INTERDEPARTMENTAL PROGRAMME OF POSTGRADUATE STUDIES IN BUSINESS ADMINISTRATION

Thesis

A CONTRIBUTION TO THE COMPARATIVE EVALUATION OF EFFICIENCY INDICATORS IN EUROZONE USING DEA METHODOLOGY

By

THOMAS E. PAPPAS

Under the supervision of

DR. ANDREAS GEORGIOU, PROFESSOR

Thesis submitted for the degree of Master in Business Administration

01/2024

Acknowledgements

I would like to express my deepest appreciation to Dr. Andreas Georgiou for his invaluable contributions to this research. As a tutor, he ignited my passion for quantitative methods during his courses and was always willing to help me and answer my questions. As a supervisor, he provided me with the chance to explore an especially intriguing topic pertinent to my previous studies. I am deeply indebted to his guidance and encouragement, without which this work would not have been possible.

I would like to extend my gratitude to Dr. Georgios Tsaples for his priceless assistance in the completion of this thesis. His eagerness to share his knowledge and expertise on my research topic was truly instrumental in the success of this project. Without his observations and suggestions, this thesis would not have been feasible.

Last but not least, I would like to thank my family, who have been a constant source of support throughout my undergraduate and postgraduate studies. Without their unwavering encouragement, this master's degree would not have been obtained.

Abstract

This study proposes a methodological framework for the construction of composite sustainability indicators. To be more specific, it suggests the application of an operations research technique, namely Data Envelopment Analysis (DEA), to weigh and aggregate economic, environmental, and social indicators into composite indices in such a way that diminishes the contingent bias that is linked with assigning weights. These indices are used to assess the intricate notion of sustainability efficiency of 20 Eurozone members comparatively, aiming to share valuable information for policymakers and governing bodies, helping them to make more informed decisions and eventually enhance the sustainability performance of the countries under assessment. A literature review of 25 papers about DEA and sustainability for the years 2021 to 2023 was carried out to select the indicators that were used in the study. The data used were retrieved from Eurostat and the World Bank databases. Afterwards, the chosen indicators were implemented under two DEA variations (classic and SBM) and 10 different scenarios, which generated 10 final indices. Those indices are separated into three sustainability indices, six eco-efficiency indices, and one socio-economic index. Ireland, Luxembourg, and Malta attained maximum efficiency scores across all sustainability scenarios, while Ireland and Luxembourg were deemed efficient under all of the eco-efficiency variations. Finally, the results of the sustainability scenarios may indicate that most of the Eurozone members do not get close to their potential output levels (GDP and Overall Life Satisfaction) in terms of their available labor force.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. Introduction

1.1 General Background

In the current era marked by economic transformation, there is a pressing need to comprehend how these changes impact the environment and society. To address this requirement, the European Union, via Eurostat, and researchers in the field have developed a multitude of indicators to gauge the sustainability efficiency of countries and regions.

These indicators are critical in understanding the interplay between economic development, environmental preservation, and social welfare, which constitute essential components of sustainable development. Analyzing these indicators can provide insights into policy formulation and implementation and facilitate comparisons between regions and countries, thereby contributing to a more comprehensive understanding of sustainable development.

Sustainability constitutes a big challenge in today's era (Sachs, 2015). By aiming to obtain dynamic and continuous harmony among ecological subsystems (environmental sustainability), social subsystems (social sustainability), and economic subsystems (economic sustainability), sustainability is by nature complex, multi-dimensional, and embedded with trade-offs among multiple sustainability dimensions (Wu, 2013).

According to Eurostat (2023), an indicator constitutes a statistical and possibly logical order of magnitude, which is naturally or arbitrarily connected with the measurement of policy activities in the broader sense of governance. The primary advantage of indicators is that they provide information in a summary form, are easy to communicate, and are subject to relative unanimity. As a rule, an indicator can be defined by its function, the means of obtaining it, its quality, and the limits on its use. (Eurostat, 2023)

Governments and corporations have utilized the indicators because of their ability to concentrate and tabulate the enormous intricacy of the contemporary dynamic environment to an easy-to-handle amount of meaningful information. (Godfrey & Todd, 2001). Indicators simplify, quantify, and share complex information by conceptualizing various phenomena, assessing the trends, and identifying the key spots (Warhurst, 2002).

According to KEI (2005, p. 1), "Indicators and composite indicators are increasingly recognized as a useful tool for policy making and public communication in conveying information on countries' performance in fields such as environment, economy, society, or technological development".

