have experienced several challenges that contributed to increasing energy prices over the continent. Firstly, the climate change is causing more extreme changes in weather, with the summers in the EU Member States getting hotter and winters getting colder. This increases the demand for energy throughout the year as the EU citizens are trying to mitigate the extreme weather conditions. The change in weather also negatively affects the ability of the EU Member States to produce its own energy from renewable sources. For example, due to the summer heatwaves, most of the hydropower sources across Europe often become usable in that season. Additionally, energy prices are increasing as a result of the recent COVID-19 pandemic recovery and the consequences of Russia‘s war in Ukraine (Rao, 2022). The gas supplies by Russia were deliberately reduced, which caused the energy prices, specifically gas prices, to increase significantly (European Council, 2023). These challenges result in volatile energy prices, which, together with various socioeconomic factors associated with poverty, cause wider energy poverty (Widuto, 2023).

In the light of these challenges, several policy measures have been implemented to tackle energy poverty and its causes. The figure below presents the most important EU policy initiatives that aim to reduce energy poverty and ensure the citizens' entitlement to essential energy services.

Figure 3. EU policy initiatives that aim to tackle energy poverty (European Commission, n.d.-a)

To addition to the policies presented above, emergency regulation was introduced in the late 2022. The emergency regulation aimed to mitigate the negative effects of energy crisis on EU citizens and businesses and in that way also indirectly affected energy poverty. The regulation applied from December 2022 to March 2023 and complemented existing EU initiatives aiming to secure EU's energy supplies. The regulation also aimed to create favourable conditions for the Member States to support the most vulnerable individuals and companies in the light of skyrocketing energy prices. The regulation presented three measures (European Council, 2023):

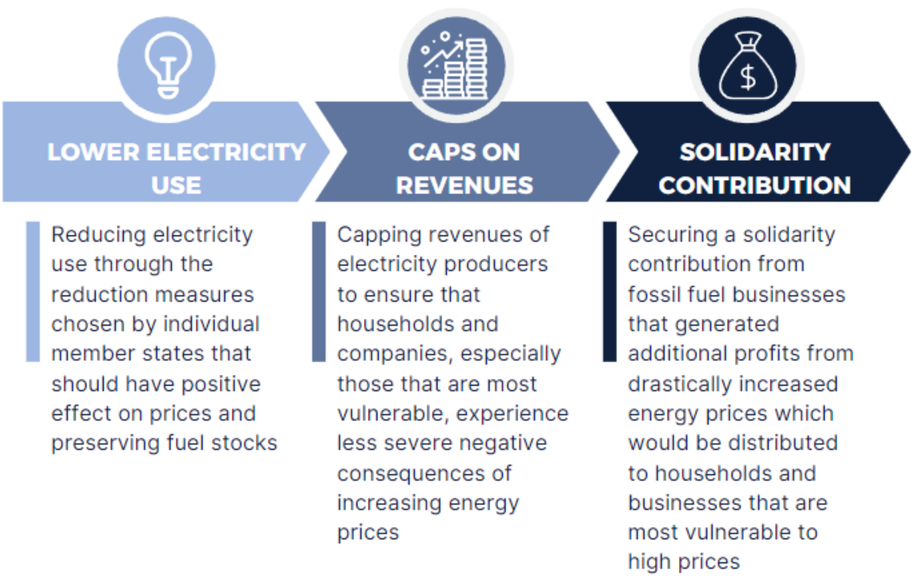

Figure 4. Measures included in emergency regulation (European Council, 2023)

Together with the mentioned emergency regulation, the EU institutions also introduced REPowerEU plan, which is a long-term strategy of the EU for tackling the energy market disruptions caused by Russia's war in Ukraine. The plan outlines the actions that will be taken to achieve Europe's independency from Russian gas and oil before 2030. While the plan does not directly target energy poverty, it affects it through the changes it introduced to the European energy market. The plan focuses on three main action areas (European Commission, 2022):

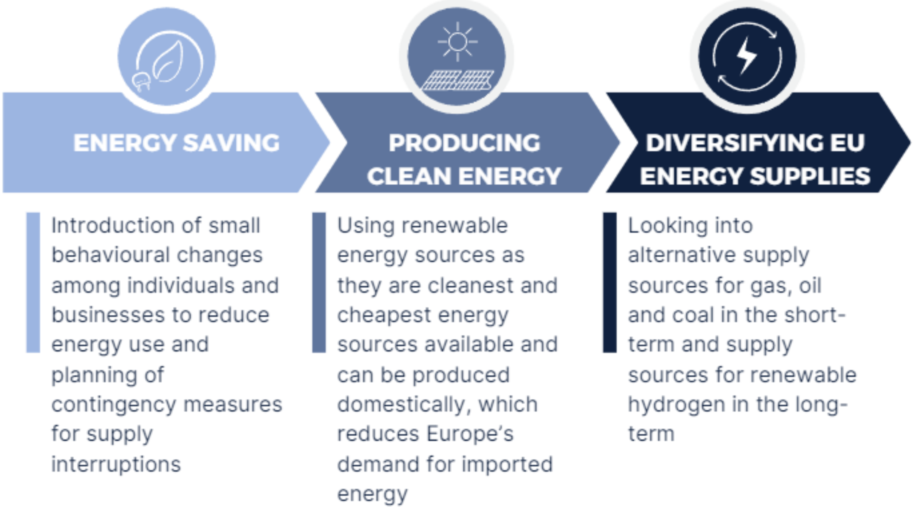

Figure 5. Action areas of REPowerEU action plan (European Commission, 2022)

The REPower EU plan relies on a significant scaling-up of renewable energy sources as it is hoped that clean energy transition will increase EU's energy independence and contribute to the reduction of energy prices over time. The plan also heavily relies on introduction of energy efficiency measures by both businesses and individual consumers, which should be identified and promoted by the EU together with international organisations, Member States and local authorities (European Commission, 2022).

Both the REPower EU plan and the emergency regulation illustrate a wider long-term EU strategy to move towards more sustainable and clean energy market. This strategy is a part of the envisioned European green transition, which main objectives include reduced green gas emissions through the promotion of circular economy, more efficient use of resources, including energy resources, and use of renewable energy sources. The envisioned green transition can be seen as the main strategy in addressing the current energy crisis in the EU (Hoyer & Küüsvek, 2022). This transition is envisioned through the European Green Deal, the action plan which addresses climate and environmental challenges. The Green Deal targets the whole EU economy with a goal to facilitate EU's transition towards a modern, resource-efficient and competitive economy. Its main objectives include "*no net emissions of greenhouse gases by 2050, economic growth decoupled from resource use, and no person and no place left behind*" (European Commission, 2019).

While the green transition and, consequently, the Green Deal targets the whole economy and does not directly tackle energy poverty, it transforms the environment in which it develops and influences habits and choices of individuals. The Green Deal action plan increased EU's ambitions for greenhouse gas emission reduction targets for 2030 and 2050. According to the increased targets the EU should reduce its greenhouse gas emissions by at least 50% by 2030 and reach climate neutrality by 2050. The Green Deal also focuses on increasing the supply of clean, affordable and secure energy, which means moving towards renewable energy sources (European Commission, 2019). As part of the Green Deal and REPower EU, in March 2023 the EU's renewable energy target for 2030 was increased to 42.5% with an ambition to surpass it and reach 45% (European Commission, 2022). The Green Deal also foresees drastic changes in construction, transport, agriculture, and other industries, with the focus on circularity, sustainability, resource and energy efficiency. It also introduces a shift in national budgets and financing priorities, shifting away from harmful subsidies towards green priorities (European Commission, 2019). This may result in the changes of the prices for different energy sources and may seriously affect the energy consumption habits of some groups (UNDP, 2010). Moreover, as already mentioned, if the green transition is introduced to energy market prematurely and is coupled with increasing energy prices, which have been observed in the EU in the light of COVID-19 and Russia's invasion to Ukraine, such situation may contribute to energy poverty trap (Belaïd, 2022).

This overview of energy poverty situation in the EU and its relevant changes highlights the need to better understand energy poverty and how it is affected by the green transition. As the presented policies, together with the national policies in the EU Member States, are causing a shift in a way energy is produced and consumed, the daily habits of energy consumers (both individuals and business) are also likely to change. The European Green Deal envisions just transition leaving no one behind. It aims to ensure that its objectives are achieved in the least burdensome and most effective way, following the oath to do no harm. It also foresees specific actions to address energy poverty and its risk during the green transition (European Commission, 2019). However, changes in energy supply and consumption trends moving towards more renewable energy sources may affect various groups in society differently, especially considering the most vulnerable groups. Hence, it is important to find a tool to evaluate energy poverty in the context of these changes. The effect of a shift towards renewable energy on the most vulnerable households should be closely monitored and this shift should be considered when assessing energy poverty. This study aims to provide a measurement of energy poverty that considers this change in energy market. This could help the policy makers better understand how the most vulnerable groups in society are affected by the changes in energy supply and ensure that the '*leave no one behind*' aim of the European Green Deal is achieved.

1.3. Existing measurements for energy poverty

Energy poverty is a complex challenge. Hence, it may be hard to assess it using a single indicator and it is usually measured by a complex index that takes into account several different indicators. However, similarly as with the definition of energy poverty, one widely accepted indicator or index does not exist. Several academics focusing on the issue present their own indices that better capture specific situation that individuals and households face. The indices focus on different data, different methodologies and different scopes. Detailed overview of the already computed energy indices is presented in the literature review synthesis matrix, available in Appendix 1. The Figure 6 below presents a summary of this literature review synthesis, with an overview of the differences of the existing energy poverty indices.

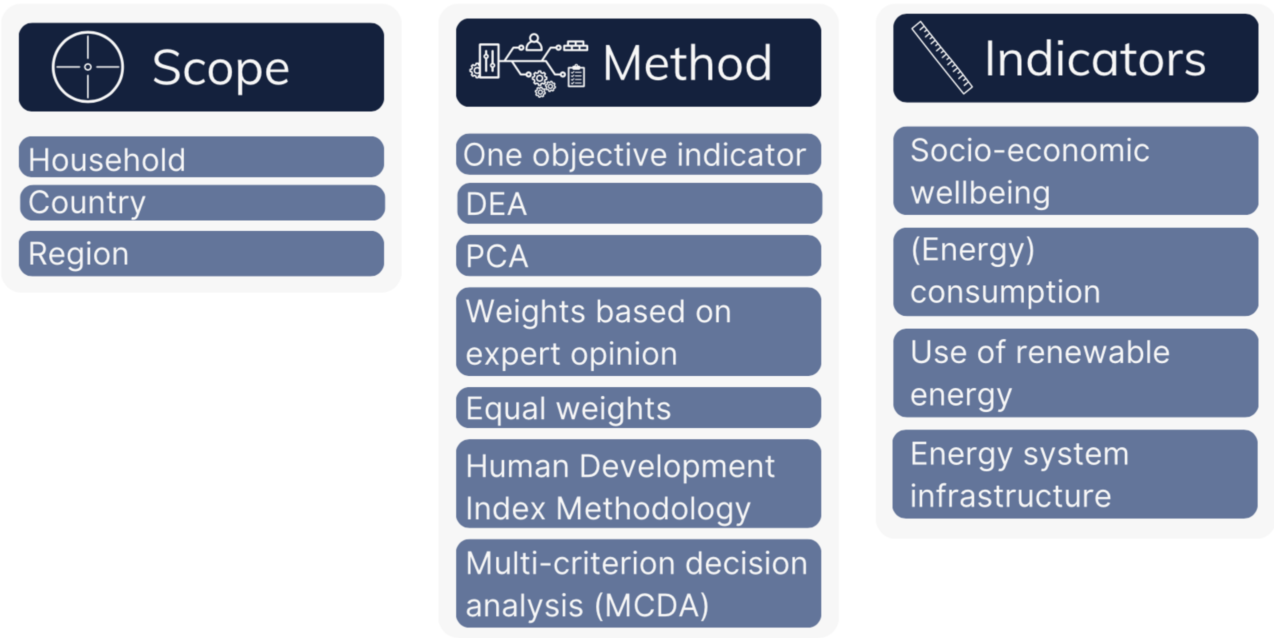

Figure 6. Overview of different scopes, methods and indicators of existing energy poverty measurements

As it can be seen, existing indices focus on different scopes, use different methods and indicators, depending on their aim and objective. Different methodological approaches are used to construct the indices, including relying on expert opinion to assign weights to different measurements of the index or using mathematical approaches for assigning weights. This section discusses the articles presented in the literature review synthesis matrix and presents different approaches to measuring energy poverty, existing academic studies that are considering use of renewable energy resources when measuring energy poverty, and the main mathematical tools to create socio-economic indices, such as energy poverty.

1.3.1. Existing indices for energy poverty

In 1991, B. Boardman proposed the 10% indicator for energy poverty. According to this measurement, energy poverty is defined as a situation when a household spends more than 10% of its income to cover energy costs (Boardman, 1991). This indicator is still often used. However, its several shortcomings have also been highlighted. For example, this measurement does not account for different expenditure on energy due to different climate. Hence, the researchers have further improved indicators for energy poverty and three main types of measurements can be identified (Kahouli & Okushima, 2021). These types of measurements include:

- **Objective factual measures** that are based on observable and measurable criteria and are connected to consumption theory. Such measures consider the amount a household spends to meet their energy needs. Such measures include already mentioned 10% indicator (Kahouli & Okushima, 2021).

- **Subjective self-reported measures** that rely on personal opinions and interpretations. Such measures are constructed using self-reported answers of the households to the questions about meeting energy needs, such as warmth of the home, and difficulties to cover energy costs that are asked in different social survey (Kahouli & Okushima, 2021).

- **Composite indices** including several individual indicators that can be seen as a compromise between one-dimensional indicators and the complex nature of energy poverty that needs to be accounted for (Okushima, 2017). Such measures tackle the main shortcomings of the one-dimensional indicators, as already mentioned 10% indicator. At the same time, they aim to present the complex nature of energy poverty through a set of sub-indicators condensed into an easily understandable indicator (Kahouli & Okushima, 2021).

The objective factual measures, as mentioned 10% indicator, are rather simple and, consequently, widely used. However, they are often criticised. For example, an author of a study assessing the fit of energy poverty measures highlights several methodological challenges when assessing energy poverty. These challenges include accounting for diversity of domestic energy services, considering housing costs, equivalising energy expenditures and household incomes, and issues regarding representativeness of data, among others. These issues cannot be addressed when using only one metric. Hence, such measurements fail to capture a complex nature of energy poverty (Herrero, 2017). Moreover, while the single indicators that are objective factual measures can be easily interpreted, they present only a narrow picture of energy poverty (Nussbaumer et al., 2012). Another study that analyses the usefulness of energy-access-consumption matrix as energy poverty indicator highlights the arbitrariness of single-metric measures (Pachauri & Spreng, 2011). There is currently no consensus on what are the minimum energy needs of a household and the process which could help derive the definition of those needs (Culver, 2017).

Subjective self-reported measures, which are based on the household responses to various survey questions are also seen controversially (Herrero, 2017). Some academics argue that such indicators effectively illustrate a perceived energy poverty and may provide valuable explicit insights into energy poverty that objective factual measures often lack (Rademaekers et al., 2016). At the same time, the reliability and accuracy of self-reported indicators are often questioned as these indicators are based on subjective understanding of comfort and satisfaction of energy needs (Herrero, 2017). The more disadvantaged individuals often tend to have lower expectations regarding their comfort and needs. They may also be ashamed of their situation and not report their struggles due to social bias (Eurostat, 2009, p. 200). Hence, the self-reported measures are subjected to 'denial of reality bias', meaning that the disadvantaged individuals may deny not being able to meet their energy needs. The self-reporting may also be affected by cultural differences in understanding the level of comfort needed and the acceptable temperature at home. These cultural differences may result in a situation where the same value of a measurement means different outcomes in different countries. Hence, such measurements may not be reliable in cross-country analyses (Nussbaumer et al., 2012). Moreover, while subjective self-reported measures may present an overview of the perceptions towards the situation, they often lack the depth needed for a more complex investigation of energy poverty (Herrero, 2017). Hence, it is often proposed that the subjective self-reporting measures are used complementarily to more objective factual measures (Rademaekers et al., 2016).

Composite indices, which aim to illustrate a complex nature of energy poverty, but also provide a simplicity of having one condensed value, are, at least partly, addressing the shortcomings of both objective factual and subjective self-reported measures. As highlighted in existing studies on multidimensional energy poverty indices, these indices aim to close or at least reduce a gap between theoretical and operational definitions of energy poverty. Multidimensional energy poverty indices capture different dimensions of energy poverty and can take into account complex and nuanced theoretical aspects of energy poverty (Pelz et al., 2018). Such indices combine several indicators that tend to be linked to energy poverty. For example, a study on home-heating energy poverty risk in Ireland proposes a composite index that utilises 10 indicators related to home heating and energy poverty. It accounts for specific characteristics of buildings, households and heating systems. Such index helps better understand the situation specific households face and in that way illustrates the complexity of energy poverty (Kelly et al., 2020).

A paper analysing energy poverty measurements across the EU proposes a Compound Energy Poverty Indicator (CEPI), which combines several factors to energy poverty. CEPI assesses three self-assessed variables referring to living conditions, namely inadequate warmth, inadequate coolness, and inadequate level of light. The indicator additionally considers arrears and leaks. The inadequate level of light is given a weight of 0.1, while other indicators are given the weight of 0.2 each. This indicator, according to the authors, offers a more complex overview of energy poverty on national levels across the EU. However, the authors also acknowledge that CEPI does not account for some of less visible aspects of energy poverty and does not consider climate differences across regions (Maxim et al., 2016).

S. Okushima in a paper published in 2017 presented a multidimensional energy poverty index (MEPI) that aims to evaluate energy poverty in developed countries from a multidimensional angle. It considers three factors related to energy poverty that are most relevant in developed countries. These factors include household income, energy efficiency of the housing, and energy costs. The author argues that while energy poverty is generally divided into availability and affordability of energy, availability challenges are more relevant in developing countries and affordability is the main concern in developed countries. Hence, the proposed index focuses on the challenges that are most relevant in developed countries and are related to energy affordability (Okushima, 2017).

S. Gupta, E. Gupta and G. K. Sarangi presents a Household Energy Poverty Index (HEPI) that was derived using principal component analysis (PCA). The index focuses on India and is constructed following a multi-dimensional energy poverty framework. The authors use 15 indicators related to different dimensions of energy consumption and economic welfare of households. The indicators belong to five groups – living standards, affordability of a household, indoor air pollution, use of clean fuels, and geographical accessibility. The authors use PCA to derive the weights for the used variables. The results are used to compute a value of HEPI. The index is then used to group households into four groups depending on their energy poverty and the geographical distribution of these groups are used to better understand energy poverty in the country (Gupta et al., 2020). This complex index with the weights derived from PCA can be more useful than individual indicators as it presents a value that takes into account data variance and different aspects of complex issue of energy poverty. A study focusing on energy poverty in Sri Lanka also uses principal component analysis to derive the weights of indicators for Multidimensional Energy Poverty Index. The authors of the study consider indicators on access to electricity and ownership of electrical devices (fridge, computer, and others) when assessing energy poverty. The authors found that the variables that mostly contribute to the energy poverty index value are use of modern cooking fuel, ownership of computer, and ownership of a fridge (Jayasinghe et al., 2021). This highlights the benefits of using PCA for combined energy poverty indices – the method allows to understand which indicators are most important for energy poverty without the potential biases that may come from the weights derived from expert opinion.

A paper exploring a nexus between energy poverty and energy efficiency propose using data envelopment analysis (DEA) and entropy method to capture a complex nature of energy poverty and its effect on energy efficiency. In the paper, the authors explore indicators of energy poverty, energy efficiency, and socio-economic wellbeing of a country. They find that energy poverty negatively affects socio-economic indicators in the country. The link between energy poverty and energy efficiency is also presented. It is highlighted that inefficient energy use contributes to higher energy poverty. The authors also present potential policy changes to address energy poverty through energy efficiency. The suggested policy changes include holistic policy framework that focuses on supporting low-income households and policies to reduce energy consumption of residential sector to improve energy efficiency (Li et al., 2021). Use of DEA and different dimensions of indicators allows the authors of the study to not only measure energy poverty, but also provide much needed context and potential steps forward, which is not always possible when using only a single indicator.

S.Kahouli and S. Okushima in their study explores a direct measurement approach that bases identification of energy poverty of a household directly on an actual use of domestic energy services (Kahouli & Okushima, 2021). This composite index addresses the shortcomings of the previously presented Compound Energy Poverty Indicator. However, it may less clearly reflect the subjective perception towards energy poverty. The measure focuses on two dimensions – use of energy services and income level. It also addresses the differences in energy needs across different regions. The measurement defines a specific level of energy use as a poverty line and considers the households that fall behind the defined poverty line as energy poor. In order to define a specific level of energy use, the measurement classifies the households to types according to factors relevant to the energy use and structural management of energy. These factors include different socio-demographic factors, such as whether a household includes the elderly, dwelling, separating the households living in apartments and in detached houses, and climate. The measurement is superior to a uniform threshold as it better assesses specific energy needs of households living under different conditions. For example, the poverty threshold is higher in colder regions where households tend to spend more on energy use. It also considers the structural management of energy use, specifically the energy sources used, as some energy sources may have poverty reduction effect (Kahouli & Okushima, 2021).

While the composite indices tackled some of the problems arising from using single metric objective factual measures or subjective self-reported measures, these indices also have some shortcomings. A study exploring European energy poverty metrics analyses the measurement of energy poverty from the metrics and measurement research side and energy poverty research (socio-political and economic research) side. The authors of the study note that from the metrics side any set of indicators pose a risk of silencing some of the aspects of the measure that are hard to quantify and amplifying the others. From energy poverty research side, there is an urgent need for an indicator that sufficiently contextualises the complex energy use issues. The authors propose an analytical framework to assess energy poverty through five dimensions. These dimensions include historical trajectories, which can be considered key characteristic of metrology shaping energy poverty metrics, data flattering and contextualised identification, that enact metrology, and new representation and policy uptake, that influences the reconfiguration of the metrics. The authors of the study stress the problematic act of measurement, noting that energy poverty, both its definition and representation, depends on what is being measured (Sareen et al., 2020).

Other studies also find that multidimensional nature of energy poverty poses a challenge of choosing the indicators that both capture the complexity of the issue and allow straightforward application of the index (Sokołowski et al., 2019). Most of the indices face a challenge of combining different dimensions and levels of different indicators into a single index to measure energy poverty (Pelz et al., 2018). One of the main questions when constructing a specific index is the weights of different indicators used. The rational weighting strategy that is theoretically sound is very difficult to construct. Most of the strategies tend to be arbitrary and value driven. Hence, weighing and aggregation of an index should be non-compensatory, and the weights should be used taking into account the meaning of coefficients (Nussbaumer et al., 2012).

Other studies highlight the lack of comparability of the values of different multi-dimensional energy poverty indices. As already mentioned, energy poverty is a complex issue of a multi-dimensional nature. However, most of the recent energy poverty measurements are too complex. Hence, they cannot be operationalised at the global level and cannot always be applied to different national contexts. This highlights the need for selection of key energy poverty indicators and dimensions that then could be adapted to specific national contexts (Pelz et al., 2018). Moreover, the study highlights that while a specific index may suit one region, it may not properly measure and highlight the most pressing aspects of energy poverty in another region (Nussbaumer et al., 2012).

Another important challenge with composite energy poverty indices is availability of sufficient quality data. To produce a good quality internationally consistent energy poverty measurement framework, data should be collected regularly, following the same data collection framework (Pachauri & Spreng, 2011). However, regular and complete data on energy poverty is rarely available. The framework for the data collection also often differs from region to region or even from country to country. Reasons for that is that energy poverty is a culturally sensitive and private issue. It is also socially and temporally dynamic. Hence, proper evaluation of this concept requires specific data, which is often lacking. This limits the ability of researchers and countries to monitor energy poverty (Thomson et al., 2017). This challenge related to data could be addressed through the implementation of specific surveys dedicated to assessing energy poverty. However, such surveys are often expensive and difficult to carry out on a large scale (Pelz et al., 2018). Moreover, as the energy poverty is a private and sensitive issue, if survey collected self-reported data, the results may be biased and hardly comparable on a cross-country level. This highlights the importance of critically assessing the available data when constructing an energy poverty index.

1.3.2. Construction of energy poverty index taking into account green transition

The overview of different poverty energy measurements points to the importance of complexity and specific context when assessing energy poverty. Hence, with the shift towards renewable energy in the EU, it is evident that this transition should be better considered when evaluating energy poverty. This section presents the existing efforts to take into account green transition when measuring energy poverty. It illustrates the progress made so far considering perspectives of the EU and other regions.

Energy indicators can be seen as closely connected to other sustainability indicators such as health, education, safe food and drink, as energy supply is necessary to ensure their availability (Pachauri & Spreng, 2011). Hence, several energy-based sustainability indicators have been developed. The Atomic Energy Agency, together with other UN agencies, presented guidelines for energy indicators for sustainable development. These guidelines recommend a list of energy indicators in social, economic and environmental dimensions that could help measure sustainable development. The mentioned three dimensions are then divided into different themes, such as equity and health, among others, and at least one energy indicator is presented for each theme. The energy indicators for each theme are chosen from a supply side and allows comparison between countries (IAEA et al., 2005). Energy-related sustainability indicators became especially relevant when the concerns about negative impact of excessive energy use became more prominent in global politics scene. Consequently, academics calculated an upper and lower limit of energy use per capita that is necessary to meet a decent standard of living. These limits were very close to each other, around 2000 W per capita per year. It was hoped that this indicator would provide some guidance for individual countries and regions in their development process (Pachauri & Spreng, 2011).

More recent studies on energy poverty tend to focus more and more on both availability of physical energy infrastructure and accessibility of different energy sources (Siksnelyte-Butkiene et al., 2021). This often includes a focus on a shift towards clean energy or a focus on climate change mitigation policies through addressing energy consumption. For example, a study on the EU climate change mitigation measures addressing energy poverty present a framework linking energy poverty and climate change mitigation measures. It highlights the importance of well-targeted climate change mitigation policy packages. The authors of a study also stress the need to link social, regulatory, economics, financial and behavioural barriers that the households face when implementing climate change mitigation measures, for example shifting towards clean energy (Streimikiene et al., 2020).

A study on energy poverty evaluation in the EU countries presents an approach to evaluating energy poverty that addresses energy demand, shift towards clean energy, and energy justice. The study proposes energy poverty framework that is based on three pillars. Firstly, it considers demand of and access to energy, focusing on society. Second pillar focuses on administration that should ensure accessibility to different energy sources and alignment between energy market and import policies of a country. Third pillar is concerned with sustainability and the level of available renewable energy sources. Such framework considers accessibility, affordability and sustainability. The study applied this framework to evaluate energy poverty in the EU Member States by weighing the criteria with Threshold-based Attribute Ratio Analysis method and evaluating the countries using Measurement Alternatives and Ranking according to Compromise Solution methodology. However, the authors of the study acknowledge the ambiguity of the proposed framework that stems from the complexity of energy poverty and related problems (Hasheminasab et al., 2023).

A study addressing the nexus between energy poverty and energy insecurity with the role of various environmental concerns, including climate change, presents a set of indicators that can help evaluate energy poverty in the light of environmental concerns. The authors of the study take into account various indicators that are often associated with energy poverty, including social, economic, energy and environmental performance indicators. They use data envelopment analysis (DEA) to define the nexus between energy poverty and environmental performance. The derived composite indicator allows to measure energy, economic, social, and environmental performance index (EPI). The new index assesses whether countries are able to tackle energy poverty question without compromising environment and contributing to climate change (Ehsanullah et al., 2021).

While the studies presenting specific energy poverty indices and measurements accounting for a shift towards more renewable energy in the EU are rather scarce, there are several studies focusing on other countries, especially China, aiming to integrate green transition to poverty measurement. For example, study analysing the role of low-carbon energy transition in mitigating energy poverty in China presents a composite energy poverty index that takes into account natural gas consumption.

The authors consider energy consumption cleanliness, energy service availability, energy affordability and efficiency, and energy management completeness. The study allows to better understand the effect of energy cleanliness and use of different energy sources on energy poverty. However, the study also highlights a few shortcomings of the index. It notes that the index does not consider nonlinear link between energy poverty and natural gas consumption (Dong et al., 2021) (Dong, et al., 2021). Another study focusing on China employs 28 indicators that can be grouped in three categories – energy service availability, residential energy efficiency and affordability, and cleanliness of energy consumption and generation. These indicators also capture policy factors relevant in China. It allows the authors to capture the complexity of energy poverty and also include green transition to their considerations. The authors find a clear link between transition towards clean energy and improvement in energy poverty. They also highlight the importance of other factors that contribute to the observed improvement (Liang & Asuka, 2022).

This overview of the existing studies that connect energy poverty with green transition and, more specifically, with the use of the renewable energy resources, highlights some of the potential gaps in the existing academic research on the nexus between energy poverty and green transition. While several indices already exist, they are not always easily interpretable. Most of the indices consider a high number of indicators that touch upon a large number of issues related to both energy poverty and green transition. This study aims to present an index that is more concrete and easier to interpret.

1.4. Main takeaways from the literature review

The literature review highlights the lack of clear definition of energy poverty and lack of consensus on how it should be measured. It is clear that the issue is multifaceted and its measurement will always have to depend on the context in which energy poverty is analysed. As presented, energy poverty is often defined as a situation where a large share of income is spent on energy needs or as an inability to meet a specific energy consumption level. Energy poverty is measured using objective factual measures, subjective self-reported measures, and composite indices that often combine the different objective and subjective indicators. Literature review revealed that all of these methods have their specific advantages and disadvantages. Therefore, the selection of measurement depends on circumstances in which energy poverty is assessed and specific objectives of that assessment.

Considering mathematical methods that are used for composite energy poverty indices, literature review presented two methods – Principal Component Analysis and Data Envelopment Analysis. Both of these methods have slightly different objectives. PCA aims to simplify the complex reality and present different variables in one indicator. The derived index can present complex dataset in one value and can be used for comparison between countries. DEA, on the other hand, evaluates each observation individually and instead of aiming to just present the situation at hand it also provides recommendations for improvement.

The literature review also presented the current situation of energy market and energy poverty in the EU. It is clear that the energy market in the EU is rapidly changing. Energy prices are fluctuating due to different problems, including COVID-19 and Russia’s war in Ukraine. These challenges also bring more uncertainty to the energy market. At the same time, the EU is focusing on green transition that promotes increased use of more sustainable energy sources. However, this transition is rarely considered when assessing energy poverty. Hence, there is a need for a clearly defined energy poverty index that takes into account the progress towards green transition. This study presents such index and provides more information for policy makers on the current situation and potential steps forward.

To sum up, the literature review, presented above, highlights the existing gap in academic research on energy poverty and green transition nexus, which is the first task of this study. The literature review presents the existing dilemmas about the definition of energy index and the ways this problem is measured ad assessed. The literature review also presents the overview of energy poverty in the EU, which points to an urgent need for an index that this study is presenting.

2. Methodology and data

This section presents information about software that is used in the study, describes in more detail statistical approaches used for constructing the indices (PCA and DEA) as well as methods to validate them. Data used in the study is also presented in this section.

2.1. Necessary software

During this study the following programming languages were used: *R* (in *RStudio* environment), *Python* (in *JupyterLab* environment).

RStudio is an integrated development environment that supports *R* and *Python*. For this study, *RStudio* was used for computations using *R*. *RStudio* was selected for using *R* due to the personal preferences of the author of the study. *R* is a programming language that is usually used for statistical analysis, computing and data visualisation. This programming language offers its user a variety of statistical and graphical techniques and can produce high quality plots. *R* language is open source and is available as free software (The R Foundation, n.d.). It is popular in academia, specifically for data science and data analytics. In this study, following *R* libraries were used: *dplyr, arules, rrcov, tidyverse*, *ggplot2*, *ggthemes*, *plyr*, *dear,* and *ggridges*.

JupyterLab is a web-based integrated development environment that supports over 40 programming languages, including *R* and *Python*. *JupyterLab* was selected for using *Python* due to the personal preferences of the author of the study. *Python* is a programming language that has recently become widely used for data analysis tasks. However, it is a general-purpose programming language that can be used for different occasions. *Python* supports object-oriented, functional, structured, procedural and reflective programming paradigms. With a vast number of developed libraries, *Python* can be used for statistical computations, machine learning, and data visualisation, among other purposes. In this study, following *Python* libraries were used: *pandas, numpy, Eurostat, ydata_profiling, sci-kitlearn, matplotlib, scipy, seaborn, folium,* and *json.*

2.2. Construction of the index

This study uses several different methods that are usually applied for the creation of the socioeconomic indices. They will be used to compute green energy poverty index. The tested methods were selected based on the literature review. The author of this study uses statistical approaches that have been used in recent studies exploring different energy poverty and socio-economic status indices and have proved to be suitable for this task. These methods are explained below and include the following:

1. Principal Component Analysis;
2. Data envelopment analysis.

2.2.1. Principal component analysis

2.2.1.1. Usefulness of the method for index construction

In early 2000s Principal Component Analysis (PCA) has become one of the most regularly used methods to construct socio economic indices (Vyas & Kumaranayake, 2006). For example, PCA was used to create a living standards index for World Bank (Gwatkin et al., 2000). This approach was also used to modify UNDP Human Development Index. Previously, Human Development Index used to be constructed by a simple aggregation procedure, which proved to be inappropriate for the index. Hence, PCA was carried out to assign more appropriate weights to the indicators (Lai, 2003).

PCA can be seen as a favourable method to construct socioeconomic index for its relative computational simplicity. This approach allows a researcher to transform a set of variables into a smaller more coherent set of orthogonal factors that account for a large part of the variation among the original data. With PCA a researcher constructing an index can avoid most of the problems that often arise when using traditional methods to compute an index, such as standardisation or aggregation (Krishnan, 2010). These problems include seasonality, recall bias, and nonlinear relationships. Moreover, PCA is computationally easy compared with alternative statistical approaches for index creation. The index derived with PCA also facilitates comparison over time or between countries (Vyas & Kumaranayake, 2006).

While PCA is commonly used for the construction of complex socioeconomic indices, the method also receives some criticism. For example, it is argued that the weights given to each indicators using PCA may fail to capture weighting preferences for each indicators for each individual case, individual, household, or, as in this study, country (Jayasinghe et al., 2021). Moreover, PCA technique is often seen as arbitrary. In most cases when constructing an index using PCA the methodology for choosing a number of components and selection of variables is not well defined. It is also important to note that the issues related to the data used will influence PCA. Hence, they should be considered when creating the index and interpreting its results (Vyas & Kumaranayake, 2006).

2.2.1.2. Presentation of the method

PCA can be described as a technique that transforms a vast number of different variables into a smaller set of principal components, uncorrelated orthogonal factors. These principal components account for variance in the dataset in the values of the original variables. Each principal component can be seen as a linear weighted combination of the initial variables. For example, first and z^{th} principal components for the dataset with variables from X_1 to X_n can be written as follows (Vyas & Kumaranayake, 2006):

$$PC_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n \quad (1)$$

$$PC_z = a_{z1}X_1 + a_{z2}X_2 + \cdots + a_{zn}X_n \quad (2)$$

In the equations a_{zn} stands for the weight for the z^{th} principal component and the n^{th} variable (Vyas & Kumaranayake, 2006).

The weights are allocated so that the principal components would be orthogonal and the first principal component would account for the largest share of the variation in the original variables in the dataset. Consequently, second principal component accounts for the maximum variation in the original variables that is not covered by the first principal component and is not correlated with the first component. Hence, all principal components are completely uncorrelated with each other and cover the variance in the original variables that is not covered by other components, representing different statistical dimensions in the original data (OECD, 2008). The figure below represents how principal components are located in respect to one another.

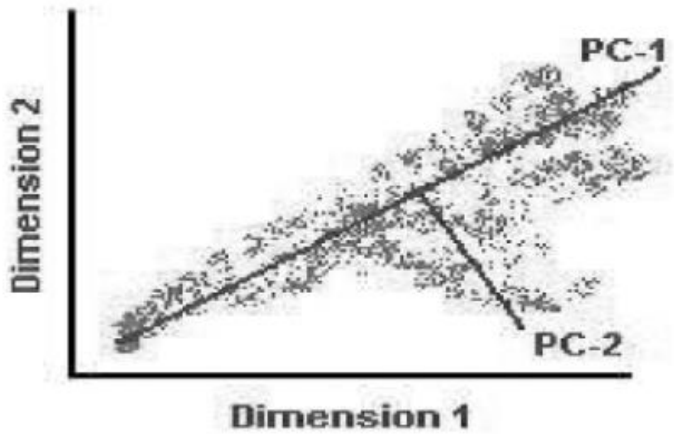

Figure 7. Two sequential dimensions of PCA (Vyas & Kumaranayake, 2006, p. 460)

It is important to keep in mind that the traditional PCA is implemented under several assumptions. Firstly, a number of observations available should be sufficient. However, there is no consensus on how many observations is sufficient. Secondly, it is assumed that the selection of indicators was not biased. Thirdly, it is assumed that the data has no outliers as their presence may influence the interpretations of the analysis. When implementing PCA it is also assumed that the data is interval and linear, and there is multivariate normality. Finally, it is assumed that there are underlying dimensions that clusters of indicators share and the data is correlated (OECD, 2008).

Even if the abovementioned assumptions are not met, PCA could still be appropriate with some modifications. For example, if a dataset used has outliers, robust PCA can be used instead of traditional PCA. Robust PCA differs from traditional PCA by a determinant it uses to decompose the original dataset.

One of the robust PCA methods that is used with datasets with outliers is a robust PCA using Minimum Covariance Determinants estimator (MCD). In traditional PCA, the principal components are computed from the decomposition of covariance matrices. The decomposition depends on the means of original variables and covariance matrix of the dataset. This makes the analysis sensitive to outliers. In the robust PCA with MCD estimator, the estimator depends on a mean and covariance matrix of subset of the dataset with the observations that have covariance matrix with the smallest determinant. This makes the estimator robust to the outliers in the data (Hubert et al., 2018).

The difference between classical and MCD estimator can be best understood through the graphic illustration. Figure 8 below presents the scatter plot of the data set of 59 Italian wines from the paper by Hubert et al. (Hubert et al., 2018). The scatter plot shows proline and malic acid quantity in all wines in the dataset and presents classical tolerance ellipse (in red) and robust tolerance ellipse (in blue).

Classical tolerance ellipse can be described as a set of *p*-dimensional points *x*. Their Mahalanobis distance, which shows how far away a specific point *x* is from the centre of the whole data cloud, taking into account its shape and size, is computed as follows (Hubert et al., 2018):

$$MC(x) = d\big(x, \bar{x}, Cov(X)\big) = \sqrt{(x-\bar{x})' Cov(X)^{-1} (x-\bar{x})} \tag{3}$$

This classical estimator using Mahalanobis distance includes most of the data points into the tolerance ellipse, as it is strongly influenced by contamination and cannot detect outliers in the data (Hubert et al., 2018).

Robust tolerance ellipse based on MCD estimate (in blue in Figure 8 below) is significantly more sensitive to outliers. Robust distances are counted as follows (Hubert et al., 2018):

$$RD(x) = d\left(x, \hat{\mu}_{MDC}, \hat{\Sigma}_{MCD}\right) = \sqrt{(x - \hat{\mu}_{MDC})^t \hat{\Sigma}_{MCD}{}^{-1} (x - \hat{\mu}_{MDC})} \tag{4}$$

Where $\hat{\mu}_{MDC}$ is the MDC estimate of location, and $\hat{\Sigma}_{MCD}$ is the MCD covariance estimate.

The location estimate can be understood as a mean of the data set with h observations that have the covariance matrix with smallest determinant. The covariance estimate is the corresponding covariance matrix of the data set with the observations h multiplied by a consistency factor. This results in the situation where tolerance ellipse is significantly more sensitive to outliers. As it can be seen in the figure below, using this estimator significantly more data points are considered to be outliers (Hubert et al., 2018).

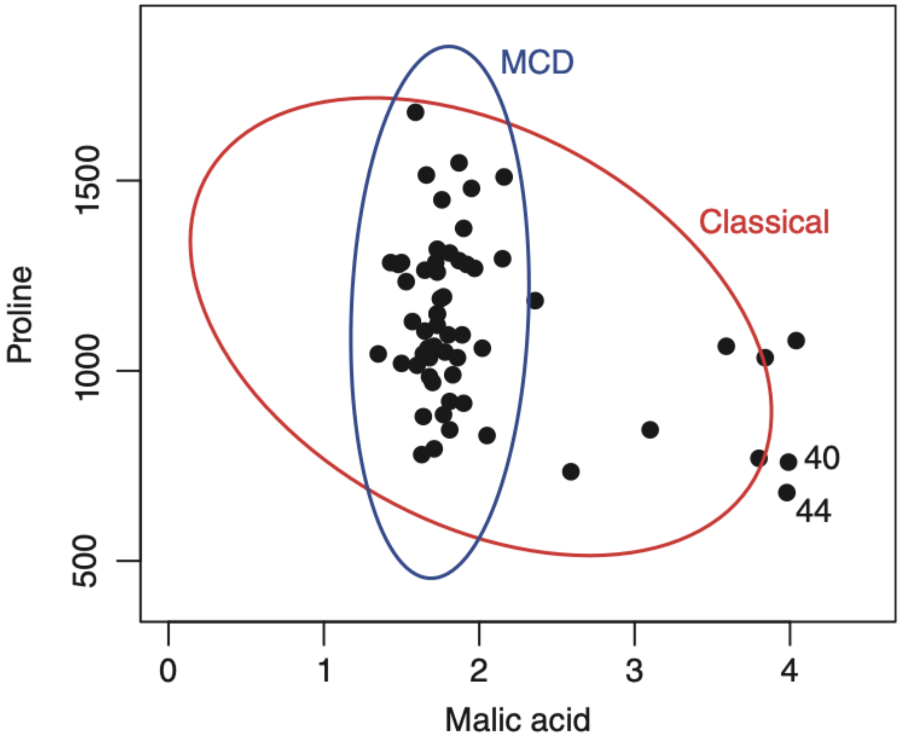

Figure 8. Tolerance ellipses using classical (red) and MCD (blue) estimates (Hubert et al., 2018)

As already mentioned, the MDC estimate of location and robust distances explained above are used in the robust PCA algorithm that was first presented by Rousseeuw & Van Driessen (Rousseeuw & Driessen, 1999). The algorithm first starts with a fixed number of random elemental subsets that are then used to construct their halfsets. After these halfsets are constructed, classical centres and covariance matrixes are computed and robust distances are computed for each point for each halfset using equation (4). After these computations, "best" halfsets are selected and the computations are carried out until the convergence. The MCD estimator is then based on the halfset that had the lowest determinant of the covariance matrix (Nisa et al., 2006). These computations then can be used to get principal components that are expressed by equations (1) and (2).

In this study, to implement the PCA method, these steps are followed:

1. The variables that may help explain and map energy poverty in the EU are identified;
2. Data is analysed to check for outliers;
3. The selected variables are checked for correlation;
4. The selected variables are standardised;
5. The PCA method is implemented;
6. The index is computed and the values for all the EU countries are calculated.

The first step is to select the most suitable indicators for the index. The indicators are selected based on the previously presented literature analysis and data availability. The data selection is discussed in more detail in the section 2.4. In general, data for PCA should not be categorical and all qualitative categorical variables should be coded into binary variables (in this study, categorical variables are not used). Missing values is also an issue for PCA. Hence, if there are a lot of missing values, different methods to impute those values should be considered (Vyas & Kumaranayake, 2006). In this study, there were no missing values in the original dataset.

Second step is analysing whether the data has outliers. Outliers may affect the results of PCA and its interpretation through its influence on correlations (Krishnan, 2010). However, it is important to also have clear criteria which points are considered outliers. In some cases, the data points that seem to be outliers from the first sight may actually include important information. For example, if the datapoints that can be considered as outliers all represent the same country in a dataset covering several countries, taking out these observations may require to take out the whole country out of the analysis. Moreover, as the analysis may be sensitive to non-linear relationships, linearity of data should also be assessed. If variables have non-linear relationship, the correlation coefficients presented may not properly reflect the strength of the relationship between the variables (Krishnan, 2010). After assessing the data, the data is standardised. Standardisation of the data is also a common practice to ensure that one variable does not have unnecessarily large influence on the PCA (OECD, 2008).

Fourth step of the PCA implementation includes assessing the appropriateness of the selected method (PCA) through examining the correlation of the variables. While some researchers do not consider multicollinearity a problem for PCA, it may result in higher standard errors (Krishnan, 2010). Still, the selected original variables have to be correlated for the PCA method to be applicable. If the original variables are not correlated, the analysis will not produce the desired results (OECD, 2008).

After data is assessed, it can be determined whether PCA is an appropriate method of analysis. If some of the assumptions are not met, for example, data includes outliers, robust PCA can be implemented instead of traditional PCA. Once the analysis is carried out, the results can be used to construct the index.

To calculate the weights for the index, it has to be firstly determined which principal components are significant and should be kept in the calculation of the weights. Once it is determined, the weights can be computed. This study uses the same computation method as already presented in the study constructing Household Energy Poverty Index for India using PCA(Gupta et al., 2020). Each variable has a weight in each principal component and each principal component has a weight that is computed dividing its eigenvalue from the sum of the eigenvalues of retained principal components. The formula to calculate the individual weights of each variable is as follows:

$$W_i = \frac{\lambda_1 f_{i1} + \lambda_2 f_{i2} + \lambda_3 f_{i3}}{\lambda_1 + \lambda_2 + \lambda_3} \tag{5}$$

Where λ_n is an eigenvalue for nth principal component, and f_{in} is a weight inside the nth principal component for *i* indicator (Gupta et al., 2020).

The weights then can be standardised to ensure that they are easier to interpret. The values are often scaled to add to 1 (Nardo et al., 2005). However, in this study it was decided to scale the weights from -1 to 1. This decision was taken as the variables used in the index present two different dimensions (energy poverty and progress in green transition). Therefore, having index with both negative and positive weights results in easier interpretability of the index weights.

2.2.2. Data envelopment analysis

2.2.2.1. Usefulness of the method for index construction

Data envelopment analysis (DAE) is another method that has been used to develop composite indices. In recent years DAE has been used to evaluate energy efficiency and energy poverty (Li et al., 2021).However, the use of DAE method in analysing socioeconomic issues is rather new. The method was firstly introduced for industrial applications to evaluate the profitability or efficiency of companies (Golany & Roll, 1989). It is a mathematical approach based on linear programming to determine a set of weights of different variables to maximise efficiency of the selected unit (country, region, company, or other) (Mariano et al., 2015).

DEA can be seen as a favourable approach due to the several approaches to extracting weights. DEA allows to extract weights from the data itself with individual weights that are most suitable for each analysed unit. However, DEA also allows to determine a common set of weights that is most advantageous for all analysed units and could be used for constructing an index (Mariano et al., 2015). The indices constructed using DEA can also be perceived as fairest for all analysed units (countries) as it tends to result in highest composite scores compared to the indices got from other weighting schemes (Nardo et al., 2005).

While DEA approach has its advantages, it also has some shortcomings. For example, if the scores are normalised in different ways, each of them may produce different weighting scheme. Individual weights may also result in a composite equal to 1 for all countries if there are no constraints imposed on weights. On the other hand, if the constraints on the weights are imposed, the analysis may provide no solution for the maximisation problem for some analysed units (countries) or it may have several solutions, making it impossible to determine the optimal set of weights. It is also important to note that the best performing unit (country) is used as a reference in the analysis. Hence, the progress of the best performer over time cannot be tracked by DEA (Nardo et al., 2005).

2.2.2.2. Presentation of the method

DAE can be described as a mathematical procedure based on linear programming. This method uses frontier approach to assess relative performance or efficiency of decision-making units (DMUs) based on fractional programming problem which is converted into a linear programming problem. DMU can be a company, organization or jurisdiction, among others (in this study, it is a country) (Cotte Poveda, 2012). DAE allows to incorporate several variables into a single value without converting those variables into the same measurement unit (Mariano et al., 2015).

Different models of DAE have been developed over the years. The first model was CCR model based on the method of frontier analysis. Other well-known models include BCC model, and window analysis assessing the performance of DMU over selected time period by looking at the DMU as a different entity in each time point (Cotte Poveda, 2012).

CCR model is also often called constant returns to scale (CRS) model. It assumes that the outputs in the model grow proportionally to inputs in the model. BCC model, which is a model of variable returns to scale (VRS), does not require proportionality. Evaluations of CRS/CCR model are based on constant returns to scale that can be represented as mathematical problem (Bowlin, 1998):

$$\text{minimise: } \theta - \varepsilon\left[\sum_{i=1}^{m} s_i^{-} + \sum_{r=1}^{s} s_r^{+}\right] \tag{6}$$

$$\text{subject to: } 0 = \theta x_{io} - \sum_{i=1}^{n} x_{ij}\lambda_j - s_i^{-} \tag{7}$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^{+} \tag{8}$$

Where θ is unrestricted in its sign, and $0 \leq \lambda_j, s_i^-, s_r^+$ for $i = 1, \ldots, m$; $r = 1, \ldots, s$; $j = 1, \ldots, n$

Evaluations of VRS/BCC model are slightly more flexible with variable returns to scale. It can be expressed as (Bowlin, 1998):

$$\text{minimise: } \theta - \varepsilon\left[\sum_{i=1}^{m} s_i^{-} + \sum_{r=1}^{s} s_r^{+}\right] \tag{9}$$

$$\text{subject to: } 0 = \theta x_{io} - \sum_{i=1}^{n} x_{ij}\lambda_j - s_i^{-} \tag{10}$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^{+} \tag{11}$$

$$1 = \sum \lambda_j$$

Where θ is unrestricted in its sign, and $0 \leq \lambda_j, s_i^-, s_r^+$ for $i = 1, \ldots, m$; $r = 1, \ldots, s$; $j = 1, \ldots, n$

The main difference between CCR and BCC models is that for CCR model DMU has to be both scale and technical efficient, while for BCC model DMU only needs to achieve technical efficiency to be considered as efficient (Bowlin, 1998). Both of these models are considered to be radial models as the efficiency index presents either equiproportional increase in outputs or equiproportional decrease in inputs needed to achieve efficiency. For this reason, these models also require selection of the orientation of the model (Mariano et al., 2015).