Numerous sustainability indices are more and more employed by policymakers, aiming to make informed policy decisions (Oras, 2005; Hezri & Dovers, 2006), and it is crucial to recognize the strengths, weaknesses, biases, and scale-dependence of these indices in order to use them effectively. (Parris & Kates, 2003; Morse & Fraser, 2005; Ness et al., 2007).

Although sustainability indices possess significant merits, they do not come without their respective demerits and limitations. For instance, numerous methodological issues should be taken into consideration in order to evaluate sustainability index performance. Some of them are the predetermined limits of the system, the quality of the data, which are included in the analysis, the normalization method, the weighting method, the aggregation method, and the comparability of the outcomes across systems. (Mayer, 2008)

With regard to weighting methods, the literature offers various approaches. Based mainly on Nardo et al. (2005), OECD (2008), Hermans et al. (2008), and Mikulić et al. (2015), these methods include equal weighting, which assigns equal importance to all variables; statistical-based methods, such as principal component analysis and regression analysis; and participatory-based methods, such as public opinion and budget allocation, which involve stakeholders in the weighting process to ensure that their perceptions are taken into account.

According to Nardo et al. (2005), equal weighting can be employed when all indicators are considered equally important or when no statistical or empirical evidence supports a different plan. However, in the complex case of sustainability, the assumption that all factors are equally important may lead to unfavorable outcomes.

Moreover, while statistical methods try to weigh each index's importance objectively, they do not come without their drawbacks. For example, according to Gan et al. (2017), in regression analysis, either multi-collinearity among indicators or an inappropriate dependent variable may lead to incorrect results. Additionally, despite the fact that they give weights on the index variables, participatory-based methods do so subjectively, and hence, the results are directly affected by the raters' opinions (Kuosmanen & Kortelainen, 2005).

Furthermore, in the context of aggregation methods, common approaches, according to Munda and Nardo (2009), Beliakov et al. (2007), OECD (2008), and Pollesch and Dale (2015) are the additive aggregation, such as the weighted arithmetic mean, the geometric aggregation, such as the weighted geometric mean and the non-compensatory aggregation methods. The non-compensatory aggregation methods are based on the properties of the aggregation functions (Pollesch & Dale, 2015) and the perspective of multi-criteria decision-making (MCDM) (Guitouni & Martel, 1998; Munda, 2005). However, the first two methods require strict and specific conditions, such as mutually preferential independence (Keeney, 1973; Keeney, 1974). Finally, concerning the third method, it is important to note that the loss of information relevant to the intensity of sustainability constitutes a potential demerit (Munda & Nardo, 2009).

Regarding the above limitations, this study aims to explore how the different sustainability indicators can be managed in order to construct a composite sustainability index that measures country sustainability in such a way that reduces the contingent bias that is generally associated with assigning weights.

In an effort to address some of these problems, numerous researchers in the field have adopted a mathematical programming technique called Data envelopment analysis (DEA). DEA is a non-parametric approach for evaluating the performance of a set of peer entities called Decision Making Units (DMUs), which transform multiple inputs into multiple outputs (Cooper et al., 2010), and hence it is particularly useful in order to aggregate indicators and assess complex notions like sustainability.

Furthermore, DEA is employed to measure the technical efficiency of those DMUs (Førsund & Sarafoglou, 2002), where technical efficiency can be seen as the capability of a DMU to convert its inputs into outputs and its prescribed as the ratio of the sum of its weighted outputs over the sum of its weighted inputs (Ishizaka & Nemery, 2013; Thanassoulis, 2001) as it is indicated in the following expression:

$$technical\ efficiency = \frac{\Sigma w_{output} * y}{\Sigma w_{input} * x},$$

where x = input level and y = output level

The fundamental bases of DEA originate from the papers of Debreu (1951), Farrell (1957), and Diewert (1973), whereas the method became well-established through the influential works of Charnes et al. (1978) and Banker et al. (1984). The use of DEA with the purpose of gauging the sustainability performance of countries or regions constitutes an increasing trend in the recent literature. Zhou et al. (2018) identified that trend and wrote a literature review about DEA and regional sustainability assessment, where they concluded that DEA constitutes an appropriate method for this purpose.

DEA, through linear programming, determines the weights of the indicators used in the model in order to produce the final index. Hence, it can weigh the indicators in such a manner that decreases the bias that is commonly connected with assigning weights. Furthermore, DEA's ability to convert multiple inputs into multiple outputs makes it particularly suitable for aggregating sustainability indicators into a composite index. This capability allows for a more reliable and robust measure of country sustainability, as it considers the complex interrelationships between different dimensions of sustainability.