DEA models can be input- or output- oriented. The orientation of the analysis depends on its goal. The input-oriented analysis focuses on the reduction of inputs. Output-oriented analysis focuses on expanding the output (Cook et al., 2014). For example, if DMU is a public space with costs for public spaces as input and quality of public spaces as output and the aim is to decrease the cost while keeping the quality of public spaces, input-oriented DEA would be most appropriate. If DMU is a country with education quality indicators as inputs and economic development indicators as outputs and the aim is to improve economic development, output-oriented DEA would be most useful.

In DEA model there can be several input and output variables for each DMU. The variables are linearly aggregated using the weights determined by the DEA methods. From that, virtual inputs and outputs are developed (Cotte Poveda, 2012):

$$Virtual\ input = v_1 x_1 + \cdots + v_n x_n = \sum_{i=1}^{n} v_i x_i \quad (12)$$

$$Virtual\ output = u_1 y_1 + \cdots + u_m y_m = \sum_{i=1}^{m} u_i y_i \quad (13)$$

In these equations v_i represents the weight assigned to an input x_i and u_i represents the weight assigned to an output y_i in the linear aggregation. These equations then can be used to maximise the ratio using the following equation (Cotte Poveda, 2012):

$$Performance = \frac{Virtual\ output}{VIrtual\ input} = \frac{\sum_{i=1}^{m} u_i y_i}{\sum_{i=1}^{n} v_i x_i} \quad (14)$$

These calculations reduce multiple inputs and outputs into a single virtual input and single virtual output that can then be used to evaluate a specific DMU. The scores vary from 0 to 1, 0 marking complete inefficiency and 1 marking absolute efficiency. The DAE model determines the weights of the outputs and inputs by aiming to maximise the ratio between the weighted virtual output and weighted virtual input. The DMUs that have maximised efficiency are then called a reference set and receive efficiency value of 1. The reference set is then compared to each of the DMUs in the dataset. The analysed DMU is assessed by its divergence from the reference set. However, it is important to note that a unique reference set for each DMU is developed and is used as its benchmark (Cotte Poveda, 2012).

In general, data envelopment analysis presents what a DMU (country) can achieve with a level of resources it has or what level of resources is needed to efficiently achieve the level of progress that is already achieved. If input-oriented analysis is used, the aim is cost/resource reduction. If output-oriented analysis is used, the aim is to enhance observed effects. Reference set is used to develop a benchmark line for a DMU that is not efficient. The benchmark is also often called best practice frontier and can be illustrated as a line in a graph.

The Figure 9 below represents DEA results. It presents the results for both constant returns to scale (CCR) model and variable returns to scale (BCC) model. The CCR model illustration is a dashed line marked ‘CCR’. The BBC model is explained by the solid line marked ‘BCC’. Points P1, P2, P3, P4, and P5 illustrate the DMUs used in the analysis. The numbers next to the points’ names present the values of those DMUs for inputs (first number) and outputs (second number). Inputs are measured by *x* axis and outputs are measured by *y* axis.

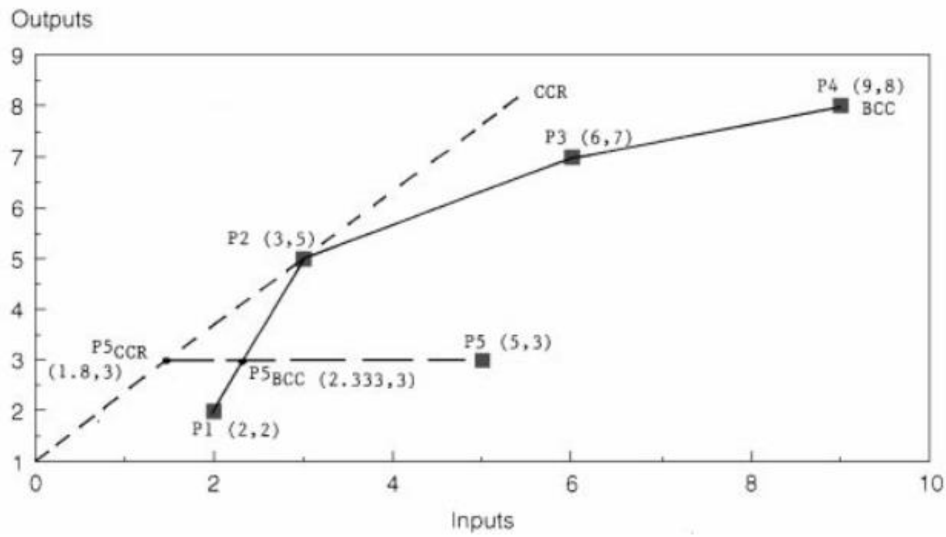

Figure 9. Illustrated example of DEA analysis (Bowlin, 1998, p. 10)
Dashed line presents the benchmark/best frontier for CRR/CRS model input-oriented model
Solid line presents the benchmark/best frontier for BCC/VRS model input-oriented model

Considering the illustration of CRR/CRS input-oriented model in the Figure 9, only DMU P2 is considered to be efficient and would have an efficiency score of 1. As presented, CRR, or constant returns to scale, model evaluates both technical and scale efficiency. Hence, it considers DMUs as efficient only when they both efficiently distribute their inputs and outputs, and operate at constant returns to scale point. In this model, DMUs P1, P3, P4, and P5 are considered as inefficient. In the graph, the distance between the DMU P5 and point $P5_{CCR}$ helps determine the efficiency score of DMU P5. In this example, the DMU P5 by CRR/CRS model receives efficiency score of 0.36. It means that to be considered efficient, this DMU would have to lower its input to 1.8, by 64%. At such point where inputs are equal to 1.8 and outputs are equal to 3 DMU P5 would be both technically and scale efficient (Bowlin, 1998).

Considering the illustration of BCC/VRS input-oriented model in the Figure 9, DMUs P1, P2, P3, and P4 are considered as technically efficient. They all receive efficiency score of 1. However, while all of these DMUs used in the reference set are technically efficient, only DMU P2 is both technically and scale efficient. This illustrates the difference between stricter CRR/CRS model and more flexible BCC/VRS model. DMU P2 has constant returns to scale, so it appears in reference sets of both models. The line between DMUs P1 and P2 represents locally increasing returns to scale. Between these points, an increase in inputs results in a larger increase in outputs. The line between DMUs P2, P3 and P4 represents locally decreasing returns to scale. Between these points, an increase in inputs result in an increase of outputs that is proportionally smaller. DMUs P1, P3 and P4 are technically efficient, but not scale efficient, because they do not have constant returns to scale. In the graph, the distance between the DMU P5 and point $P5_{BCC}$ can be used to calculate the efficiency score of DMU P5 by BCC/VRS input-oriented model. Efficiency score for this DMU in this example is 0.47. It indicates that to achieve efficiency DMU P5 should reduce its inputs by 53% (Bowlin, 1998).

Data envelopment analysis has several assumptions that should be met so that this analysis approach could be used. Firstly, it is assumed that selected DMUs are to some extent homogenous. They should have the same inputs and outputs so that they would be comparable. Conditions under which the DMUs operate should also be similar to some degree so that the comparison between them would be logical. Secondly, it is assumed that the amount of data provided is sufficient for a critical evaluation. If a number of DMUs is too small, the results of the analysis may be too optimistic (Golany & Roll, 1989). There are no strict mathematical rules how many DMUs should be included in the analysis. General rule is to have at least twice as many DMUs as inputs and outputs combined (Cook et al., 2014). However, Banker, Charnes, Cooper, Swarts and Thomas stated that there should be at least three times more DMUs than the number of inputs and outputs combined not to lose discrimination power in the analysis (Banker et al., n.d.). Thirdly, it is assumed that the data is in appropriate scale so that no single input or output is dominating the analysis. However, Cook, Tone and Zhu argue that mixture of raw data and percentiles or ratios can be allowed if the data is properly mixed, for example, ratio values are used as inputs and total values are used as outputs. The authors point out that in VRS/BCC model with ratio variables projections of the analysis remain in the range from 0 to 100, but when using CRS/CRR model one should exercise caution as the projection values may exceed 100 (Cook et al., 2014). Considering the values of the input and output variables it is also important to make sure that there are no negative values or zero values. DEA is concerned with the ratios of the presented variables, so negative or zero values may present misleading or undefined results of the analysis.

In this study the DAE method will be implemented following these steps:

1. The DMUs are chosen, meaning the scope of the analysis is determined;
2. Inputs and outputs are defined;
3. Appropriate parameters for DEA model are set and the model is applied;
4. The EU Member States are classified according to the newly developed efficiency index.

Considering selection of DMUs, it is important to keep in mind that DEA approach aims to assess comparable units with an aim to improve their performance in assessed area. The number of units analysed is also important. While a larger set allows to more accurately identify relationships between inputs and outputs, it also results in higher heterogeneity within the dataset, potentially causing the results to be influenced by exogenous factors that are not the focus of the analysis (Golany & Roll, 1989). In this study, the selected units for analysis are 27 EU Member States.

After DMUs are selected, inputs and outputs can be defined. As mentioned, the general rule is to include at least twice as many units as the number of inputs and outputs included in the analysis (Golany & Roll, 1989). While this rule does not necessarily have to be satisfied, having more inputs and outputs may diminish the discrimination power. At the same time, it is important to ensure that relevant inputs and outputs are included in the analysis. Inputs and outputs generally depend on the specific problem. While there may be different scenarios that are solved with DEA, the aim of such analysis, in most cases, is to minimise the inputs and maximise the outputs (Cook et al., 2014). In this study the same approach is followed. Energy poverty indicators are treated as inputs and the aim is to minimise them, while indicators of the use of renewable energy sources are treated as outputs and higher quality of them is desired.

Once the DMUs, inputs and outputs are determined, DEA model can be chosen, and the analysis can be implemented. DEA model’s specifics depend on the problem analysed. In this study, only a basic model of DEA is considered as it is deemed appropriate for the aim of the analysis. Variable returns to scale input-oriented model is chosen. It is also important to determine what data is used for the analysis. In DEA one can choose which observations are evaluated and which observations are used as reference data. Evaluated observations get results about their performance, while reference observations are used to understand how efficiency may differ in the countries and determine what is considered to be efficient and what is not efficient. Same observations can be used as evaluated and reference set. However, they can also differ. The selection of model and its justification is presented in more detail in results section 3.3.

2.3. Validation of the index through correlation analysis

In this study, correlation analysis is used as a method to validate the values of the newly computed indices. Correlation analysis is a useful tool to examine associative relationships between different variables. In this context, correlation analysis is used to understand whether the newly created indices are related with other indicators that are related to the same issue of energy poverty. The chosen indicators are presented in section 3.4. If correlation analysis indicates correlation between the newly created indices and selected parameters, it will indicate that the newly constructed indices are measuring aspects of energy poverty, as it should.

Correlation can be assessed through several different methods. The most commonly known method is Pearson's correlation. This correlation is computed as follows (Makowski et al., 2020):

$$r_{xy} = \frac{cov(x,y)}{SD_x SD_y} \tag{15}$$

In this equation, *cov(x,y)* stands for the covariance of the two variables in question, *x* and *y*, while SD_x and SD_y stand for standard deviations for *x* and *y* variables (Makowski et al., 2020).

Pearson's correlation is used to assess linear relationship between two variables and cannot detect non-linear relationships. It is sensitive to outliers and assumes that the variables analysed are normally distributed. This correlation analysis method is widely used and easy to interpret.

If data is not normally distributed, there are outliers or the relationship between the variables is suspected to be non-linear, Spearman's rank correlation coefficient can be used. It is a non-parametric correlation measurement, which assesses monotonic relationships. Spearman's rank correlation can be also expressed as Pearson's correlation between the rank scores of the two variables in question. It is computed as follows (Makowski et al., 2020):

$$r_{s_{xy}} = \frac{cov(rank_x, rank_y)}{SD(rank_x) * SD(rank_y)} \tag{16}$$

In this equation, $cov(rank_x, rank_y)$ stands for the covariance of the ranks of two variables in question, *x* and *y*, while $SD(rank_x)$ and $SD(rank_y)$ stand for standard deviations of the ranks of *x* and *y* variables respectfully (Makowski et al., 2020).

While Spearman's rank correlation coefficient is less sensitive towards outliers, does not require normal distribution and can detect non-linear trends, it is important to keep in mind that it is less sensitive to linear trends than Person's correlation.

More correlation analysis methods exist. However, in this study only Pearson's and Spearman's correlations were considered for its straightforward interpretability. Moreover, initial descriptive statistical analysis indicated that the data used in the study does not require more elaborate correlation analysis methods.

2.4. Data for the index

The analysis of this study covers 27 EU Member States – Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden. Selected time period includes 13 years – from 2010 to 2022. This time period was selected due to data availability. This results in a dataset with 351 observations.

The aim of this study is to construct a green energy poverty index that can be used to evaluate energy poverty in the EU in the light of green transition. The index in this study has indicators for two different issues – energy poverty and progress in green transition or sustainability, measured by the use of renewable energy. The selected variables are presented in the Figure 10 below.

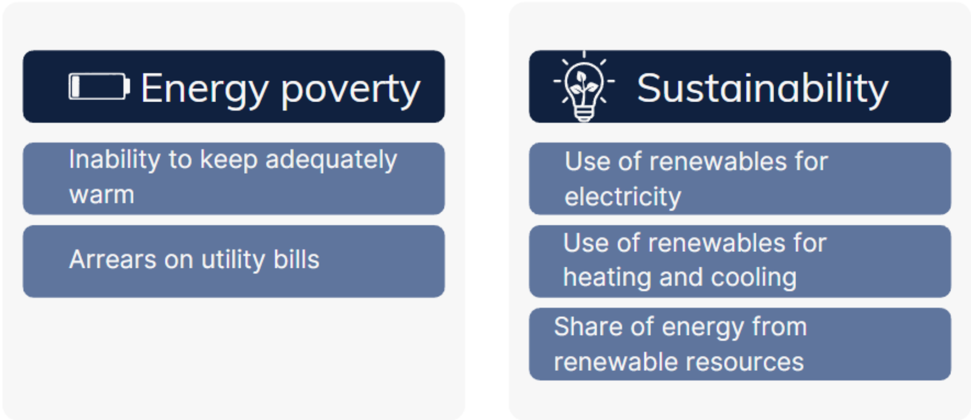

Figure 10. Indicators used to construct green energy poverty index

Considering specific indicators for energy poverty dimension, recent energy poverty indices consider metrics such as energy consumption, affordability, reliability, quality, availability and safety. However, there is no consensus on which of these metrics are necessary to adequately measure energy poverty and what thresholds for each of the mentioned dimensions are desirable in energy poverty index (Pelz et al., 2018). As presented in the Literature review section, in developed countries, as the countries in the EU, one of the most important dimensions is affordability, while dimensions as infrastructure or availability are of secondary importance (Ruiz-Rivas et al., 2022).

Consequently, the green poverty energy index focuses on the indicators of affordability dimension. The EU Energy Poverty Advisory Hub proposes to focus on inability of a household to keep their home adequately warm, high share of disposable income being spent on electricity, low absolute expenditure on electricity, and having to borrow to pay the bills for energy (Widuto, 2023). This study uses inability of a household to keep their home adequately warm and having to borrow to pay the bills for energy.

Share of disposable income being spent on electricity has been considered. However, Principal Component Analysis showed that the variable is not significant as the weights given to the variable by first two principal components (PCs that had Eigenvalue higher than one) were small compared to other variables. Absolute expenditure on electricity has been excluded because the change in its value should be interpreted differently than of other indicators. It means that while a higher value of selected measures indicates higher energy poverty, higher absolute expenditure on electricity indicates lower energy poverty. While this would not be problematic for the Principal Component Analysis, it would require additional data manipulation for Data Envelopment Analysis. As the variable is very strongly corelating with the share of disposable income being spent on electricity, the variable which was deemed insignificant, it was decided to exclude absolute expenditure on electricity as well.

To illustrate renewable energy use, the constructed index includes three variables that directly show the use of renewable energy sources. These variables are the use of renewables for electricity, the use of renewables for heating and cooling, and share of energy from renewable resources. Higher values of all of these variables indicate the higher level of green transition.

3. Results and discussion

This section presents the results of the implemented approaches to constructing the green energy poverty indices. The section is constructed as follows: initial data analysis is presented, the experiments conducted are explained in detail, and the results and their interpretation are presented.

3.1. Initial data analysis and manipulation

All of the variables selected are available in Eurostat database and were extracted through *Python* using *eurostat* library. The main descriptive statistics and other information about the selected variables are presented in the table in Appendix 2. The graphs showing the change of all indicators' values for each country in the data set are also available in the Appendix 2. Descriptive statistics indicate a great level of variance in the data. For majority of the indicators chosen, the standard deviation is greater than both mean and median of the data set. The chosen indicators also differ greatly in values, which indicates a potential need for data standardisation before PCA.

As it can be seen from the graphs presenting different countries (available in Appendix 2), some of the countries have significantly higher values for some of the variables than others throughout the whole observed period. For example, Greece has significantly higher percentage of population that is unable to keep their homes adequately warm. At the same time, Germany has significantly higher use of renewables for electricity and for heating and cooling. In the analysis, the values for these variables for these countries will likely appear as outliers. This can be confirmed looking into the box plots also available in Appendix 2. However, taking them out may reduce the usability of the index for the EU. Hence, to deal with the outliers robust PCA is used instead of traditional PCA when constructing the index. While DEA may generally be sensitive to outliers, the closer examination of the data for individual countries (presented in Appendix 2) indicates that the higher values of specific indicators illustrate specific individual countries. Hence, looking from individual DMU (country) perspective, those values could not be considered as outliers. Hence, they are not problematic for data envelopment analysis.

Variables that account for energy poverty (inability to keep adequately warm and arrears on utility bills) are percentages from population, while variables concerned with progress towards green transition (use of renewables for electricity, use of renewables for heating and cooling, and share of energy from renewable resources) are measured in absolute values. As the variables are measured in different scales, standardisation is highly recommended for PCA. For DEA, non-standardised values can be used as percentage values and absolute values are used as inputs and outputs respectively and are not mixed. For PCA, data was scaled using median absolute deviation (MAD). This means that when computing scaled values, the distance between each data point and the median is measured instead of measuring the distance to the mean. This makes the scale more robust to outliers. The following formula (15) is used to scale the data:

$$X_{i_{scaled}} = \frac{(X_i - median(X_i)}{MAD_{X_i}} \tag{17}$$

$$\text{where } MAD_{X_i} = median(|X_i - \tilde{X}|) \tag{18}$$

Where X_i is a specific observation, and $\tilde{X}$ is median of X.

This scaling method was chosen due to the existence of outliers in the dataset. Scaling was incorporated in the *PcaCov* function of the *rrcov R* package used for PCA.

Finally, the selected dataset was assessed for correlation. As already mentioned, selected variables should be correlated with each other for PCA to be an appropriate analysis approach. However, multicollinearity may be a problem and should be addressed before the analysis and collinear terms should be eliminated or combined before applying PCA (OECD, 2008). In the figure and table below, the correlation matrix for the variables, which are used in the green energy poverty index, is presented. *Spearman's ρ* is used to evaluate relationships between the variables. This correlation statistic was chosen because it is not-parametric, meaning it does not assume that the data follows normal distribution, and robust to outliers. Normality test results are available in Appendix 4 and indicate that indeed the variables are not normally distributed. As it can be seen, the correlation between the variables differs. Weakest positive correlation appears between arrears on utility bills and share of energy from renewable resources (0.019). Highest positive correlation is detected between use of renewables for electricity and use of renewables for heating and cooling (0.899). Weakest negative correlation exists between inability to keep adequately warm and share of energy from renewable resources (-0.118). Highest negative correlation appears between arrears on utility bills and use of renewables for electricity (-0.329). The only potentially concerning relation is between use of renewables for electricity and use of renewables for heating and cooling. Correlation between other variables does not exceed 0.7 or -0.7, indicating that the correlation between the variables likely does not imply collinearity.

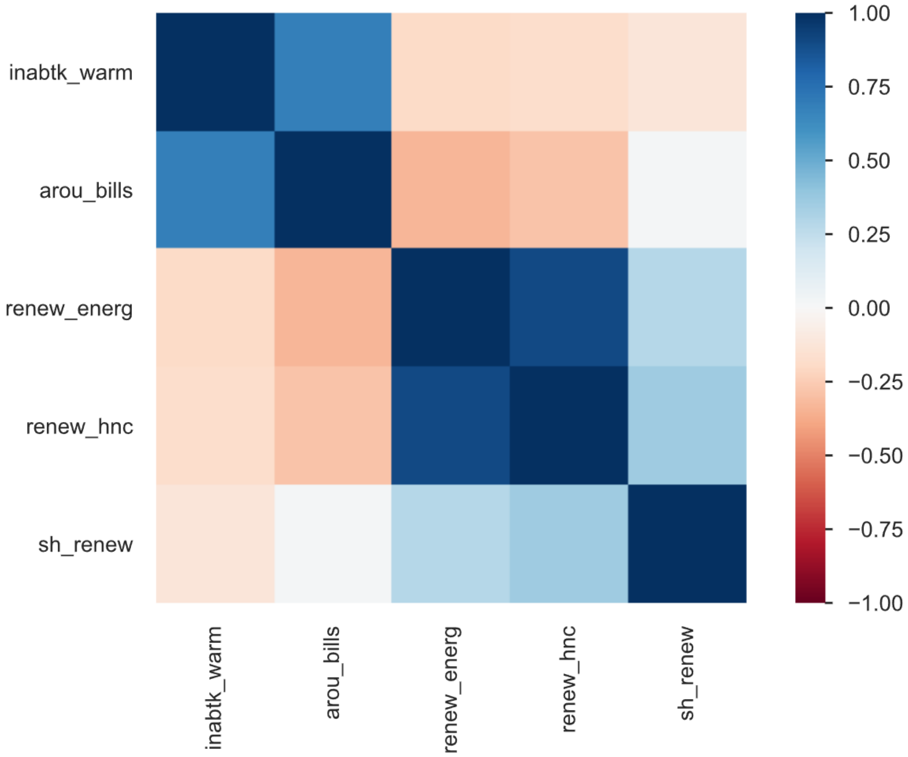

Figure 11. Correlation heatmap (Spearman's ρ) for variables used for the index

	Inability to keep adequately warm	Arrears on utility bills	Use of renewables for electricity	Use of renewables for heating and cooling	Share of energy from renewable resources
Inability to keep adequately warm	1.000	**0.687 (0.000)**	**-0.192 (0.000)**	**-0.172 (0.000)**	**-0.118 (0.027)**
Arrears on utility bills	**0.687 (0.000)**	1.000	**-0.329 (0.000)**	**-0.289 (0.000)**	0.019 (0.716)
Use of renewables for electricity	**-0.192 (0.000)**	**-0.329 (0.000)**	1.000	**0.899 (0.000)**	**0.288 (0.000)**
Use of renewables for heating and cooling	**-0.172 (0.001)**	**-0.289 (0.000)**	**0.899 (0.000)**	1.000	**0.358 (0.000)**
Share of energy from renewable resources	**-0.118 (0.027)**	0.019 (0.716)	**0.288 (0.000)**	**0.358 (0.000)**	1.000

Table 1. Correlation (Spearman's ρ) for variables used for the index
p-values are presented in the brackets
Bolded correlations are significant

To conclude, correlation analysis indicates that variables used for the construction of the index are corelating with each other. Majority of the correlation relationships is moderate, which is necessary for PCA. However, as some of the correlations are strong, there is a risk of multicollinearity. This means that the performed PCA analysis should be also checked for multicollinearity by examining the condition index values for different principal components. This is discussed in section 3.2.

3.2. Results of PCA

Descriptive data analysis indicated that the selected dataset does not meet some of the assumptions of traditional PCA. More specifically, the dataset includes several outliers that cannot be omitted without losing an important share of information. Hence, robust PCA has been chosen for the analysis. To implement robust principal component analysis, *rrcov* package in *R* was used. As already mentioned, a robust scaler with median absolute deviation (MAD) was used to standardise the data for PCA. To address the issue of outliers in PCA itself, robust PCA with MCD estimator was used.

The results of the robust PCA are presented in the Table 2 below. As five variables were used, the analysis resulted in five principal components. The first principal component explains 68.97% of variance in the original dataset, second PC explains 16.39% of variance, adding to the total of 85.36%, and the third PC explains 10.81% of variance in the original dataset, adding to the total of 96.17% of variance. The rest of the principal components explain less than 5% of variance each and can be seen as less relevant for the analysis. First principal component, which assigns largest weights to inability to keep warm and arrears on utility bills, explains more than half of the variance in the original dataset. Hence, it is clear that these two indicators may be some of the most important in understanding and explaining the variance between countries.

While interpreting principal components, it is also important to consider condition index, which helps determine whether there are issues of multicollinearity. If the value of the condition index is between 10 and 30, it signals potential multicollinearity. If the value of the index for the PC is higher than 30, it indicates strong potential multicollinearity. It may also mean that the principal component with an index higher than 30 contains residual noise rather than necessary information (Kim, 2019). From the results in the Table 2 below, it seems that only the data in the last principal component of the principal components may have a problem of multicollinearity. It is likely that this component contains residual noise as it is the last component. However, as this principal component is not retained in the further analysis, its potential problem of multicollinearity is irrelevant. The decision which principal components are retained is explained later on in this section.

	PC1	PC2	PC3	PC4	PC5
Statistics of the PCs					
Eigenvalue	6.075	1.444	0.952	0.329	0.010
Proportion of variance	68.97%	16.39%	10.81%	3.74%	0.09%
Cumulative variance	68.97%	85.36%	96.17%	99.91%	1.000
Condition index	1.000	2.051	2.526	4.298	27.300
Standard deviation	2.465	1.202	0.976	0.573	0.090
Weights of each variable in each principal component					
Inability to keep adequately warm	0.795	-0.469	0.377	0.071	-0.039
Arrears on utility bills	0.606	0.587	-0.530	-0.065	0.044
Use of renewables for electricity	-0.013	0.155	0.013	0.732	-0.663
Use of renewables for heating and cooling	-0.000	0.181	0.184	0.626	0.736
Share of energy from renewable resources	0.030	0.615	0.736	-0.252	-0.121

Table 2. Results of robust PCA

The plot in Figure 12 below shows how the observations are distributed according to the first and second principal components. PCA was applied to the whole data set, meaning observations from all 27 countries of 13 years period were included. It is not surprising that the observations from different year for the same country are mostly close together. The first principal component has larger weights for indicators for energy poverty (inability to keep warm and arrears on utility bills) and very low weights for all three indicators for the progress in green transition (use of renewables for electricity, use of renewables for heating and cooling and share of energy from renewable resource). Hence, it is not surprising that countries like Bulgaria (BG) or Greece (EL) that generally tend to have more socio-economic challenges score high on the axis of the first PC. Second principal component assigns large weight for one indicator of energy poverty (arrears on utility bills), which may explain the high score of Greece (EL) on the second PC axis as well. However, as this component also has larger weights for three indicators indicating the progress in green transition, countries that are well known for their focus on renewable energy resources, such as Germany (DE) and Sweden (SE) also have high scores on the PC2 axis.

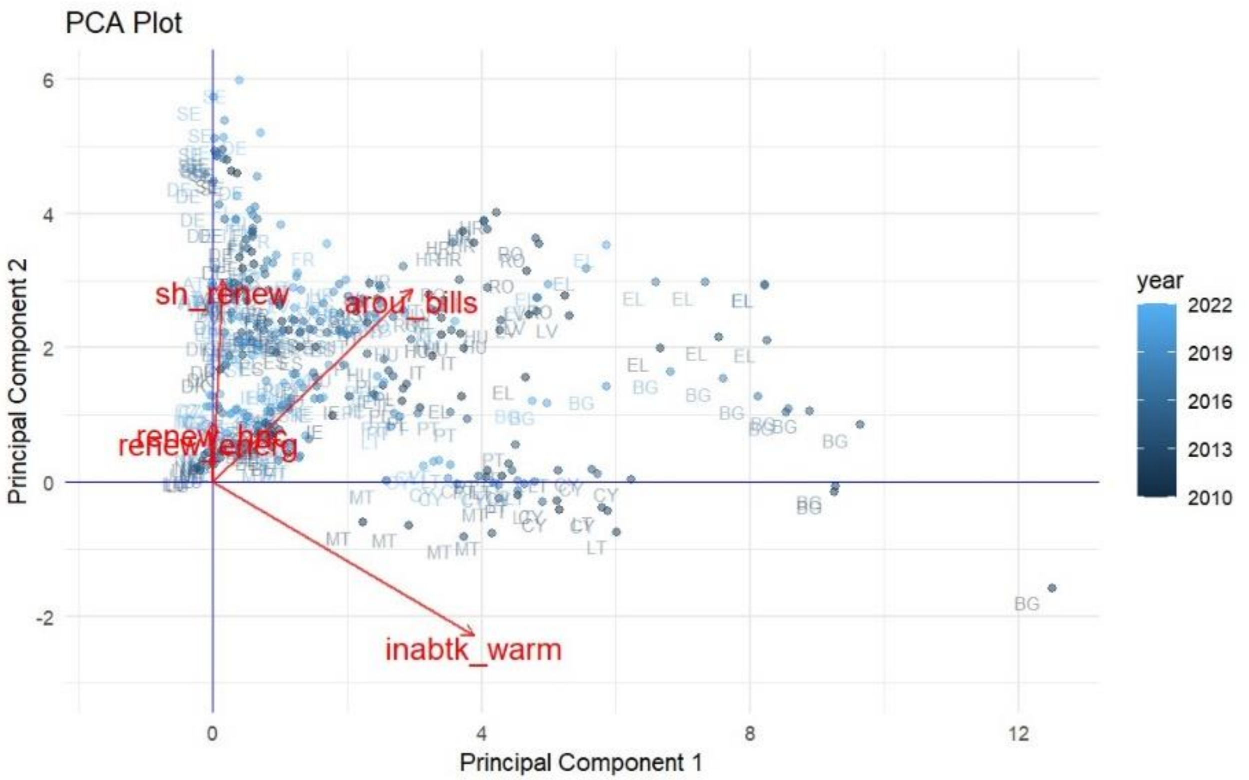

Figure 12. PCA plot for robust PCA

While the results of the robust PCA already shines some light on how the variance in the original dataset and difference between the countries can be explained, additional analysis is necessary before constructing the index. Principal component analysis resulted in five different principal components. However, not all of the components are relevant for further analysis. In most cases, principal components are retained if they have eigenvalues higher than 1 (Kaisen criterion), explain more than 10% of variance in the original dataset, or together account for a set specific share of variance (target usually varies from 70% to 90%) (Nardo et al., 2005). As it can be seen from Table 2 above, first two components have eigenvalues higher than 1. The share of explained variance is higher than 10% for first two components. First two components accumulatively account 85.36% of the variance, while first three components explain 96.17% of the variance. This indicates that an optimal number of principal components is two.

In this study, to determine the number of PCs to keep, it was decided to use *findPC* package in R that uses computational methods to identify elbow points to determine which PCs could be kept. The elbow method is often used to determine the number of relevant PCs. According to this method, optimal number of principal components is identified from the scree plot of standard deviations of principal components. The point where the elbow in the curve occurs is considered as the optimal number of principal components. The used package has six different computational methods to find the optimal number of PCs (Zhuang et al., 2022). As it can be seen from the graph below, all methods suggest two principal components as optimal. Hence, in further computations only first two principal components are used.

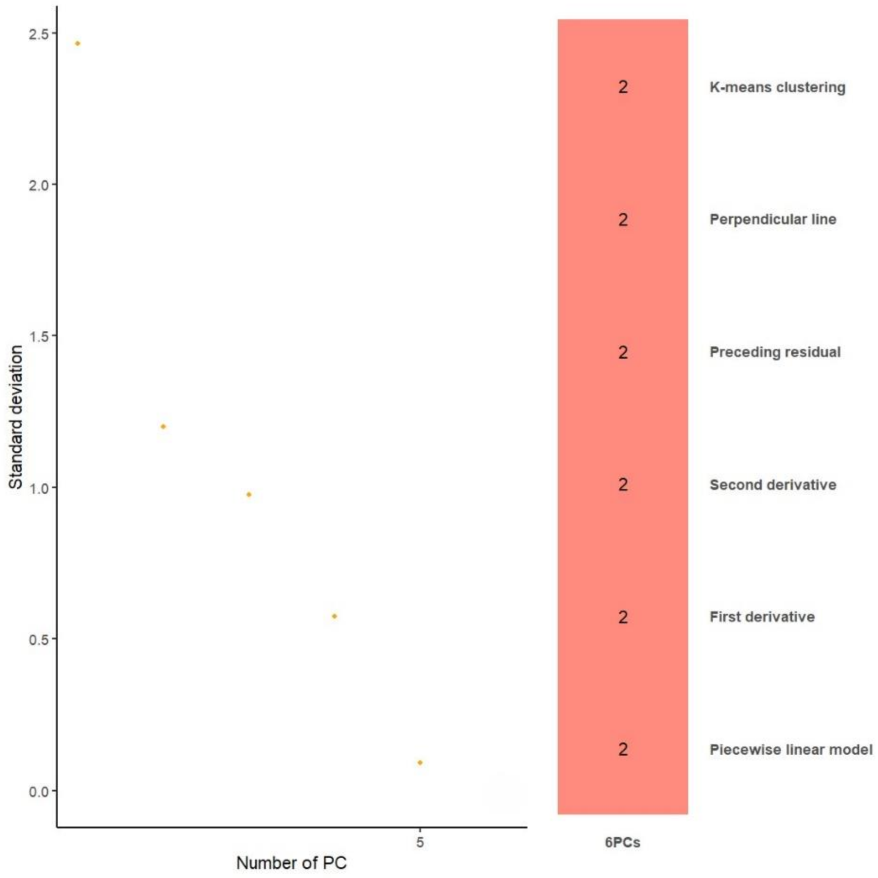

Figure 13. Elbow method to determine an optimal number of PCs

The weights of each indicator in the overall index are computed using eigenvalues of the principal components and the weights they contribute to individual indicators. This study uses the same method as already presented in the study constructing Household Energy Poverty Index for India using PCA (Gupta et al., 2020). This means that formula (5) is used to compute the weights for each indicator.

The weights for each indicator for green poverty index are presented in the table below. The weights are scaled to have values from -1 to 1 for better interpretability. While usually the scaling is done in a way to ensure that the weights add to 1 (Nardo et al., 2005), in this study it was decided to take a different approach. The variables used in the study present two different dimensions – energy poverty and progress towards the green transition. As can be seen from the presentation of PCA results, two main principal components, which have the most weight in index creation, also to some extent present these two different dimensions. Therefore, having positive and negative index weights makes the index easier to interpret. Indicators with negative weights (energy poverty indicators) contribute to a lower value of an index, while indicators with positive weights (sustainability indicators) contribute to a higher value of an index.

	Weight
Inability to keep adequately warm	-0.826
Arrears on utility bills	-1.000
Use of renewables for electricity	1.000
Use of renewables for heating and cooling	0.949
Share of energy from renewable resources	0.580

Table 3. Weights for the Green Energy Poverty Index

Using the weights and values of the chosen indicators, it is possible to calculate the values of the index for each country for each year available in the dataset. However, it is important to keep in mind that the PCA was carried out using scaled data, which means that to calculate the values of the index the data also has to be scaled using the same scaler (using median absolute deviation (MAD)). The scaled values of the indicators can be computed using the formula (17). For that, median of each indicator and MAD of each indicator is needed. MAD is computed using formula (18). These statistics are presented in the table below.

	Median	MAD
Inability to keep adequately warm	6.1	3.700
Arrears on utility bills	7.4	3.800
Use of renewables for electricity	11718.408	10490.139
Use of renewables for heating and cooling	1666.2	1381.579
Share of energy from renewable resources	17.852	6.993

Table 4. Median and MAD for variables used in the green energy poverty index

With these values, the index value can be easily calculated for a selected country for selected year. The following formula should be used:

$$GEPI = renew\_energ + 0.949 * renew\_hnc + 0.580 * sh\_renew \quad (19)$$

$$-0.826 * inabtk\_warm - arou\_bills$$

Values of the green energy index for 27 EU Member States for 2022 are available in Appendix 3. As it can already be seen from the weights of the index, lower values mean that the country is energy poor and lagging behind in green transition, while higher values indicate low energy poverty and more progress towards green transition.

To make the values of the index more interpretable, it was decided to scale the values to fit into the scale from 0 to 1. Standardised index values for the last year available (2022) can be found in Appendix 3. The values of the standardised index range from 0, the most energy poor and the least advanced in green transition, to 1, the least energy poor and most advanced in the green transition process through the use of renewable energy resources. However, it is important to keep in mind that the values of the index were scaled, so both 0 and 1 do not refer to theoretical situations of complete energy poverty or complete green transition. These values refer to the worst and best performance observed in 27 EU Member States in the period between 2010 and 2022.

Between 2010 and 2022 period, the best performing country was Germany in 2022, when the value of standardised green energy poverty index for Germany was equal to 1. The worst performing country was Bulgaria. It obtained a standardised green energy poverty index value of 0 in 2010. The Table 5 below presents the descriptive statistics for the Green Energy Poverty Index. Average value for the index is equal to 0.402. Standard deviation is 0.163, which means that the variance in the data is moderate. As only 25% of the observations have an index value higher than 0.500, in majority of the observations the countries can be considered as moderately to highly energy poor and lagging behind in green transition.

Descriptive statistics	Green energy poverty index by PCA
Count	351
Mean	0.402
Standard deviation	0.163
Minimum	0.000
1st quartile	0.304
Median	0.365
3rd quartile	0.500
Maximum	1.000

Table 5. Descriptive statistics for Green Energy Poverty Index

The histogram in Figure 14 below shows the distribution of the index values. As it can be seen, majority of values are clustered between 0.300 and 0.500, which is also seen from the values of 1st (0.304) and 3rd (0.500) quartiles presented in Table 5. This means that majority of the observations vary from highly energy poor and highly lagging behind in green transition to mildly energy poor and mildly lagging behind in green transition.

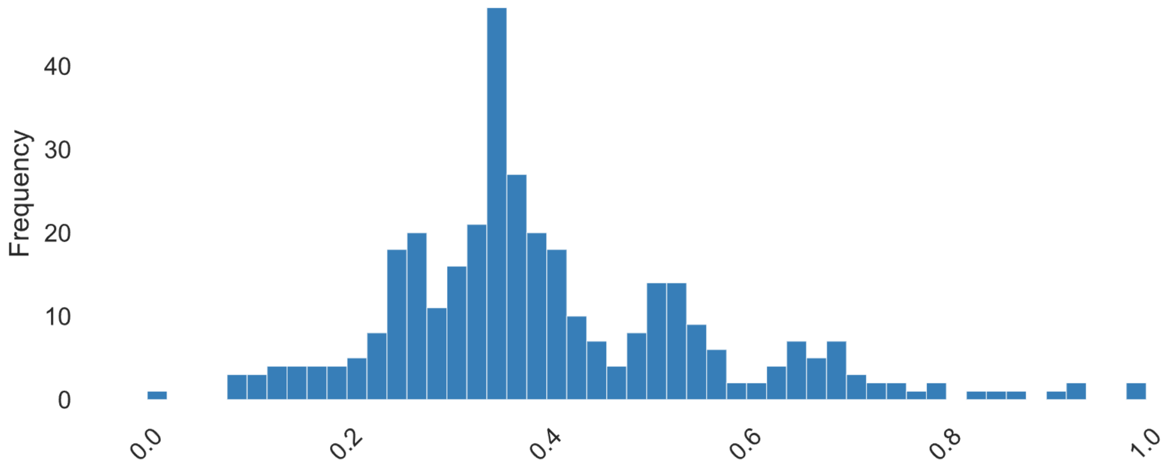

Figure 14. Histogram for Green Energy Poverty Index

To summarise, PCA was implemented to better understand how the selected variables contribute to explaining the variance observed in the data. First two principal components, that explain 85% of the whole variance, have been determined to be significant components and have been used to derive the weights for selected variables in the Green Energy Poverty Index. First principal component focused more on energy poverty dimension, while second principal component was more concerned with the progress towards green transition. The weights of the index were standardised for simpler interpretation. The variables for energy poverty (inability to keep adequately warm and arrears on utility bills) received negative weights, while variables for renewable energy use (use of renewables for electricity, use of renewables for heating and cooling, share of energy from renewable resources) received positive weights. This means that lower values of the index are related to greater energy poverty and lagging in green transition. Majority of the observations from EU Member States between 2010 and 2022 point to moderate or high energy poverty. More detailed interpretation of the index values is available in section 3.5.

3.3. Results of DEA

The first step of data envelopment analysis is to select the scope of the analysis through the selection of DMUs. The aim of the DEA is to assess similar units, in this case, countries, and provide suggestions on how to improve their performance. Therefore, selected units of analysis should be similar to be comparable, but some differences between them are desirable for meaningful insights (Golany & Roll, 1989). The selected units for analysis (DMUs) in this study are 27 EU Member States. These countries are similar to each other considering their geographical location (all of them are in Europe), culture, development level (all countries are developed countries). However, they are different in climate and some socio-economic factors.

Second step of DEA is to determine which variables are used as inputs and which are defined as outputs. It was decided to treat energy poverty indicators (inability to keep adequately warm, arrears on utility bills, and housing cost overburden) as inputs and green transition indicators (use of renewables for electricity, use of renewables for heating and cooling, and share of energy from renewable resources) as outputs. The variables are already presented in section 2.4. As there are two inputs and three outputs and 27 DMUs, the selection satisfies the general rule to have at least twice as many DMUs as inputs and outputs (Golany & Roll, 1989). Energy poverty indicators are chosen as inputs and green energy indicators as outputs as the aim of DEA is to minimise energy poverty (inputs) and the greater amount of renewable energy sources (outputs) is desired. This means that the analysis follows the classical use of DEA where the aim is to minimise inputs and maximise outputs.

The manipulation of the data (scaling) before the analysis in this case is not required. While the used dataset includes both percentage values and total values, it is unlikely that it negatively affects the analysis. As pointed out by Cook et al., mixture of absolute and ratio values can be used in some cases, for example, when ratio values are used as inputs and total values are used as outputs. The DEA projections also should not exceed 100 when VRS/BCC model is used (Cook et al., 2014). This approach is followed in this study. The ratios (inability to keep adequately warm, arrears on utility bills) are used as inputs, and absolute values (use of renewables for electricity, use of renewables for heating and cooling, and share of energy from renewable resources) are used as outputs.

The DEA model used in the analysis is BCC/VRS input-oriented model. The analysis was carried out using *R* package *DeaR*. Variable returns to scale model has been chosen for its flexibility. In this model the DMU, in this case country, needs to be only technically efficient to be considered as efficient as scale efficiency is not required. As scale efficiency is not the aim of this exercise, CRS/CRR model is deemed as not suitable. The input-oriented model was chosen as the analysis focuses on the reduction of energy poverty (inputs).

As already presented in the literature review section, use of renewable energy sources and progress in green transition may compensate for the negative effects of energy poverty in the development of a country (Adom et al., 2021). Use of renewable energy sources may also result in lower energy poverty as it may improve availability, efficiency and affordability of energy (Dong et al., 2021). However, at the same time, if transition towards renewable energy is too rapid or not well thought-out, it may further worsen a problem (Hussain et al., 2023), and have long-lasting negative consequences (Adom et al., 2021). Hence, the main aim of DEA in this study is to determine which countries are efficiently using their renewable energy resources to achieve lower energy poverty. The countries that receive an efficiency score of 1 can be considered as best performers. They have achieved the lowest energy poverty possible with the amount of renewable energy resources they are using. Countries with an efficiency score lower than 1 have not achieved the highest efficiency possible, meaning they could have lower energy poverty indicators considering the amount of renewable energy resources used.

The data envelopment analysis was carried out three different times:

1. Using the whole dataset for both evaluation and reference;
2. Using data from 2022 for evaluation, and using the whole dataset as reference;
3. Using data from 2022 for both evaluation and reference.

The first analysis focuses on the whole dataset. 27 countries are evaluated for the whole 13 years period. Each observation is treated as a separate DMU. The figure below shows the distribution between efficient and inefficient observations in the whole dataset.

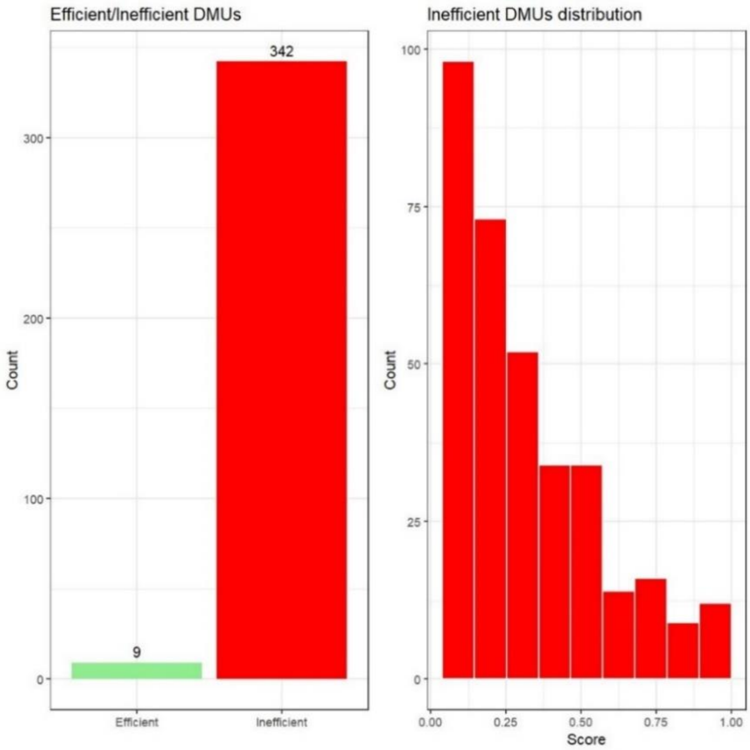

Figure 15. Distribution of efficient and inefficient observations (whole dataset)

Out of 351 observations, efficiency was achieved in nine observations by four countries:

- Luxembourg in 2010;
- Sweden in 2013, 2014, 2015, 2021 and 2022;
- Germany in 2019, 2021 and 2022;
- The Netherlands in 2021.

These best performing countries appeared in the reference set of the observations when a specific country was not efficient in their use of renewable energy to achieve lower energy poverty. This means that every time inefficiency was observed, the combination of these best performers was used to construct the line of marking an optimal amount of inputs and outputs and the distance between observed inefficient country and the constructed line is used to determine the efficiency targets. While the model is input-oriented, it focuses primarily on reduction of the inputs, energy poverty. However, as several inputs and outputs are used, the model presents target values not only for inputs, but also for outputs. In all cases at least one output remains the same in the target point and observed values and other outputs in the target point are slightly higher.

As can be seen from Figure 16 below, the Netherlands appeared in the reference sets the most. It was used as a reference for 279 observations. Luxembourg has appeared in 260 reference sets. Sweden as a country appeared in more reference sets than the Netherlands. However, it was through five different years, hence, as five different observations. Its performance in 2013 did not appear in any of the reference sets, its performance of 2014 appeared in 17, of 2015 – in 26, results of 2021 appeared in 42 reference sets, and Sweden's performance of 2022 appeared in 272 reference sets. Germany was included in the reference sets of inefficient observations the least. Germany's data from 2019 did not appear in any of the reference sets, its data from 2021 was used in 1 reference set, and its performance of 2022 appeared in 54 reference sets.

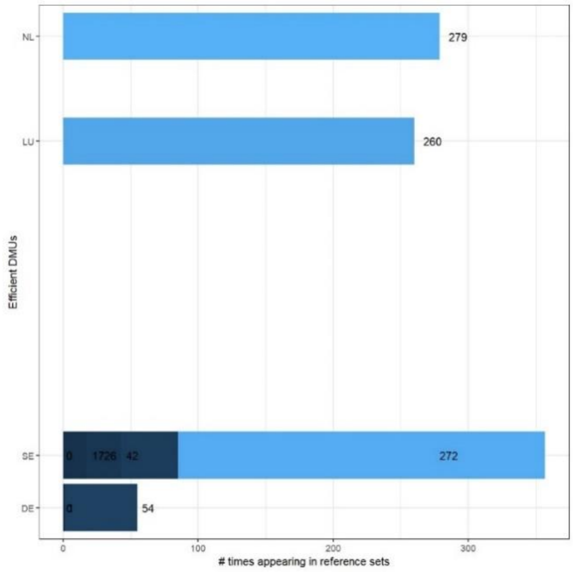

Figure 16. Best performers' appearance in refence sets (whole dataset)

Figure 15 presents not only the division between efficient and inefficient observations, but also the distribution of efficiency scores of the observations that are considered to be inefficient (have an efficiency score lower than 1). As it can be seen, around half of the observations have efficiency score between 0 and 0.25. As presented in descriptive statistics in Table 6, the lowest efficiency value is 0.039, which is the efficiency score of Bulgaria in 2010. It means that in 2010 Bulgaria would have had to decrease its energy poverty by around 96% to be considered as technically efficient. Average of all index values is 0.339. This means that on average in all observations the countries should have had around 64% lower energy poverty indicators to be considered as efficient. Standard deviation is equal to 0.256. The standard deviation being approximately two thirds of the mean points to a rather high variance in the data. This could be caused by unequal distribution of data points. Two thirds of the data being distributed in the efficiency values from 0.039 to 0.498 (3rd quartile), meaning the countries at the given time period were from mildly to highly inefficient in using their renewable energy resources in a way that contributes to energy poverty. At the same time, the remaining 25 % of the data is distributed between the values of 0.498 and 1.

Descriptive statistics	Efficiency index (DEA)
Count	351
Mean	0.339
Standard deviation	0.256
Minimum	0.039
1st quartile	0.137
Median	0.264
3rd quartile	0.498
Maximum	1

Table 6. Descriptive statistics for efficiency index (whole dataset)

The histogram in Figure 17 presents the distribution of the efficiency index values in slightly more details than Figure 15. However, it points to the same conclusions. Majority of the countries over the observed years were from inefficient in a way they are using renewable energy if the goal to improve energy poverty is considered.

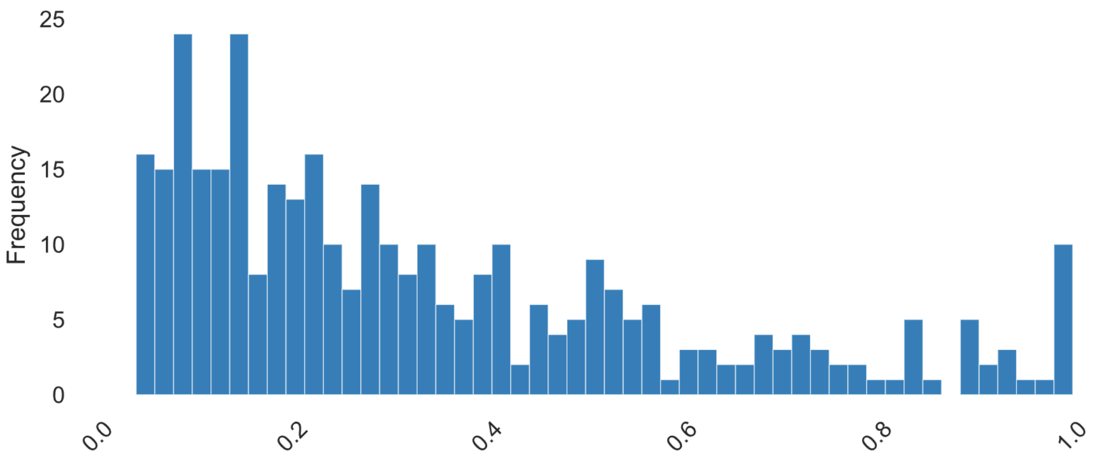

Figure 17. Histogram for efficiency index (whole dataset)

After the whole dataset was analysed, it was decided to examine only the data from the latest available year – 2022. Firstly, 2022 observations were used as DMUs for evaluation, and the rest of the observations were used as DMUs for reference. The analysis presents the same results as the analysis where the whole data set is used for both evaluation and reference. It finds that only two countries in 2022 can be considered as efficient (Germany and Sweden). The efficiency values for all countries for 2022 are available in the appendix 3. The Figure 18 below shows the number of efficient and inefficient countries and distribution of efficiency scores of inefficient DMUs.

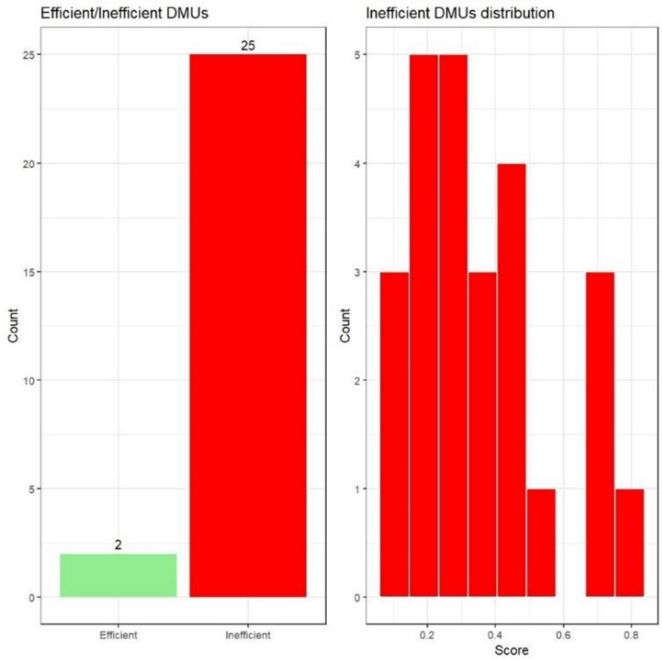

Figure 18. Distribution of efficient and inefficient observations (data from 2022 evaluated)

The Figure 18 indicates that majority of the countries have the efficiency score of 0.6 or lower. It means that all the countries besides Germany and Sweden could have had at least 40% lower values of energy poverty indicators with the same or slightly higher use of renewable energy use if renewable energy was used with the focus on improving accessibility and affordability of energy.

While only Germany and Sweden are considered as efficient in the 2022 dataset, reference sets for inefficient countries include also other observations that are considered as efficient in the whole dataset and are presented above. Figure 19 below presents which observations were used in reference sets and how often. Figure 20 present for which country each of the benchmark observations are used as reference. This diagram illustrates where the counted appearances from Figure 19 actually appear.

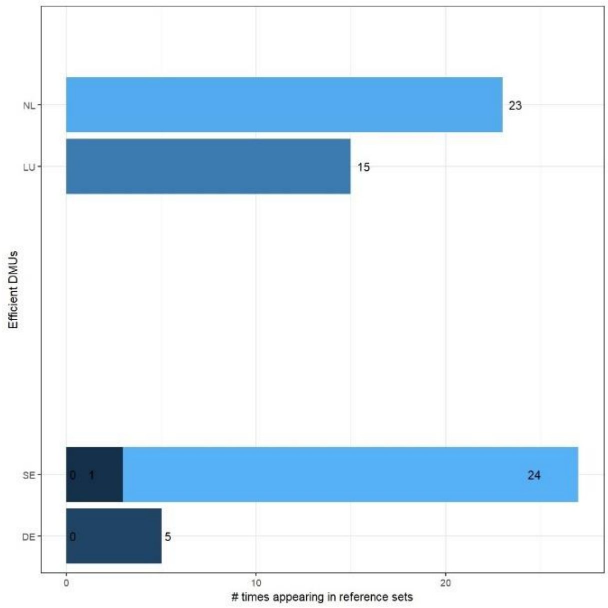

Figure 19. Best performers' appearance in refence sets (data from 2022 evaluated)

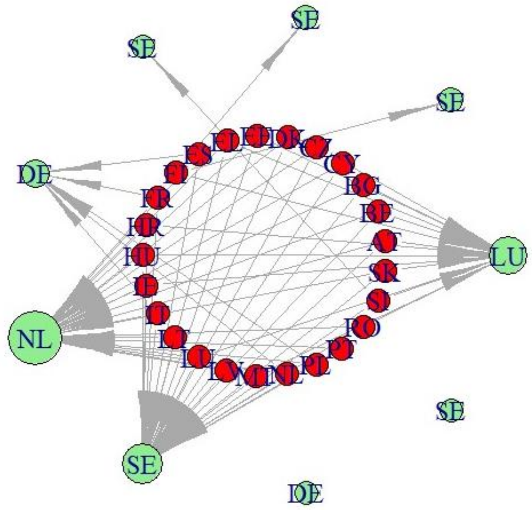

Figure 20. Use of best performers in reference sets (data from 2022 evaluated)

Points DE (Germany) and SE (Sweden) with no connections are Germany 2021 and Sweden 2013 observations respectfully. These best performers were not used in any of the reference sets. Sweden's performance from 2014, 2015 and 2021 (smaller SE points at the top) are used only in one reference set each. Two of these observations are used in the reference set for Greece (ES), while the third one is used for the reference set for Slovenia (SI). Sweden's performance in 2022 was used in the reference sets for majority of the countries, except for Finland (FI), Hungary (HR) and Lithuania (LT). Germany's performance from 2022 is included in the reference sets for Greece (ES), France (FR), Ireland (IE), the Netherlands (NL) and Poland (PL).

Finally, the analysis where only data from 2022 is used was carried out. It presents more optimistic results than the analysis of 2022 data with the whole dataset as a reference. In this analysis, six countries are considered to be efficient (Austria, Czech Republic, Germany, Finland, Netherlands, and Sweden). The results differ because of the amount of data used in the analysis. As known from the literature, if a number of evaluated DMUs is smaller, the results of the analysis may be overly optimistic (Cook et al., 2014). Hence, while in this section both, the analysis with the whole dataset and analysis with only the 2022 data, are presented, in later sections only the results of the analysis where the whole dataset is considered as reference DMUs are considered.

Figure 21 resents the division between efficient and inefficient countries and provides the distribution of efficiency scores for inefficient countries. As it can be seen, even though there are less inefficient countries, their scores are quite similar to the efficiency scores computed when the whole dataset is used as reference DMUs. Majority of the countries still have efficiency score of 0.6 or less, meaning they could have had at least 40% lower values of energy poverty indicators if renewable energy resources were used in a more efficient way, meaning focusing on improving accessibility and affordability of energy.

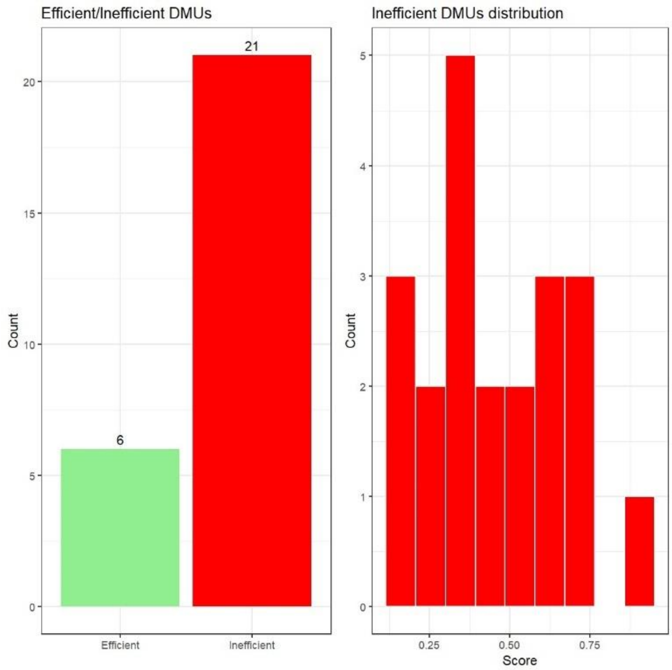

Figure 21. Distribution of efficient and inefficient observations (only data from 2022)

Figure 22 indicates how many times each of the best performing countries were used in the reference set for the countries that are considered inefficient. Figure 23 illustrates these connections between inefficient countries and the best performers used in their reference sets. As it can be seen, Czech Republic is used in the largest number of reference sets (17), while Austria and Germany are used in the smallest number of reference sets (4). This shows that not using observations from previous years as reference DMUs changes the relations between the data and presents different efficiency scores. While these scores are also informative, having more references presents a more critical view. For this reason, in the later sections the results for 2022 from the analysis with the whole dataset are referenced.

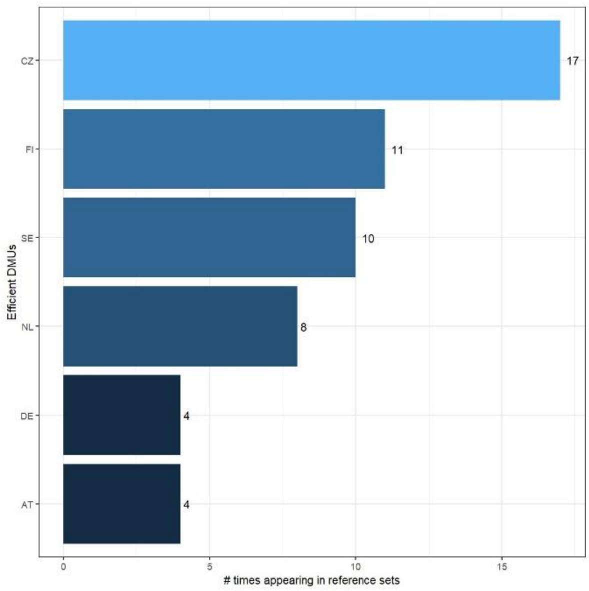

Figure 22. Best performers' appearance in refence sets (only data from 2022)

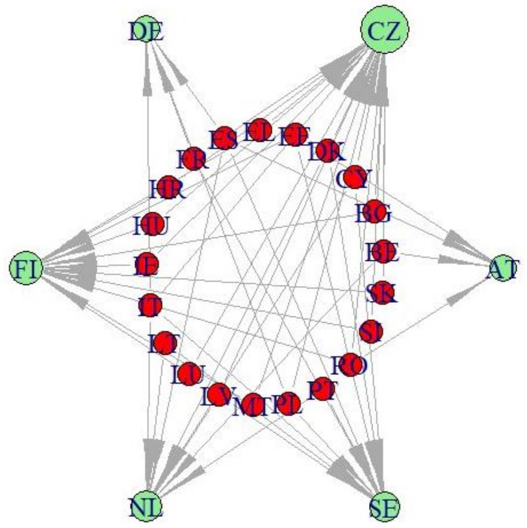

Figure 23. Use of best performers in reference sets (only data from 2022)

DEA results reveal how efficiently green energy resources are used to reduce energy poverty. However, it is important to keep in mind that the comparison of the results of different countries is rather complex. For example, if two countries have an efficiency score of 1, it does not mean that they have the same energy poverty. However, it means they both have the lowest energy poverty with the renewable energy resources they are using.

3.4. Validation of the Green Energy Poverty Indices

After the principal component analysis and data envelopment analysis approaches have been implemented and the indices are computed, the derived values can be validated. Correlation analysis is used for validation of the results of PCA and DEA. In this section, the results of this validation exercise are presented.

For correlation analysis, the index values are compared with the following indicators:

- Absolute expenditure on electricity – lower values of this indicators point to potential energy poverty;
- Disposable annual household income – lower values of this indicator point to potential energy poverty;
- Population considering their dwelling as too dark – higher values of this indicator point to potential energy poverty.

These variables were selected as they are mentioned as potential indicators for energy poverty in the literature. As already mentioned, low absolute expenditure on electricity is one of the indicators that the EU Energy Poverty Advisory Hub proposes to focus on when evaluating energy poverty (Widuto, 2023). As this indicator is not used to compute the green energy poverty index and efficiency scores, it is suitable for the validation of the index. Disposable annual household income and population considering their dwelling as too dark variables are also mentioned in the EU Energy Poverty Advisory Hub report from 2023 as indicators that can also provide information about energy poverty. Income is considered to be one of the main drivers of energy poverty. Therefore, disposable annual household income can help identify in which countries households are more vulnerable to energy poverty. Population considering their dwelling as too dark is also a relevant indicator because it can stem from the lack of natural or artificial lighting. If there's too little artificial lighting, this may be a result of low energy consumption that is often a sign of energy poverty (Gouveia et al., 2023).

All of these indicators are available in Eurostat. They were extracted in the same way as the indicators used for constructing the index through Eurostat library in *Python*. The descriptive statistics of these indicators is available in Appendix 2. Shapiro-Wilk test showed that all variables do not follow normal distribution. Normality test results are available in Appendix 4. While absolute expenditure on electricity has values for all countries (27 EU Member States) for the investigated period (2010-2022), other two indicators have missing values. Disposable annual household income variable has 295 values (compared with 351 index values) and population considering their dwelling as too dark variable has 319 values (compared with 351 index values). However, the number of missing values was deemed not high enough to affect correlation analysis.

Green Energy Poverty Index, which was constructed using principal component analysis, assesses how energy poor and how much lagging behind in the green transition an EU Member State is. As presented in section 3.2, the value of the index ranges from 0 (very energy poor and highly lagging behind in the green transition) and 1 (low energy poverty and great progress towards the green transition). Considering the meaning of the values of the index, the index should positively corelate with absolute expenditure on electricity and disposable annual household income and negatively corelate with the population considering their dwelling as too dark.

Data envelopment analysis produced an efficiency score that can be used to understand whether the countries are using their renewable energy resources in a way that contributes to decreasing energy poverty or in a way that is not socially inclusive. The values of the efficiency score can range from 0, completely inefficient, to 1, completely efficient. The interpretations of the values are presented in section 3.3. The interpretation of the index is slightly more complex that the interpretation of the index computed using PCA. Higher value, indicating higher efficiency, does not necessarily indicate lower energy poverty, but indicates that the energy poverty is lowered as much as possible taking into account the renewable energy resources used by the country. Hence, the efficiency score may have lower correlation with variables that indicate energy poverty. Still, the correlation should exist. Considering the meaning of the values of the index, the efficiency score should positively corelate with absolute expenditure on electricity and disposable annual household income and negatively corelate with the share of population considering their dwelling as too dark. The created indices should also correlate with each other. Correlation should be positive as for both of the indices worse performing countries should have lower values, while countries with low energy poverty and great progress towards green transition should have values close to 1.

Figure 24 and Table 7 below present the results of the correlation analysis that includes both of the constructed indices (green energy poverty index based on PCA and DEA efficiency scores). *Spearman's ρ* is used to evaluate relationships between the variables. This method was chosen as it does not assume normal distribution and is not as sensitive to outliers. The results of the correlation analysis are as expected – both indices positively correlate with absolute expenditure on electricity and disposable annual household income and negatively corelate with the population considering their dwelling as too dark. The indices also positively correlate with each other. The correlation between green energy poverty index based on PCA and selected variables varies from weak to strong. Correlation between the efficiency scores based on DEA and selected variables vary from moderate to strong. As already mentioned, energy poverty is a complex problem that could be assessed from different angles. Hence, not all variables related to energy poverty are highly correlated, as they examine different sides of this multifaceted problem. Still, as the correlation between the computed indices and the selected variables is as expected, even if it is weak with some variables, it can be stated that the indices are measuring the issue that they should be measuring – energy poverty.

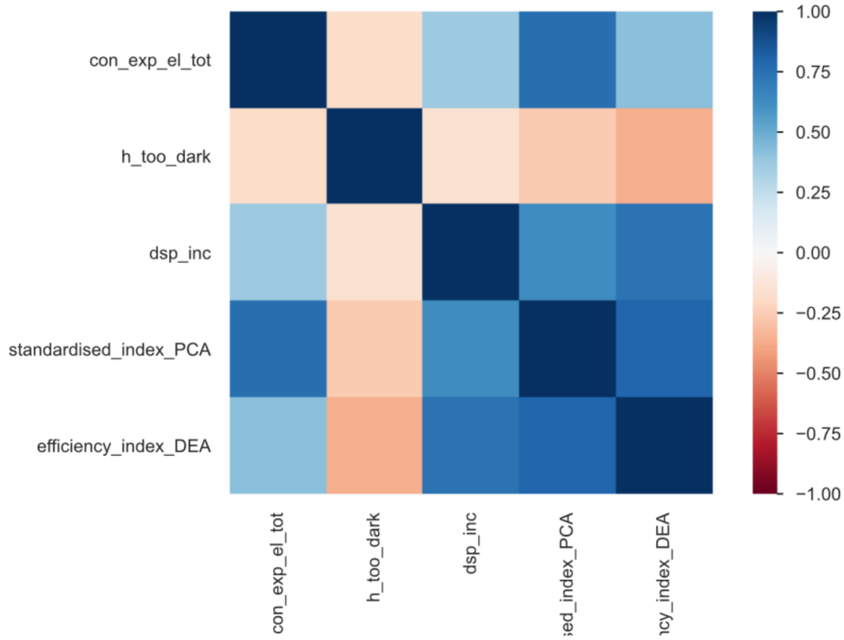

Figure 24. Correlation matrix (Spearman's ρ) between the constructed indices and chosen variables

Correlation with the DEA efficiency scores	Green Energy Poverty Index based on PCA	Efficiency scores based on DEA
Absolute expenditure on electricity	**0.750 (0.000)**	**0.421 (0.000)**
Disposable annual household income	**0.621 (0.000)**	**0.738 (0.000)**
Population considering their dwelling as too dark	**-0.262 (0.000)**	**-0.363 (0.000)**
Green Energy Poverty Index based on PCA	1.000	**0.795 (0.000)**
Efficiency scores based on DEA	**0.795 (0.000)**	1.000

Table 7. Correlation (Spearman's ρ) between the constructed indices and chosen variables
p-values are presented in the brackets
Bolded correlations are significant

The correlation between the indices is also strong and positive (0.795), which shows that they are measuring the same issue. The distribution of the values of both indices is presented in the scatter plot in Figure 25 below. The values for the latest available year (2022) are in dark blue, while past values are presented by light blue points. The scatter plot indicates that the observations with high efficiency scores do not necessarily have high score of green energy poverty index based on PCA, while the observations with high green energy poverty index scores generally also have high efficiency scores. This can be because of a slightly different objectives of the indices. Green energy poverty index based on PCA shows whether the country has low energy poverty and is well progressed in green transition. The efficiency scores, on the other hand, show whether the country has the lowest energy poverty indicators considering the progress it has made towards green transition. Hence, it may be a case that the country still has high energy poverty and is slightly lagging behind in the green transition, so it has relatively low score of green energy poverty index, but it also used its renewable energy sources in a way that focuses on increasing energy affordability and accessibility, so the country is awarded high efficiency score from DEA.

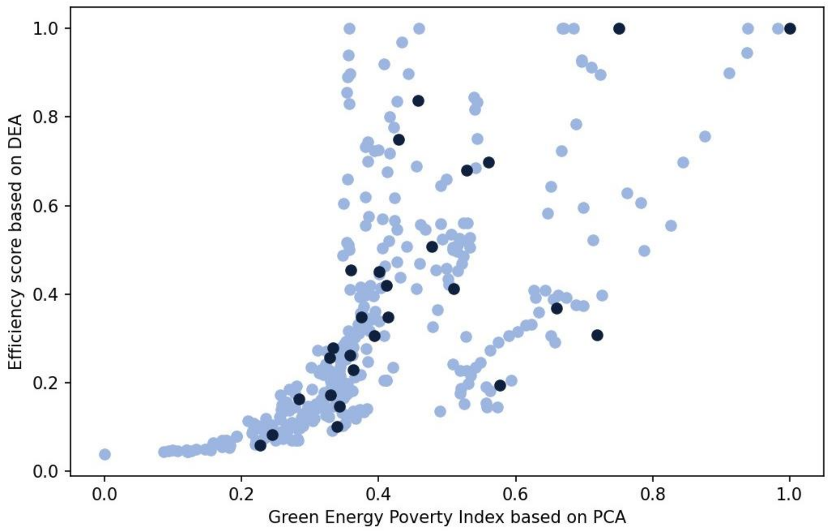

Figure 25. Scatter plot of Green Energy Poverty Index and efficiency scores

In conclusion, correlation analysis confirmed that the two computed indices – Green Energy Poverty Index based on PCA and efficiency scores from DEA – are measuring energy poverty. They correlate with selected variables that are associated with energy poverty. More specifically, both indices positively correlate with absolute expenditure on electricity and disposable annual household income, and negatively correlate with a share of population that considers their dwelling as too dark. The indices also positively correlate with each other, which further confirms that they are both measuring energy poverty.

3.5. Interpretation of the results and policy implications

Previous sections present the results of PCA and DEA and briefly explains how the results can be understood. However, they are more concerned with the statistical meaning of the results and their reliability. This section focuses on the policy implications of the results. It presents how the indices could be used to better understand the connection between energy poverty and green transition. The section also presents a closer look into the related policies in the best and worst performing countries to derive policy suggestions for the countries that wish to improve their situation.

3.5.1. Green Energy Poverty Index based on PCA

The Green Energy Poverty Index based on PCA has values from 0 to 1. The values of the index were standardised, meaning that both 0 and 1 are not theoretical worst and best scenarios, but refer to the lowest and highest existing values of non-standardised index. Index score 0 is given to Bulgaria in 2010, when the country is most energy poor and the least advanced in green transition from the whole dataset. Index score 1 is given to Germany in 2022, when the country has lowest energy poverty indicators and highest indicators of renewable energy use.

This index differs from other energy poverty indices because it considers not only energy poverty indicators, but also use of renewable energy resources. As already discussed in literature review section 1.1.2, if high use of renewable energy resources is followed by proper policies that promote socially inclusive green transition, their use may contribute to higher availability, affordability and efficiency of energy, reducing energy poverty (Dong et al., 2021). Hence, the index takes into account not only the observed energy poverty, but also this potential to reduce energy poverty through green transition.

Index values for all 27 EU Member States for the last year available (2022) can be found in Appendix 3. The values are also illustrated in the map in the Figure 26 below and ranked in Table 8. Countries that are not coloured (beige) are not included in the index. As it can be seen, majority of the countries have low to moderate index scores, as they are coloured light green. This indicates that there is still a lot to be done in these countries to ensure that they progress in green transition in an inclusive way that contributes to lower energy poverty. The best performing country is Germany (in dark green), while worst performing countries are Cyprus, Bulgaria and Greece (light yellow).

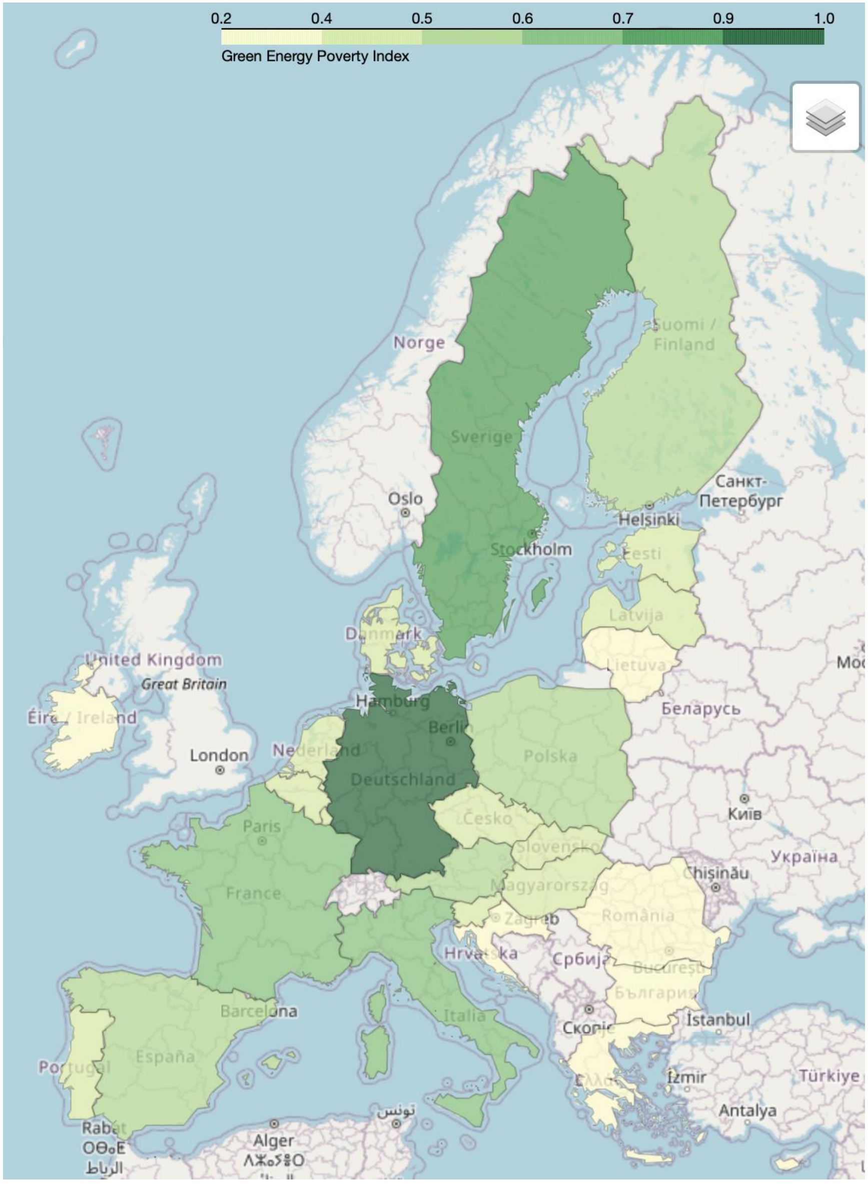

Figure 26. Green Energy Poverty Index based on PCA map for 2022

In general, while most of the countries are moderately to highly energy poor and lagging behind in green transition, the overall situation in European Union is improving. The Figure 27 below shows the change of the index values in EU27 between 2010 and 2022. As it can be seen, EU average is gradually increasing over the years except for a slight decrease from 2021 to 2022. This is not surprising considering different policy measures that have been taken over the years to address energy poverty in the EU and ensure inclusive green transition. These policies are already discussed in section 1.2. Specific measures tackling energy poverty include introducing the concept of energy poverty in EU policy measures in 2009, launching Energy poverty Observatory in 2016,and presenting Commissions recommendation on energy poverty in 2020, to name a few (European Commission, n.d.-a). Most significant policy measures focusing on just and rapid green transition include European Green Deal (European Commission, 2019), and REPowerEU (European Commission, 2022). The slight improvement in the Green Energy Poverty Index shows that the measures taken are, at least to some extent effective.

The slight decrease between 2021 and 2022 can be explained by several challenges that could have contributed to worsening energy poverty faced by the EU Member states in recent years, which are already discussed in section 1.2. For example, COVID-19 pandemic and Russia’s invasion of Ukraine resulted in higher energy prices (Rao, 2022).

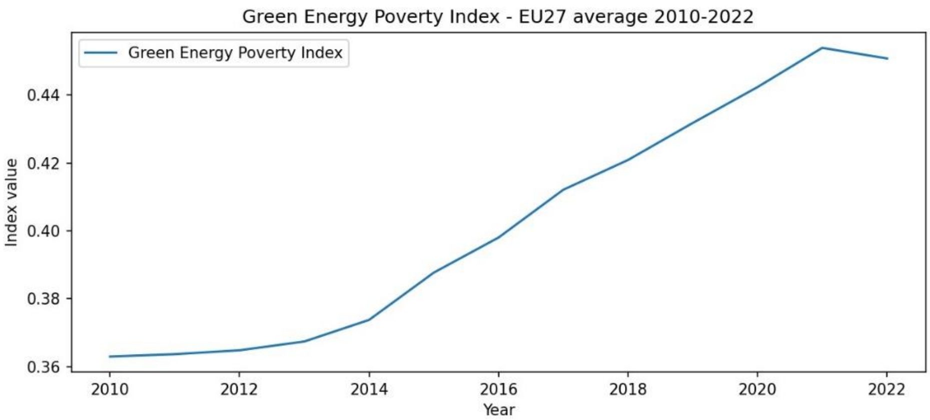

Figure 27. Green Energy Poverty Index in EU27 (average) between 2010 and 2022

The index values can provide guidance for policy makers in decision making process specifically focusing on energy poverty, climate policies and green transition. They are most informative together with other indicators related to energy poverty and green transition as the index highlights the relation between the two issues. For example, if the country has a relatively low value for the index, it can mean that the country is energy poor and lagging in green transition. However, if the indicators related to green transition are examined at the same time and they indicate more progress in green transition than the index, this may also mean that the country is not ensuring that the green transition is inclusive and contributing to lower energy poverty instead of the other way around.

3.5.2. Efficiency scores based on DEA

The efficiency scores from DEA show whether the country achieved the lowest energy poverty possible with the renewable energy resources it has. It means that if the country has effective policies and its green transition is well-thought out to improve affordability and availability of energy, efficiency value for that country is 1 or close to 1. At the same time, if the green transition has been premature and climate polities implemented are not socially inclusive, efficiency score will be low.

Index values for all 27 EU Member States for the last year available (2022) can be found in Appendix 3. The values are also illustrated in the map in the Figure 28 below. The best performing country is Germany (dark green), while worst performing country is Greece (light green).

The Figure 28 shows that majority of the countries are highly or moderately ineffective in using their renewable energy resources to reduce their energy poverty. As known from existing academic literature, increased use of renewable energy resources may positively affect energy poverty if it is paired with policies that ensure that this change in use of energy resources improve availability and affordability of energy (Dong et al., 2021). However, if the green transition is premature and followed by poor climate policy that does not promote social inclusiveness, it may worsen energy poverty (Belaïd, 2022). Hence, low efficiency scores signal that a large share of the EU27 countries may have not properly prepared for the green transition and did not focus on improving efficiency and affordability of energy when increasing the amount of renewable energy resources used.

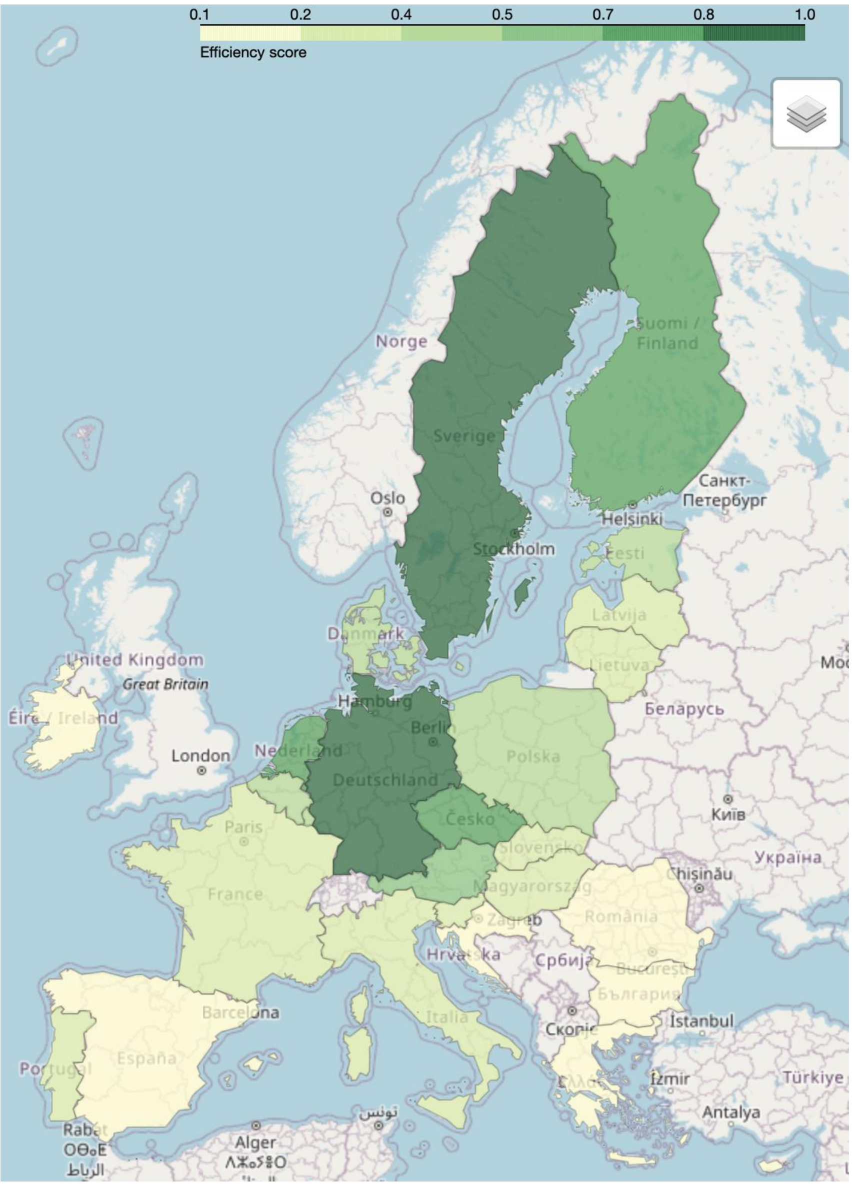

Figure 28. Efficiency score based on DEA map for 2022

The efficiency scores can be used as a great addition to the Green Energy Poverty Index based on PCA that is presented above. While the index shows how well the country is performing considering energy poverty and green transition, the efficiency scores help to understand whether a country is achieving the best possible results with the resources it has. For example, Denmark and the Netherlands have similar values of green energy poverty index, 0.478 and 0.458 respectively. This would mean that they are similarly energy poor and similarly lagging behind in green transition. However, efficiency scores for these countries differ. Denmark efficiency score of 0.507, while the Netherlands has efficiency score of 0.838. This means that while both countries are performing similarly, the Netherlands is using its renewable resources more efficiently to tackle energy poverty. With their use of renewable resources, the Netherlands could improve its energy poverty indicators by 17.2%, while Denmark could have 49.3% lower energy poverty indicators if it had more inclusive green transition policies and focused its use on renewable energy on improving energy affordability and effectiveness.

In general, situation in the EU is slightly improving considering the efficient use of renewable energy with an aim to reduce energy poverty. Figure 29 below shows the average efficiency scores in the EU between 2010 and 2022. While the average scores point to moderate inefficiency in the whole EU, it is also clear that situation is slightly improving. The same is seen from the average Green Energy Poverty Index. The improvement in average efficiency scores further confirms that policies focusing on just green transition, such as Green Deal, are ensuring that green transition consider vulnerable society in groups and that efficiency and affordability of energy is improved when more renewable energy resources are introduced in the energy market.

As with the average Green Energy Poverty Index, there is a decrease in efficiency scores from 2021 to 2022. This can be a result of the already mentioned challenges and higher energy prices. Moreover, climate change resulted in changes in energy demand and reduced ability of the EU Member States to produce energy from renewable energy sources due to harsher weather conditions (Rao, 2022).

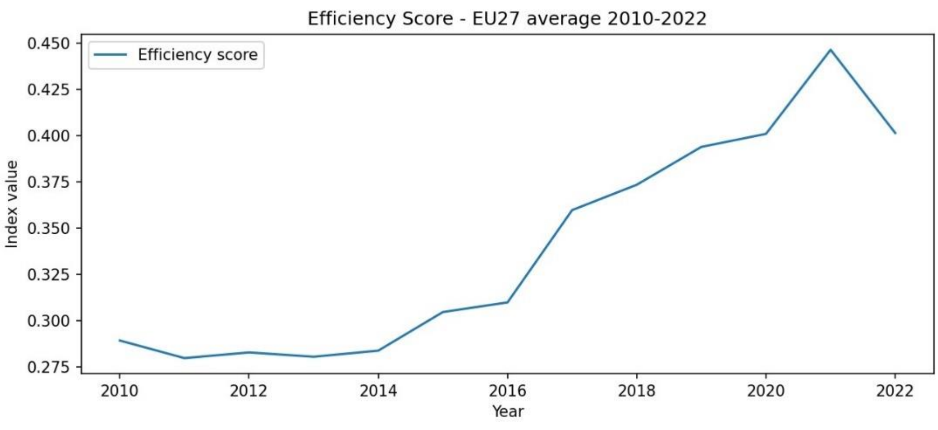

Figure 29. Efficiency scores in EU27 (average) between 2010 and 2022

While discussing the relevance of DEA results for policy making, it is also important to consider the target values that the analysis produces. Target values show what values of all used indicators the country could have if it was more efficient. More specifically, they indicate how much the country could reduce its energy poverty if it used its renewable energy resources with a focus on improving energy availability and affordability. These target values can be informative to policy makers and could be used as realistic targets in action plans for energy poverty reduction. The target values for all countries for 2022 are available in Appendix 5.

3.5.3. Policy implications

The scores of the indices show that majority of the countries in the EU27 are still highly or moderately energy poor and their green transition, evaluated through the use of renewable energy resources, is not as just as envisioned. However, while some countries are still significantly lagging behind, others have managed to progress in green transition and reduce its energy poverty in the process.

The Table 8 below presents the values of the green energy poverty index based on PCA and efficiency scores based on DEA for EU27 countries for 2022. The countries are ranked from the best performing to the most lagging behind. While the ranking of both indices differs due to different objectives of the indices, both of them show that Germany is the best performer, while Greece is the most lagging behind. Closer look to these countries can provide meaningful insights on the possible steps struggling countries could take to improve their energy poverty while progressing in green transition.

	Country	Green Energy Poverty Index (PCA)		Country	Efficiency Score (DEA)
1.	Germany	1.000	1.	Germany	1.000
2.	Sweden	0.751	2.	Sweden	1.000
3.	France	0.719	3.	The Netherlands	0.838
4.	Italy	0.659	4.	Czech Republic	0.750
5.	Spain	0.577	5.	Finland	0.699
6.	Finland	0.561	6.	Austria	0.679
7.	Austria	0.528	7.	Denmark	0.507
8.	Poland	0.509	8.	Luxembourg	0.454
9.	Denmark	0.478	9.	Estonia	0.451
10.	The Netherlands	0.458	10.	Belgium	0.421
11.	Czech Republic	0.429	11.	Poland	0.412
12.	Portugal	0.413	12.	Italy	0.369
13.	Belgium	0.411	13.	Slovenia	0.349
14.	Estonia	0.401	14.	Portugal	0.348
15.	Latvia	0.394	15.	France	0.308
16.	Slovenia	0.375	16.	Latvia	0.307
17.	Hungary	0.363	17.	Lithuania	0.279
18.	Luxembourg	0.359	18.	Slovakia	0.262
19.	Slovakia	0.359	19.	Malta	0.258
20.	Croatia	0.343	20.	Hungary	0.229
21.	Romania	0.339	21.	Spain	0.195
22.	Lithuania	0.333	22.	Ireland	0.172
23.	Ireland	0.329	23.	Cyprus	0.164
24.	Malta	0.328	24.	Croatia	0.148
25.	Cyprus	0.284	25.	Romania	0.101
26.	Bulgaria	0.245	26.	Bulgaria	0.083
27.	Greece	0.226	27.	Greece	0.059

Table 8. EU27 countries ranked by their green energy poverty index and efficiency scores for 2022

The scores of Green Energy Poverty Index indicate that Germany is the least energy poor country with the most progress in green transition. The efficiency score of 1 indicates that the country has also achieved the lowest energy poverty it can achieve given its use of renewable resources. Such good performance can be related to several policy measures that have been taken in Germany before and during the analysed period. Firstly, the government includes energy poverty in a comprehensive set of social policies focusing on poverty in general. For example, basic social support includes energy costs in its consideration of living expenses of a household. Individuals can also apply for government backed loans to cover debts on energy payments so that they would avoid energy disconnection and, in some cases, long-term debts caused by high energy costs may be taken over by the government. The government also may issue subsidies for various improvements that improve energy efficiency. Moreover, research community in the country has been actively focusing on energy poverty since early 2000, with recent work focusing also on the effect that the changes in energy market have on energy poverty. Finally, there are several local level initiatives that tackle energy poverty. These include improving energy efficiency, support for local governments with energy audits, and social tariff implemented by local energy suppliers, among others (EU Energy Poverty Observatory, 2020).

Germany also invests a lot of effort into green transition that is well thought-out and just, which may explain efficiency score of 1. Two main policies that are currently leading green transition in Germany include Climate Action Programme 2030 and Climate Action Act. These policies focus on use of renewable energy resources, improvement of buildings to increase their energy efficiency, and transformation of transport sector. The main goals of these actions include reducing greenhouse emissions by 55 % by 2030, phasing out coal as a resource for energy production, and restructuring transport sector (Die Bundesregierung, n.d.). These policies and the whole energy transition in Germany also heavily focuses on ensuring energy affordability in the process (EU Energy Poverty Observatory, 2020).

The actions presented above may explain high scores Germany received for both of the indices. The policy makers acknowledge energy poverty as a complex problem that is a part of a broader poverty issue and addresses it through a broad comprehensive set of social policies. The issue of energy poverty also has been closely analysed and monitored since early 2000s, which allowed the government to make informed decisions on different measures that may tackle this problem. Moreover, actions for green transition are implemented with an aim to improve energy efficiency and affordability in the country. As already discussed in the Literature review, green transition actions with such focus tend to result in lower energy poverty, among other benefits (Dong et al., 2021).

The worst scores for both computed indices were given to Greece. These scores may be a result of a long-term struggles the country has faced. Energy poverty, as poverty in general, increased in the country after the financial crisis in 2008. The situation has been further worsened by generally low energy efficiency of residential buildings and increasing energy prices. Worsening situation in the country since the financial crisis required a comprehensive policy framework that targets energy poverty and related challenges. However, Greece lacked such framework during the observed time period. The country also lacks systems to accurately measure and monitor energy poverty and assess the effectiveness of different implemented measures that should address energy poverty. Research on energy poverty in Greece is also scarce, which results in a lack of understanding of the factors that contribute to energy poverty in the country and makes it difficult for individuals and policy makers to find relevant and up-to-date information about the issue (EU Energy Poverty Observatory, 2019).

In recent years specific actions were taken to address the challenges that Greece face in tackling energy poverty. Energy Poverty Observatory was established in 2014 to develop research activities on the topic of energy poverty in Greece. The Observatory aims to provide clear framework for measuring and evaluating energy poverty in the country as well as policy measures to tackle it. It strives to also provide clear and up-to-date information for individuals and policy makers about energy poverty, main factors contributing to it, and most important factors to consider when tackling this issue (EU Energy Poverty Observatory, 2019). Between 2010 and 2020 Greece also aimed to introduce specific measures that tackle the causes of energy poverty. These measures focused on improving energy access and efficiency. They included social electricity tariff introduced in 2010, heating oil allowance since 2013, regulatory measures for the protection of energy poor household introduced in 2015, energy efficiency obligation schemes active since 2017, energy upgrade programme for dwelling available since 2020, and scheme to replace heating oil boilers, introduced in 2015 (ONPE, 2021). In 2021 the government introduced an Action Plan to Combat Energy Poverty. This action plan can be seen as a first step towards a comprehensive policy framework that aims to tackle energy poverty in Greece. The main objectives of the action plan include mapping and analysing the households affected by energy poverty, introducing a comprehensive set of policies to tackle energy poverty, and setting up a system for monitoring energy poverty and the policies that should tackle it. Foreseen set of policies tackling energy poverty considers three dimensions – information and training, consumer protection, and development perspective with a focus on improving energy efficiency and increasing use of renewable energy sources (FAO, n.d.). However, the effect of these recent policies is not yet visible, as indicated by the Green Energy Poverty Index and efficiency scores.

Considering the green transition, Greece has recently introduced policies aiming to reduce emission from energy use across different sectors, facilitate the move from traditional energy sources to renewable ones across different sectors, and mitigate negative impact of climate change on individuals and businesses. While there is hope that transition from fossil fuels to renewable energy prices will increase availability of reliable and affordable energy, clear measures to mitigate potential negative effects from rapid green transition are missing from the current policy framework (Leidecker et al., n.d.). Hence, increased energy prices, consequently increased prices of some goods, especially those intensive in carbon, and other challenges may further worsen energy poverty trap among Greek households.

The political and economic situation of Greece may explain the low Green Energy Poverty Index and efficiency scores. As the research on energy poverty in a country is scarce, it is difficult to understand the underlying causes of energy poverty and identify effective solutions to tackle them. This lack of knowledge may explain why previously implemented policy measures tackling energy poverty seems to be ineffective. The measures aiming to improve energy efficiency seem to lack focus on the poorest households, while the financial support for debts related to energy costs tackle the results of energy poverty, not its causes. A comprehensive policy framework focusing on energy poverty has been introduced only in 2021, which means that its potential effect is not yet visible. However, during the majority of the observed time period such policy framework was not in place. Moreover, while the country introduced several actions to progress in green transition, these actions lack focus on improving availability and affordability of energy, which means that they may exacerbate energy poverty in the country.

Data envelopment analysis produced target values for inefficient countries, including Greece. These values show how much a country could decrease its energy poverty if the use of its renewable energy resources would be accompanied by the policies that ensure affordability, efficiency and availability of energy during the green transition. These values are presented in the Table 9 below.

	Current values	Target
Inability to keep adequately warm	18.700	1.108
Arrears on utility bills	34.100	2.020
Use of renewables for electricity	23742.000	39549.446
Use of renewables for heating and cooling	1666.200	3718.718
Share of energy from renewable resources	22.678	22.678

Table 9. Current and target values of the chosen indicators for Greece

The principal component analysis resulted in a formula (19) that can be used to easily calculate non-scaled values of the green energy poverty index, if the values of the needed indicators are available. The indicators should be scaled using formula (17) and median absolute deviations and medians of each indicator provided in Table 4. The formula can be used to calculate the index value for Greece with its target values to better understand how much the country could improve with better thought-out green transition policies. The scaled target values are presented in the Table 10.

	Target values	Scaled target values
Inability to keep adequately warm	1.108	-1.349
Arrears on utility bills	2.020	-1.416
Use of renewables for electricity	39549.446	2.653
Use of renewables for heating and cooling	3718.718	1.486
Share of energy from renewable resources	22.678	0.690

Table 10. Non-scaled and scaled target values of the chosen indicators for Greece

The Green Energy Poverty Index for Greece with its target values is calculated as follows:

$$GEPI = 2.653 + 0.949 * 1.486 + 0.580 * 0.690 - 0.826 * (-1.349) - (-1.416) = 6.994$$

Table 12 in Appendix 3 presents non-scaled values for the Green Energy Poverty Index for 2022. As it can be seen, value for Greece is -5.594. The index score calculated with the target values would place Greece between Austria and Finland, making it 7^{th} best performing country in 2022, as seen from Table 8 above. This shows how much progress Greece could make with the right actions and policies in place.

To sum up, the analysis of the situation of best and worst performing countries, Germany and Greece respectively, revealed which policy measures may help countries progress in their green transition in a way that also results in lower energy poverty. Three most important actions were identified. Firstly, the issue of energy poverty should be thoroughly researched to better understand its underlying causes and the most suitable solutions. Secondly, policy actions targeting energy poverty should be a part of a comprehensive long-term policy framework that focuses on both energy poverty and broader related socioeconomic challenges, such as overall poverty. Finally, actions towards green transition, such as increased use of renewable energy sources, should be introduced with an aim to also improve energy efficiency and affordability.

3.6. Limitations of the study and topics for future research

This study presents valuable findings that can inform policy decision making process focusing on energy poverty in the light of green transition. However, the study has some limitations. These limitations mainly stem from the complexity of the issue in question – energy poverty – and the availability of the data. The challenges include complex relationship between indicators used in the analysis and use of created indices for the future computations.

Firstly, it is important to remember that while the Green Energy Poverty Index and efficiency scores present the progress made towards green transition and the effect it has towards energy poverty, the values of the indices should be considered critically. While the connection between energy poverty and progress towards green transition exist, there may be other factors affecting energy poverty in a country that are not accounted for in the analysis. For example, as already discussed in literature review section 1.1.2, energy poverty in the light of green transition is strongly affected by the climate policies and political actions leading to green transition (Belaïd, 2022). As quality and nature of policies implemented in the EU Member States are not included in the index, some of the specific connections may be overlooked. DEA efficiency scores and target values partly account for this limitation. The efficiency scores show whether the country achieved the lowest energy poverty possible with the renewable energy resources it has. Still, to better understand how the desired effect can be achieved, inclusion of quantitative policy analysis is required. Therefore, further analysis is needed to better understand the relation between energy poverty and green transition as well as the factors influencing this relation.

Second important limitation is concerned specifically with the index computed by PCA. The aim of the index to consider energy poverty in the light of green transition. Hence, the index evaluates if the country is energy poor and at the same time lags behind in green transition. The indicators of energy poverty have positive weights in the index, while the indicators for the use of renewable energy resource use have negative weights. In general, this satisfies the objective of the index to evaluate both energy poverty and progress in green transition. It also makes the index easy to interpret. However, this also may result in situations where energy poverty is not properly presented by the index because of high use of renewable energy resources that masks the real level energy poverty. The index computation allows to compensate for high energy poverty with high use of renewable energy. Hence, best performing countries have to be analysed closely to make sure that it is not a case.

Third important limitation is the sensitivity of the index to the size of a country. This is specific for the index created using PCA. As already mentioned, the index uses absolute values for renewable energy use. This means that the values for these indicators are likely larger for the countries that are larger and have more residents. Hence, in some cases, smaller countries may seem to be performing worse than they actually perform compared with the larger countries. This could be addressed by using the share of renewable energy used from the total consumption for specific purpose, namely electricity or heating and cooling. However, such data is not available for all countries for the whole time period analysed. The number of missing values was too great to employ interpolation without compromising the results of the analysis. Therefore, this limitation needs to be kept in mind when examining the values of the Green Energy Poverty Index.

Final important limitation stems from the index computation. Both Green Energy Poverty Index based on PCA and efficiency scores from DEA were calculated using data from 2010 to 2022. Index based on PCA is also normalised, so the normalised values can be only compared in this dataset. If more data is added, the standardisation has to be repeated including both the already used data and new data. This may change the values of the standardised index for the observations already included in this study. This does not compromise the reliability of the index values, especially non-scaled values. However, it is an important point to keep in mind if the index is used in the future. Considering the replication of data envelopment analysis, it is important to keep in mind that if the past values are not used in DEA, the results may be more optimistic. Moreover, the analysis of the progress of the country over the years would also be slightly less straightforward.

These limitations highlight the topics that could be a focus of future research on energy poverty and green transition. These topics include:

- Complexity of the relationship between energy poverty and green transition;
- Factors influencing complex relationship between energy poverty and green transition;
- Availability of data for green transition and potential proxies.

Conclusions

1. Energy poverty is highly discussed issue. However, the consensus on definition of energy poverty and the most suitable method to measure it is lacking. The challenge to find a universal definition and measurement is also complicated by frequent changes in the energy market that increase its uncertainty. For example, the green transition in the EU is slowly changing energy market in the EU and the energy consumption habits of the EU citizens. However, this is rarely acknowledged and taken into account when assessing energy poverty. While there are some attempts to better capture the relationship between energy poverty and green transition, they often do not have a strong mathematical background. Hence, there is a clear need for an energy poverty assessment that sufficiently accounts for the green transition. This study fills this gap by presenting Green Energy Index based on PCA and efficiency scores derived from DEA analysis of the energy poverty and green transition nexus.
2. The study focused on five indicators: inability to keep adequately warm, arrears on utility bills, use of renewables for electricity, use of renewables for heating and cooling, and share of energy from renewable resources. First two indicators are used to assess energy poverty and are recommended for this aim by the EU Energy Poverty Advisory Hub, which makes them most suitable indicators of energy poverty in the EU context. Indicators of the use of renewable energy resources were chosen as proxies for the progress in green transition. While green transition is significantly more complex than the switch from traditional energy sources to renewable sources, these indicators were seen as most suitable as the focus is on energy market.
3. PCA was used to create a Green Energy Poverty Index. This method of analysis helps simplify the complex dataset and can be used to create an index that is easy to interpret. The index had the following weights for chosen indicators:
    - Inability to keep adequately warm: -0.826;
    - Arrears on utility bills: -1.000;
    - Use of renewables for electricity: 1.000;
    - Use of renewables for heating and cooling: 0.949;
    - Share of energy from renewable resources: 0.580.

   The index values were standardised. The standardised values range from 0 to 1. 0 indicates that country is highly energy poor and is lagging behind in the green transition, while 1 means that the country has low energy poverty and is well progressed towards green transition.
4. DEA was used to calculate the efficiency scores for the 27 EU Member States. This analysis presents the current situation and suggests possible improvements. Values of the efficiency index may range from 0 to 1. Countries that receive value of 1 are considered as efficient. It means that they are moving towards the green transition in a way that is inclusive and tackles energy poverty. Countries that receive the value lower than 1 are using their renewable energy resources in a way that is not necessarily focused on increasing affordability and accessibility of energy and could have lower energy poverty considering the amount of renewable energy resources they use. The analysis also provides the target values for energy poverty indicators, meaning values that could be achieved with the same amount of renewable energy used if there was more focus on improving efficiency and affordability of energy.
5. The created indices are validated through correlation analysis. Correlation analysis includes variables that are closely related to energy poverty. These variables are absolute expenditure on electricity, disposable annual household income, and population considering their dwelling as too dark. The correlation analysis revealed that constructed indices are correlating with these

variables in a way that was expected. The newly computed indices also strongly correlate to each other. Hence, it is clear that the newly constructed indices are measuring the problem for which they are constructed – energy poverty.

6. The study presents two indices that aim to assess energy poverty while taking into account green transition. Index based on PCA presents the current situation and provides values that facilitates cross-country comparison in the analysed time period. The standardised Green Energy Poverty Index ranges from 0, high energy poverty and lagging in green transition, to 1, low energy poverty and great progress towards green transition. The analysis revealed that majority of the countries have scores lower than 0.5, meaning they are moderately to highly energy poor and lagging behind in green transition. Considering 2022 values, the best performing country is Germany, while the worst performing country is Greece.

   DEA produced efficiency scores are more related to individual observations and provides clear targets for minimising energy poverty. Efficiency can have values between 0 and 1. Value of 1 indicates that a country is using its renewable energy resources in a way that is inclusive and contributes to reducing energy poverty. Values lower than 1 indicate that a country could achieve lower energy poverty if its green transition progress, namely increase in renewable energy use, would be more focused on social inclusivity, energy affordability, and energy accessibility. Considering 2022 values of the index, two countries can be considered as efficient – Germany and Sweden. The most inefficient country in 2022 is Greece.

   Both of the indices presented in the study should be considered together in policy discussions. The PCA index presents an overall situation and facilitates the comparison over the years and between the countries. The DEA index allows to move from the general overview and comparison and presents not only the current situation, but also realistic targets for energy poverty that could be achieved by a country if its current consumption of renewable energy was more inclusive and oriented towards increasing energy affordability and efficiency.

   A closer look into the policies implemented by the best performing and worst performing countries in 2022, Germany and Greece respectively, provides clear guidelines on how EU Member States could effectively tackle energy poverty and progress in just green transition. Firstly, there should be an active research community focusing on energy poverty to ensure that this problem is well understood and decisions taken to address it are well informed. Secondly, as energy poverty is a complex problem it should be tackled by a comprehensive set of policies that are a part of a larger framework focusing on other related socio-economic issues, such as general poverty. Finally, when progressing in green transition, including increasing use of renewable energy resources, countries should ensure that this progress also improves energy efficiency and affordability.

7. While the study contributes to the knowledge gap on the relationship between energy poverty and green transition, the complexity of the issue should be further researched. Hence, the future research could focus on better understanding the specificities of the relationship between the energy poverty and green transition that the study highlighted, focusing on other factors that may be influencing this relationship. The study also highlighted a clear limitation in the availability of data to measure both energy poverty and green transition. Hence, future research could explore suitable proxies for these issues.

List of references

1. Adom, P. K., Amuakwa-Mensah, F., Agradi, M. P., & Nsabimana, A. (2021). Energy poverty, development outcomes, and transition to green energy. *Renewable Energy*, *178*, 1337–1352. https://doi.org/10.1016/j.renene.2021.06.120
2. Bajomi, A. Z., Feldmár, N., & Tirado-Herrero, S. (2021). Will Plans to Ease Energy Poverty Go Up in Smoke? Assessing the Hungarian NECP through the Lens of Solid Fuel Users' Vulnerabilities. *Sustainability*, *13*(23), Article 23. https://doi.org/10.3390/su132313047
3. Banker, R., Chames, A., Cooper, W. W., Swarts, J., & Thomas, D. A. (n.d.). *An introduction to data envelopment analysis with some of its models and their uses*.
4. Belaïd, F. (2022). Implications of poorly designed climate policy on energy poverty: Global reflections on the current surge in energy prices. *Energy Research & Social Science*, *92*, 102790. https://doi.org/10.1016/j.erss.2022.102790
5. Boardman, B. (1991). *Fuel poverty: From cold homes to affordable warmth*. Pinter Pub Limited. https://cir.nii.ac.jp/crid/1130000794123994880
6. Bonatz, N., Guo, R., Wu, W., & Liu, L. (2019). A comparative study of the interlinkages between energy poverty and low carbon development in China and Germany by developing an energy poverty index. *Energy and Buildings*, *183*, 817–831. https://doi.org/10.1016/j.enbuild.2018.09.042
7. Bowlin, W. (1998). Measuring Performance: An Introduction to Data Envelopment Analysis (DEA). *The Journal of Cost Analysis*, *15*. https://doi.org/10.1080/08823871.1998.10462318
8. Cook, W. D., Tone, K., & Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega*, *44*, 1–4. https://doi.org/10.1016/j.omega.2013.09.004
9. Cotte Poveda, A. (2012). Estimating Effectiveness of the Control of Violence and Socioeconomic Development in Colombia: An Application of Dynamic Data Envelopment Analysis and Data Panel Approach. *Social Indicators Research*, *105*(3), 343–366. https://doi.org/10.1007/s11205-010-9772-7
10. Culver, L. C. (2017). Energy Poverty: What You Measure Matters. *Pre-Symposium White Paper for Reducing Energy Poverty With Natural Gas: Changing Political, Business, and Technology Paradigms (California, CA: Stanford University)*.
11. Die Bundesregierung. (n.d.). *What the government is doing for the climate*. Website of the Federal Government | Bundesregierung. Retrieved 24 May 2024, from https://www.bundesregierung.de/breg-en/issues/climate-action/government-climate-policy-1779414
12. Dong, K., Ren, X., & Zhao, J. (2021). How does low-carbon energy transition alleviate energy poverty in China? A nonparametric panel causality analysis. *Energy Economics*, *103*, 105620. https://doi.org/10.1016/j.eneco.2021.105620
13. 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
14. Ehsanullah, S., Tran, Q. H., Sadiq, M., Bashir, S., Mohsin, M., & Iram, R. (2021). How energy insecurity leads to energy poverty? Do environmental consideration and climate change concerns matters. *Environmental Science and Pollution Research*, *28*(39), 55041–55052. https://doi.org/10.1007/s11356-021-14415-2

15. EU Energy Poverty Observatory. (2019). *EU Energy Poverty Observatory Case Study: Energy Poverty Observatory in Greece—A tool for the effective measurement and monitoring of energy poverty*. https://energy-poverty.ec.europa.eu/document/download/d059db54-d79b-4f27-a05c-e840916e8ffb_en?filename=case_study_epov_greece.pdf
16. EU Energy Poverty Observatory. (2020). *Member State Report—Germany*. https://energy-poverty.ec.europa.eu/system/files/2021-10/epov_member_state_report_-_germany.pdf
17. European Commission. (n.d.-a). *Energy poverty*. Retrieved 7 April 2024, from https://energy.ec.europa.eu/topics/markets-and-consumers/energy-consumer-rights/energy-poverty_en
18. European Commission. (n.d.-b). *The European Pillar of Social Rights in 20 principles—Employment, Social Affairs & Inclusion—European Commission*. Retrieved 6 April 2024, from https://ec.europa.eu/social/main.jsp?catId=1606&langId=en
19. European Commission. (2019). *Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU*. http://data.europa.eu/eli/dir/2019/944/oj/eng
20. European Commission. (2020). Commission Recommendation (EU) 2020/1563 of 14 October 2020 on energy poverty. In *OJ L* (Vol. 357). http://data.europa.eu/eli/reco/2020/1563/oj/eng
21. European Commission. (2021). *Proposal for a directive of the European Parliament and of the Counil on energy efficiency (recast)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0558
22. European Commission. (2022, May 18). *REPowerEU*. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/repowereu-affordable-secure-and-sustainable-energy-europe_en
23. European Council. (2023). *Energy crisis: Three EU-coordinated measures to cut down bills*. https://www.consilium.europa.eu/en/infographics/eu-measures-to-cut-down-energy-bills/
24. European Parliament, & The Council of the European Union. (2018). *Regulation (EU) 2018/1999 of the European Parliament and of the Council of 11 December 2018 on the Governance of the Energy Union and Climate Action, amending Regulations (EC) No 663/2009 and (EC) No 715/2009 of the European Parliament and of the Council, Directives 94/22/EC, 98/70/EC, 2009/31/EC, 2009/73/EC, 2010/31/EU, 2012/27/EU and 2013/30/EU of the European Parliament and of the Council, Council Directives 2009/119/EC and (EU) 2015/652 and repealing Regulation (EU) No 525/2013 of the European Parliament and of the Council*. http://data.europa.eu/eli/reg/2018/1999/2021-07-29/eng
25. Eurostat. (n.d.). *Inability to keep home adequately warm*. Eurostat. Retrieved 6 April 2024, from https://ec.europa.eu/eurostat/databrowser/view/ILC_MDES01__custom_6198699/default/table?lang=en
26. Eurostat. (2009). *What can be learned from deprivation indicators in Europe? - Products Statistical working papers—Eurostat*. Eurostat. https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/ks-ra-09-007
27. Faiella, I., & Lavecchia, L. (2021). Energy poverty. How can you fight it, if you can't measure it? *Energy and Buildings*, *233*, 110692. https://doi.org/10.1016/j.enbuild.2020.110692

28. FAO. (n.d.). *Greece—Action Plan to Combat Energy Poverty*. https://www.fao.org/faolex/results/details/en/c/LEX-FAOC211210/
29. Golany, B., & Roll, Y. (1989). An application procedure for DEA. *Omega*, *17*(3), 237–250. https://doi.org/10.1016/0305-0483(89)90029-7
30. Gouveia, J. P., Bessa, S., Palma, P., Mahoney, K., & Sequeira, M. (2023). *Energy Poverty National Indicators Uncovering New Possibilities for Expanded Knowledge*. https://energy-poverty.ec.europa.eu/document/download/b00326ad-da4b-43cd-a99e-574eb587fce7_en?filename=EPAH2023_2nd%20Indicators%20Report_Final_0.pdf
31. Gupta, S., Gupta, E., & Sarangi, G. K. (2020). Household Energy Poverty Index for India: An analysis of inter-state differences. *Energy Policy*, *144*, 111592. https://doi.org/10.1016/j.enpol.2020.111592
32. Gwatkin, D., Rutstein, S., Johnson, K., RP, P., & Wagstaff, A. (2000). Socio-economic differences in health, nutrition, and population in Burkina Faso. In *The Journal of American Medical Association (JAMA)* (Vol. 298). https://doi.org/10.1001/jama.298.16.1943
33. Hasheminasab, H., Streimikiene, D., & Pishahang, M. (2023). A novel energy poverty evaluation: Study of the European Union countries. *Energy*, *264*, 126157. https://doi.org/10.1016/j.energy.2022.126157
34. Herrero, S. T. (2017). Energy poverty indicators: A critical review of methods. *Indoor and Built Environment*, *26*(7), 1018–1031. https://doi.org/10.1177/1420326X17718054
35. Hoyer, W., & Küüsvek, A. (2022). *A green transition is vital to solving Europe's energy problems*. European Investment Bank. https://www.eib.org/en/stories/green-transition-energy
36. Hubert, M., Debruyne, M., & Rousseeuw, P. J. (2018). Minimum covariance determinant and extensions. *WIREs Computational Statistics*, *10*(3), e1421. https://doi.org/10.1002/wics.1421
37. Hussain, S. A., Razi, F., Hewage, K., & Sadiq, R. (2023). The perspective of energy poverty and 1st energy crisis of green transition. *Energy*, *275*, 127487. https://doi.org/10.1016/j.energy.2023.127487
38. IAEA, U. Department of Economic and Social Affairs, IEA, Eurostat, & European Environment Agency. (2005). *Energy indicators for sustainable development: Guidelines and Methodologies*.
39. IEA. (2012). *World Energy Outlook 2012*. OECD Publishing. https://doi.org/10.1787/weo-2012-en
40. IEA. (2015). *World Energy Outlook 2015: Methodology for Energy Access Analysis* (pp. 1–8) [Tech. rep.]. International Energy Agency.
41. Jayasinghe, M., Selvanathan, E. A., & Selvanathan, S. (2021). Energy poverty in Sri Lanka. *Energy Economics*, *101*, 105450. https://doi.org/10.1016/j.eneco.2021.105450
42. Kahouli, S., & Okushima, S. (2021). Regional energy poverty reevaluated: A direct measurement approach applied to France and Japan. *Energy Economics*, *102*, 105491. https://doi.org/10.1016/j.eneco.2021.105491
43. Kelly, J. A., Clinch, J. P., Kelleher, L., & Shahab, S. (2020). Enabling a just transition: A composite indicator for assessing home-heating energy-poverty risk and the impact of environmental policy measures. *Energy Policy*, *146*, 111791. https://doi.org/10.1016/j.enpol.2020.111791
44. Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, *72*(6), 558–569. https://doi.org/10.4097/kja.19087

45. Krishnan, V. (2010). *Constructing an Area-based Socioeconomic Index: A Principal Components Analysis Approach **. https://www.semanticscholar.org/paper/Constructing-an-Area-based-Socioeconomic-Index-%3A-A-Krishnan/b1152cdc9e9ca217434db492a9e1c4ba519640ce
46. Lai, D. (2003). Principal Component Analysis on Human Development Indicators of China. *Social Indicators Research*, *61*(3), 319–330. https://doi.org/10.1023/A:1021951302937
47. Lan, J., Khan, S. U., Sadiq, M., Chien, F., & Baloch, Z. A. (2022). Evaluating energy poverty and its effects using multi-dimensional based DEA-like mathematical composite indicator approach: Findings from Asia. *Energy Policy*, *165*, 112933. https://doi.org/10.1016/j.enpol.2022.112933
48. Leidecker, T., Bulman, T., Levin, I., & Blake, H. (n.d.). *Transitioning to a green economy in Greece* (OECD Economics Department Working Papers 1757). OECD Publishing. https://www.iobe-ec-conference.gr/2021/Factsheet_on_Green_Policies_Sep2021.pdf
49. Li, W., Chien, F., Hsu, C. C., Zhang, Y., Nawaz, M. A., Iqbal, S., & Mohsin, M. (2021). *Nexus between energy poverty and energy efficiency: Estimating the long-run dynamics*.
50. Liang, Q., & Asuka, J. (2022). A multidimensional energy poverty measurement in China—Based on the entropy method. *Energy for Sustainable Development*, *71*, 554–567. https://doi.org/10.1016/j.esd.2022.11.005
51. Makowski, D., Ben-Shachar, M., Patil, I., & Lüdecke, D. (2020). Methods and Algorithms for Correlation Analysis in R. *Journal of Open Source Software*, *5*(51), 2306. https://doi.org/10.21105/joss.02306
52. Mariano, E. B., Sobreiro, V. A., & Rebelatto, D. A. do N. (2015). Human development and data envelopment analysis: A structured literature review. *Omega*, *54*, 33–49. https://doi.org/10.1016/j.omega.2015.01.002
53. 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), Article 5. https://doi.org/10.3390/su8050483
54. Mirza, B., & Szirmai, A. (2010). *Towards a new measurement of energy poverty: A cross-community analysis of rural Pakistan*. http://collections.unu.edu/view/UNU:275
55. Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005). *Tools for Composite Indicators Building*.
56. Nisa, K., Herawati, N., Setiawan, E., & Nusyirwan. (2006). *Robust Principal Component Analysis Using Minimum Covariance Determinant Estimator* [International Conference on Mathematics and Natural Sciences (ICMNS)]. http://repository.lppm.unila.ac.id/6621/1/Nisa%20ICMN06.pdf
57. 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
58. OECD. (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/economics/handbook-on-constructing-composite-indicators-methodology-and-user-guide_9789264043466-en
59. Okushima, S. (2017). Gauging energy poverty: A multidimensional approach. *Energy*, *137*, 1159–1166. https://doi.org/10.1016/j.energy.2017.05.137

60. ONPE. (2021). *Context and schemes to tackle energy poverty in Europe: Greece—Background and governance*. https://onpe.org/sites/default/files/2024-02/ONPE%20Greece%20EN.pdf

61. Pachauri, S., & Spreng, D. (2011). Measuring and monitoring energy poverty. *Energy Policy*, *39*(12), 7497–7504. https://doi.org/10.1016/j.enpol.2011.07.008

62. Pelz, S., Pachauri, S., & Groh, S. (2018). A critical review of modern approaches for multidimensional energy poverty measurement. *WIREs Energy and Environment*, *7*(6), e304. https://doi.org/10.1002/wene.304

63. Rademaekers, K., Yearwood, J., Ferreira, A., Pye, S., Hamilton, I., Agnolucci, P., Grover, D., Karásek, J., & Anisimova, N. (2016). *Selecting Indicators to Measure Energy Poverty*.

64. Rao, D. (2022, September 25). *Europe's energy crisis, explained*. Theweek. https://theweek.com/climate-change/1016912/climate-changes-role-in-europes-energy-crisis

65. Rousseeuw, P. J., & Driessen, K. V. (1999). *A Fast Algorithm for the Minimum Covariance Determinant Estimator*.

66. Ruiz-Rivas, U., Tahri, Y., Arjona, M. M., Chinchilla, M., Castaño-Rosa, R., & Martínez-Crespo, J. (2022). Energy Poverty in Developing Regions: Strategies, Indicators, Needs, and Technological Solutions. In C. Rubio-Bellido & J. Solis-Guzman (Eds.), *Energy Poverty Alleviation: New Approaches and Contexts* (pp. 17–39). Springer International Publishing. https://doi.org/10.1007/978-3-030-91084-6_2

67. Sareen, S., Thomson, H., Tirado Herrero, S., Gouveia, J. P., Lippert, I., & Lis, A. (2020). European energy poverty metrics: Scales, prospects and limits. *Global Transitions*, *2*, 26–36. https://doi.org/10.1016/j.glt.2020.01.003

68. Siksnelyte-Butkiene, I., Streimikiene, D., Lekavicius, V., & Balezentis, T. (2021). Energy poverty indicators: A systematic literature review and comprehensive analysis of integrity. *Sustainable Cities and Society*, *67*, 102756. https://doi.org/10.1016/j.scs.2021.102756

69. 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

70. Sokołowski, J., Kiełczewska, A., & Lewandowski, P. (2019). *Defining and measuring energy poverty in Poland*.

71. Sokołowski, J., Lewandowski, P., Kiełczewska, A., & Bouzarovski, S. (2020). A multidimensional index to measure energy poverty: The Polish case. *Energy Sources, Part B: Economics, Planning, and Policy*, *15*(2), 92–112. https://doi.org/10.1080/15567249.2020.1742817

72. Sovacool, B. K. (2014). Defining, measuring, and tackling energy poverty. In *Energy poverty: Global challenges and local solutions* (2nd ed., pp. 21–53). https://books.google.lt/books?hl=en&lr=&id=HIJ1BQAAQBAJ&oi=fnd&pg=PA21&dq=50.%09Sovacool,+B.+K.+(2014).+Defining,+measuring,+and+tackling+energy+poverty.+Energy+poverty:+Global+challenges+and+local+solutions,+2,+21-53.&ots=A6GM7O7UqT&sig=VOnt1Qau9OSwWDpfecV4o8xqjZU&redir_esc=y#v=onepage&q&f=false

73. Streimikiene, D., Lekavičius, V., Baležentis, T., Kyriakopoulos, G., & Abrhám, J. (2020). Climate Change Mitigation Policies Targeting Households and Addressing Energy Poverty in European Union. *Energies*, *13*(13), 3389.

74. The R Foundation. (n.d.). *R: What is R?* Retrieved 15 April 2024, from https://www.r-project.org/about.html
75. 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
76. 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
77. UNDP. (2010). *Human Development Report 2010—The Real Wealth of Nations: Pathways to Human Development*. UNDP. https://hdr.undp.org/system/files/documents/human-development-report-2010-complete-english.human-development-report-2010-complete-english
78. *United Nations Millennium Development Goals*. (n.d.). United Nations. Retrieved 6 April 2024, from https://www.un.org/millenniumgoals/bkgd.shtml
79. Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning*, *21*(6), 459–468. https://doi.org/10.1093/heapol/czl029
80. Widuto, A. (2023). *Energy poverty in the EU | Think Tank | European Parliament*. https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2022)733583
81. Zhuang, H., Wang, H., & Ji, Z. (2022). findPC: An R package to automatically select the number of principal components in single-cell analysis. *Bioinformatics*, *38*(10), 2949–2951. https://doi.org/10.1093/bioinformatics/btac235

Appendices

Appendix 1. Literature review matrix

Source	Issue addressed	Sample	Method	Measurement	Indicators
Boardman, 1991	There is a difference between poverty and fuel poverty, and it needs to be addressed differently. However, for the fuel poverty to be effectively addressed, there is a need to find a way to measure it.	-	More than 10% household income spent on energy needs is considered as energy poor	10% indicator	Household spending on energy exceeds 10% its income
Mirza & Szirmai, 2010	Authors developed a complex energy poverty index using data from Pakistan to measure energy poverty specifically among rural households.	Pakistan	The authors combined an average of energy shortfall, index of energy inconvenience (combined from unweighted indices and then weighing their combinations by the share of the energy source use in the whole energy use of a household), and percentage of the minimum basic level of energy consumption.	Energy Poverty Index (EPI)	Data collected through the survey specifically designed for the study
Nussbaumer et al., 2012	While the most existing indicators focus on energy access or development level related to energy, the authors present an index assessing	Selected African countries	Methodology is derived from Oxford Poverty and Human Development Initiative, so there is a focus on absence of opportunities. The index measures energy poverty	Multidimensional Energy Poverty Index (MEPI)	Type of cooking fuel Food cooked on store or open fire if using any fuel beside electricity, LPG, natural gas, or biogas Has access to electricity

	lack of access to modern energy sources.		analysing *d* variables in a population of *x* individuals. The indicators are weighted unevenly with the weighting vector.		Has a fridge Has a radio OR television Has a phone land line OR a mobile phone
Thomson & Snell, 2013	The authors address the knowledge gap about fuel poverty in the EY. Using EU-SILC survey data they examine fuel poverty in the EU in the contex of rising fuel prices and accession to the EU of several former social states.	EU	The authors choose three proxies to estimate fuel poverty and create logistic regression models to determine what variables contribute to higher fuel poverty.	Fuel poverty proxy indicators	Ability to pay to keep the home adequately warm Arrears on utility bills, and the presence of a leaking roof Damp walls or rotten windows
IEA, 2015	There was a need to look into energy poverty and development in a comprehensive manner	-	The index was constructed by calculating an average of three normalised components (weighted equally).	Energy Development Index (EDI)	Per capita commercial energy consumption Share of commercial energy in total final energy use Electrification rate
Dubois & Meier, 2016	The authors present an analytical framework to assess energy poverty across the EU at the macro-scale. The framework is based on energy services deprivation and, secondly, on analysis of energy inequality, taking into account that different households may be differently	EU	The authors focus on macro level indicators. Three dimensions are used: ability to keep home adequately warm, energy affordability and energy efficiency. Second and third dimensions are indices made from 3 different indicators that are equally weighted. All three	Energy services deprivation indicator	Ability to keep home adequately warm Ability of the household to purchase necessary quantity of energy without suffering undue financial hardship (3 indicators) Energy efficiency (3 indicators)

	affected by deprivation of energy services.		dimensions are weighted equally to derive energy services deprivation indicator.		
Snell et al., 2015	Authors consider relationship between fuel poverty, disability and policy changes.	UK	Authors use two 10% measures of fuel poverty with two classifications of income (full and basic income). They also use Low Income High Costs indicator for fuel poverty where to be considered as fuel poor households have required costs above average or their income after spending on fuel are below poverty line.	Measurement for fuel poverty	Data from EHS (cross-sectional study on people's housing circumstances and energy efficiency of housing)
Maxim et al., 2016	Energy poverty is a complex issue that requires a complex measurement. The paper offers an improved energy poverty index focusing on the EU to better assess this issue.	The EU	The weights were assigned following Energy Poverty Index weights	CEPI	Inadequate living conditions (not warm, not cool, dark) Arrears Leaks
Gupta et al., 2020	Authors present a novel measurement of energy poverty on a household level. The new index aims to form an analytical basis for energy policies in India.	India	Principal Component Analysis (PCA)	Household Energy Poverty Index (HEPI)	15 indicators that present different dimensions of energy and energy poverty

Faiella & Lavecchia, 2021	The authors present a multidimensional indicator that was developed by them earlier and was adopted by Italian government to measure energy poverty.	Italy	The authors determine a threshold for needed heating level using adjustment factor and find the level of income needed to afford it. Then they create a matrix for mapping the minimum needed heating expenditure for each combination of household type, geographical area and city size.	Measure of absolute fuel poverty	The authors estimate the heating demand of a household through the integration of technical heating requirement information with expenditure data from Italian Household Budget Survey
Li et al., 2021	Authors analyse the relationship between energy efficiency and energy poverty in developed and developing countries to better understand how energy poverty may affect country's development.	USA, India, Russia, Italy, UK, Norway, Qatar, Kuwait, Germany, Thailand, Austria, South Korea, Spain, Indonesia	Data Envelopment Analysis and entropy analysis	Measurement of energy poverty effect on social welfare	Energy poverty Energy efficiency Socio-economic indicators
Kelly et al., 2020	The authors propose a new composite index that assesses energy poverty in Ireland taking into account these characteristics.	Ireland	Weighting based on an arbitrary, a priori basis using Relative Risk Ratio (RRR) calculations	HH-EPRI	Heating requirements (3 indicators) Building characteristics (2 indicators) Household characteristics (5 indicators)

Sokołowski et al., 2020	Authors propose a multidimensional energy poverty index that considers multi-faceted nature of energy poverty but at the same time can be easily used for poverty mapping and in policy decision making process.	Poland	Authors assess join distribution of deprivations that is defined by a set of deprivations D. Then authors use dual cut-off approach. A household is considered as poor if the weighted deprecation score is higher than the set poverty cut-off. To derive the value of the index, the authors calculate the headcount ratio that considers multi-dimensionally poor households.		Household income and household energy expenditure used to calculate low income, high costs and high share of energy expenditure in income Inability to keep the home adequately warm Presence of leaks, damp, or rot Inability to pay utility bills
Ehsanullah et al., 2021	Authors address the nexus between energy poverty and energy insecurity with the role of various environmental concerns.	G7	Data envelopment analysis (DEA)	Energy, economic, social, and environmental performance index (EPI)	Set of indicators on energy economics and environmental concerns related to energy poverty
Jayasinghe et al., 2021	Authors create an energy poverty index for Sri Lanka to better understand this problem in a country.	Sri Lanka	Principal Component Analysis (PCA)	Multidimensional Energy Poverty Index (MEPI)	Set of indicators from household income and expenditure survey: use of modern cooking fuel, access to electricity, having a fridge, having a radio or television, having a landline or mobile

					phone, having a computer, having access to electric fan
Lan et al., 2022	Authors develop a multidimensional energy poverty index for five Asian countries. The index is used to better examine the consequences of energy poverty in Asia.	Pakistan, India, Sri Lanka, Bangladesh, Nepal	Authors used multi-criterion decision analysis (MCDA) and standard-weight data envelopment (DEA) related model to assign weights to selected indicators.	Multidimensional energy poverty index (EPI)	Rural people's electricity access Urban people's electricity access Alternative and nuclear energy Electric power consumption Net energy imports Energy use Fossil fuel energy utilization Energy consumption (GDP per unit) R&D Output of renewable electricity Renewable energy utilization The time it takes to get energy Access to clean cooking fuels and technology Primary energy's energy intensity level At ($1.90)/per day, the poverty headcount ratio Gini Index
Liang & Asuka, 2022	There is no consensus on how to measure energy poverty. As	China	Entropy method	Multidimensional energy index for	Household energy consumption (5 indicators)

	China is undergoing socioeconomic transformations and energy transition, new index to capture China's energy poverty is needed.			China. Used to calculate energy poverty in China from 2014 to 2019	Household energy supply (5 indicators) Energy consumption structure (4 indicators) Air pollutants from residential energy consumption (4 indicators) Energy/electrical appliances (5 indicators) Residential energy affordability (4 indicators)
Hasheminasab et al., 2023	There is a need for a comprehensive energy poverty assessment for the EU that considers sustainability.	EU	Multiple Criteria Decision-Making methodology – ITARA methodology	Energy poverty index taking into account changing energy market in the EU	Primary energy consumption Final energy consumption Final energy consumption in household per capita Energy productivity Share of renewable energy in gross final energy consumption by sector Energy import dependency by products Population unable to keep home adequately warm by poverty status Greenhouse gas emissions intensity of energy consumption

Appendix 2. Variables used for the green energy poverty index and its validation

Descriptive statistics

Variables	Source	Measurement unit	Median	Average	Standard deviation	Min value	Max value
Energy poverty							
Inability to keep adequately warm	Eurostat Code: ilc_mdes01	Percentage of population struggling to maintain adequate temperature at their home because of financial difficulties	6.1	10.253	9.962	0.5	66.5
Arrears on utility bills	Eurostat Code: ilc_mdes07	Percentage of population that are unable to pay their utility bills on time because of financial difficulties	7.4	10.387	8.575	1.2	42.2
Use of renewable energy sources							
Use of renewables for electricity	Eurostat Code: nrg_ind_ured	Produced electricity from renewable resources as a share of the consumed elevtricity	11718.408	32683.445	47076.169	0.68	260603.5
Use of renewables for heating and cooling	Eurostat Code: nrg_ind_urhcd	Produced energy for heating and cooling from renewable resources as a share of the consumed energy	1666.2	3590.347	4092.031	4.32	18628
Share of energy from renewable resources	Eurostat Code: nrg_ind_ren	Produced energy from renewable resources as a share of the consumed energy	17.852	20.816	11.819	0.979	66.002
Variables for the validation of the index							
Absolute expenditure on electricity	Eurostat Code: nama_10_co3_p3	Total household expenditure on electricity	4654.000	11187.524	16938.217	105.500	91487.000

Disposable annual household income	Eurostat Code: ilc_mdho04	Total income received by all household members from all sources	5.600	5.829	1.945	2.600	12.100
Population considering their dwelling as too dark	Eurostat Code: nama_10r_2hhinc	Percentage of population that consider their dwelling as not having enough light or being too dark	13200.000	14330.094	7295.245	3100.000	38300.000

Table 11. Descriptive statistics

Variables presented by country over time

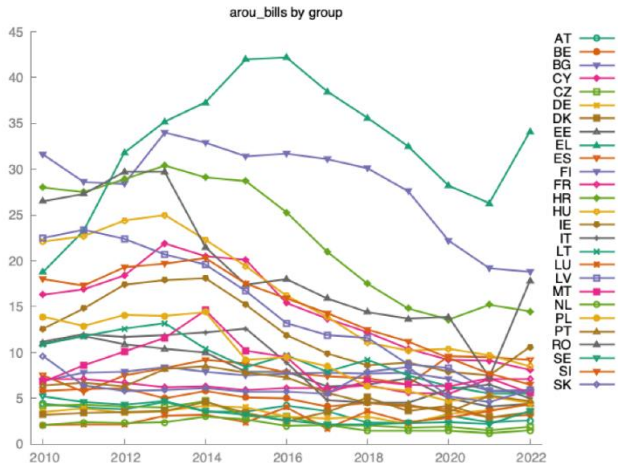

Figure 31. Arrears on utility bills in EU27 2010-2022

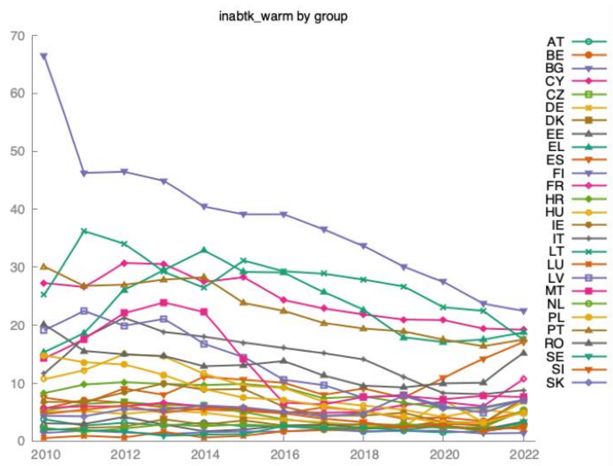

Figure 30. Inability to keep adequately warm in EU27 2010-2022

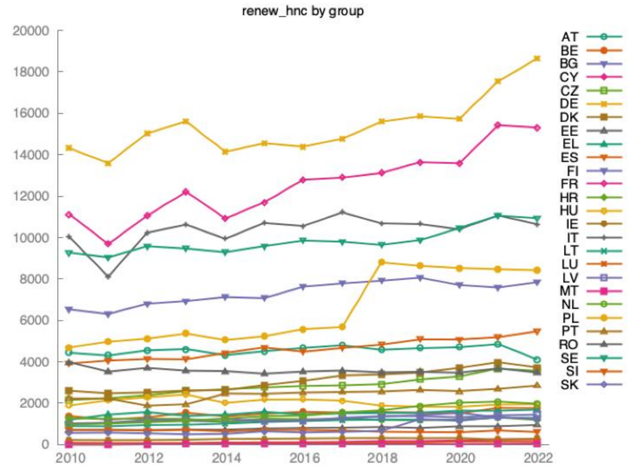

Figure 34. Use of renewables for heating and cooling in EU27 2010-2022

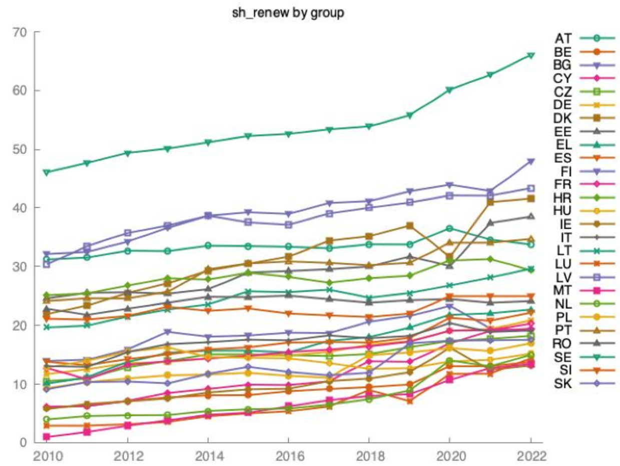

Figure 32. Share of energy from renewable resources for EU27 2010-2022

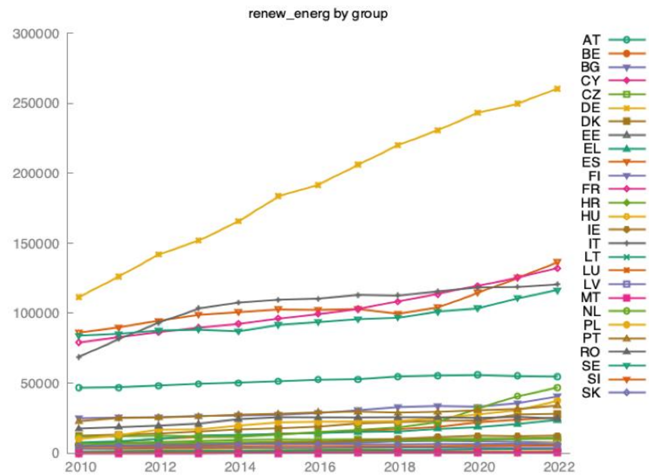

Figure 33. Use of renewables for electricity in EU27 2010-2022

Box plots for outliers' detection

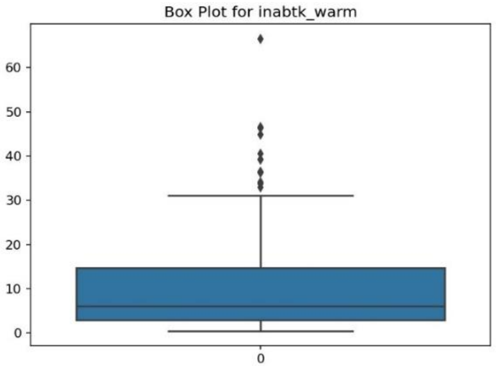

Figure 38. Box Plot for Inability to keep warm variable

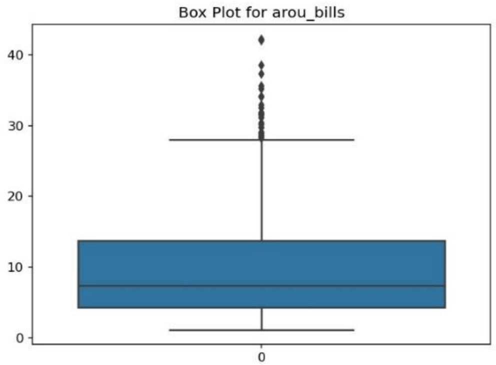

Figure 39. Box Plot for Arrears for utility bills variable

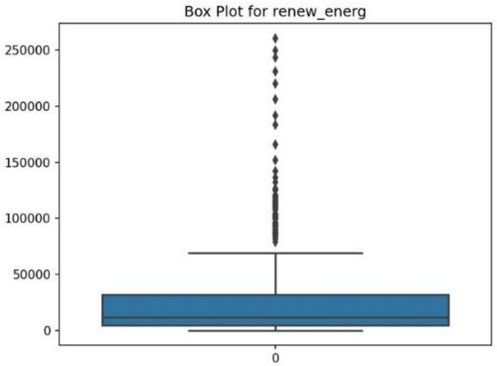

Figure 37. Box Plot for Use of renewables for electricity variable

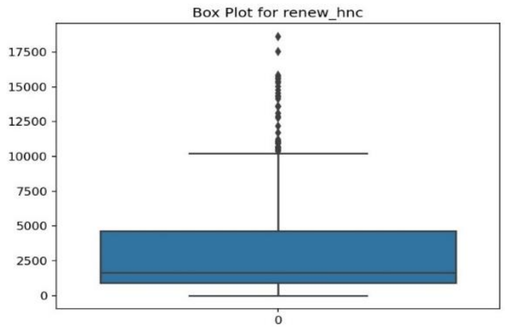

Figure 36. Box Plot for Use of renewables for heating and cooling variable

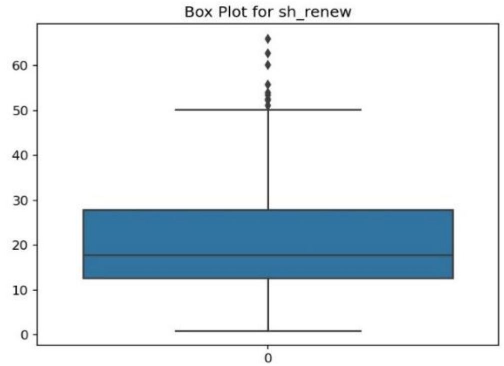

Figure 35. Box Plot Share of energy from renewable resources variable

Appendix 3. Results of index for 2022

Country	Country code	Green Energy Poverty Index (PCA) not scaled	Green Energy Poverty Index (PCA) scaled	Efficiency Score (DEA)
Austria	AT	6.155	0.528	0.679
Belgium	BE	1.578	0.411	0.421
Bulgaria	BG	-4.866	0.245	0.083
Cyprus	CY	-3.367	0.284	0.164
Czech Republic	CZ	2.293	0.429	0.750
Germany	DE	24.485	1.000	1.000
Denmark	DK	4.177	0.478	0.507
Estonia	EE	1.188	0.401	0.451
Greece	EL	-5.594	0.226	0.059
Spain	ES	8.038	0.577	0.195
Finland	FI	7.401	0.561	0.699
France	FR	13.562	0.719	0.308
Croatia	HR	-1.057	0.343	0.148
Hungary	HU	-0.287	0.363	0.229
Ireland	IE	-1.595	0.329	0.172
Italy	IT	11.242	0.659	0.369
Lithuania	LT	-1.444	0.333	0.279
Luxembourg	LU	-0.447	0.359	0.454
Latvia	LV	0.929	0.394	0.307
Malta	MT	-1.638	0.328	0.258
Netherlands	NL	3.410	0.458	0.838
Poland	PL	5.400	0.509	0.412
Portugal	PT	1.679	0.413	0.348
Romania	RO	-1.201	0.339	0.101
Sweden	SE	14.817	0.751	1.000
Slovenia	SI	0.187	0.375	0.349
Slovakia	SK	-0.450	0.359	0.262

Table 12. Results of PCA and DEA analyses – values for 2022

Appendix 4. Normality test results

Variables	Shapiro-Wilk test statistic	Shapiro-Wilk test p-value
Energy poverty		
Inability to keep adequately warm	0.802	0.000
Arrears on utility bills	0.829	0.000
Use of renewable energy sources		
Use of renewables for electricity	0.687	0.000
Use of renewables for heating and cooling	0.782	0.000
Share of energy from renewable resources	0.940	0.000
Variables for the validation of the index		
Absolute expenditure on electricity	0.640	0.000
Disposable annual household income	0.964	0.000
Population considering their dwelling as too dark	0.935	0.000
Indices		
Green Energy Poverty Index based on PCA	0.943	0.000
Efficiency score based on DEA	0.885	0.000

Table 13. Results for Shapiro-Wilk normality test

Appendix 5. Target values for 2022 from DEA analysis

	Target values					Actual values				
Country	Inability to keep adequately warm	Arrears on utility bills	Use of renewables for electricity	Use of renewables for heating and cooling	Share of energy from renewable resources	Inability to keep adequately warm	Arrears on utility bills	Use of renewables for electricity	Use of renewables for heating and cooling	Share of energy from renewable resources
Austria	1.834	1.766	66445.957	5801.465	33.758	2.700	2.600	54820.600	4104.400	33.758
Belgium	2.146	1.347	38677.653	2190.193	13.759	5.100	3.200	25334.500	1742.500	13.759
Bulgaria	1.870	1.563	43620.015	3136.158	19.095	22.500	18.800	7744.900	1263.200	19.095
Cyprus	2.309	1.330	49770.328	3236.976	19.429	19.200	8.100	893.600	237.600	19.429
Czech Republic	2.174	1.424	50790.809	3539.900	21.161	2.900	1.900	10900.100	3539.900	18.195
Germany	6.700	4.300	260603.500	18628.000	20.796	6.700	4.300	260603.500	18628.000	20.796
Denmark	1.997	1.776	80994.355	7243.756	41.601	5.100	3.500	28175.000	3711.400	41.601
Estonia	1.535	1.986	70108.847	6632.152	38.472	3.400	4.400	2895.600	938.800	38.472
Greece	1.108	2.020	39549.446	3718.718	22.678	18.700	34.100	23742.000	1666.200	22.678
Spain	2.344	1.796	136487.900	9582.015	22.116	17.100	9.200	136487.900	5452.000	22.116
Finlan	0.979	3.985	82566.443	8808.225	47.886	1.400	5.700	40586.100	7828.500	47.886
France	2.444	2.183	219492.818	15295.700	20.259	10.700	7.100	132287.700	15295.700	20.259
Croatia	1.034	2.143	49218.338	4927.065	29.354	7.000	14.500	10438.100	1187.500	29.354
Hungary	1.075	1.944	27234.540	2352.928	15.190	4.700	8.500	7353.000	1939.300	15.190
Ireland	1.241	1.828	26058.522	1989.020	13.107	7.200	10.600	12452.500	278.000	13.107
Italy	2.405	1.847	154689.882	10625.600	19.131	8.800	5.000	120601.400	10625.600	19.131
Lithuania	2.166	1.534	64092.372	5074.832	29.599	17.500	5.500	3385.500	1266.400	29.599
Luxembourg	0.954	1.999	24362.399	2190.121	14.356	2.100	4.400	1065.400	151.400	14.356
Latvia	1.973	1.810	83409.528	7553.679	43.316	7.100	5.900	3895.800	1433.300	43.316
Malta	1.959	1.444	35716.247	2108.506	13.404	7.600	5.600	297.000	51.600	13.404

Netherlands	2.375	1.257	46918.100	2673.478	14.972	5.300	1.500	46918.100	1968.300	14.972
Poland	2.019	1.854	97937.571	8403.900	41.146	4.900	4.500	37190.800	8403.900	16.879
Portugal	2.095	1.636	71243.536	5992.495	34.677	17.500	4.700	34009.100	2837.200	34.677
Romania	1.540	1.804	47413.497	4024.272	24.140	15.200	17.800	24631.800	3465.600	24.140
Sweden	3.300	3.600	116494.700	10928.700	66.002	3.300	3.600	116494.700	10928.700	66.002
Slovenia	0.907	2.268	41182.378	4182.271	25.002	2.600	6.500	5480.200	590.200	25.002
Slovakia	1.862	1.547	40981.511	2845.296	17.501	7.100	5.900	6404.400	1224.200	17.501

Table 14. Target values and actual values from DEA analysis for 2022