Another advantage of DEA is that it does not need to identify the process of how inputs transform into outputs, while it only needs information about output and input quantities, not prices (Papathanasiou et al., 2021). It also needs less information compared to parametric methods. For instance, it does not require information about the statistical distribution of the data. (Hajiagha et al., (2016) ; He et al., (2016)). The abovementioned constitute significant advantages of DEA, as they enable its application in situations where such information is not available or where the relationships between inputs and outputs are complex or non-linear.

1.2 Thesis Purpose and Objectives

The purpose of the current thesis is to suggest a methodological framework, more specifically to propose the implementation of Data Envelopment Analysis in order to aggregate and weigh various sustainability sub-indicators that have been applied in the literature and eventually develop composite sustainability indices and gain insights from their comparisons. More precisely, these sustainability indices will be created with the aim of measuring the relative sustainability efficiency of 20 countries of the Eurozone and simultaneously discovering and sharing valuable information for governments, policymakers, and other researchers.

The objectives of this study are :

1. To create different sustainability indices using DEA, taking into consideration several economic, environmental, and social indicators, in order to measure the intricate and multidimensional concept of sustainability.
2. To use these indices to assess the sustainability efficiency of 20 European countries that use the euro as their official currency comparatively.
3. To utilize the results of the final indices in order to identify the countries that are relatively efficient, which afterwards can be used as role models for the less efficient countries.

All of the above eventually aim to help governments and policymakers make more informed decisions, apply more successful policies, and finally enhance the sustainability efficiency of the Eurozone countries.

1.3 Thesis Structure

The rest of the thesis is structured as follows.

Section 2 of the study is dedicated to a literature review on the use of Data Envelopment Analysis (DEA) for sustainability assessment. The focus is on exploring how authors have utilized DEA methodology for aggregating sustainability indicators into composite sustainability indices.

In **section 3**, the methodology of Data envelopment analysis is described theoretically and mathematically.

In **section 4**, the DEA variations that will be utilized are chosen, formulated mathematically, and employed with the twofold purpose of creating several composite sustainability indices and evaluating the relative sustainability efficiency of 20 Eurozone countries.

In **section 5**, the findings obtained through the implementation of the proposed models are further analyzed and discussed.

Finally, in **section 6,** conclusions, lessons learned from the research, and propositions for further research are presented and discussed.

2. Literature Review

2.1 Sustainability, Sustainable Development and European Union

The concept of sustainable development was first introduced by the United Nations Conference on the Human Environment in 1972 (Marta Negri et al., 2021), but it became well-known in 1987 in the report of the World Commission for the Environment and Development (WCED) also known as Brutland. This report was the turning point globally in an approach to issues relevant to the socio-economic and environmental dimensions of development processes.

The WCED's report (1987) became the first proposition of a complete approach to subjects of social and economic development from an environmental point of view. The term sustainable development, as prescribed in the Brutland report, is defined as meeting the needs of present generations without jeopardizing the ability of future generations to meet their own needs.

This specific definition focuses on two fundamental principles. The first is meeting the basic needs of the people below the line of poverty, while the second emphasizes the need to maintain the ability of the environment to meet both present and future needs through the use of technology and social organization.(WCED, 1987)

Wu & Wu (2012) support the fact that sustainability concerns the ability to retain a paired human-nature system at a covetable state for multiple generations in front of anthropogenetic and environmental perturbations and uncertainties. Bearing in mind the complexity arising from the multiplicity of components and their intricate interactions, it becomes challenging to find a clear definition of sustainability in specific terms without controversy. (Wu & Wu, 2012, chapter 4)

The phenomenon of sustainable development is quite complex, thus making the comparison and the valuation of advances of European Union member states in the implementation of its objectives specifically tricky. Mariola Grzebyk and Małgorzata Stec (2015) tried to establish a synthetic measure of the level of sustainable development, taking into consideration concurrently the economic, social, and environmental factors. There is a large number of authors in the literature who accept that these three pillars can effectively describe sustainability or sustainable development, such as Elkington (1998), Trianni et al. (2017), and Nikolaou et al. (2021).

Regarding economic and social development, a political scientist called Robert Gibson (2001) argues that it should not be measured solely by material gains since they do not guarantee human well-being. He also suggests that, except for the three dimensions that reflect the disciplines of those who study sustainability, it would have been helpful to include cultural and political dimensions as well. In addition to the above, the author does not accept the general idea of pillars; instead, he develops seven principles on which sustainability could be based.

Another definition of sustainability is given by Kuhlman and Farrington (2010), who describe it as the ability to maintain well-being for a long or even indefinite period. This sentence covers mainly the environmental factor of the three-dimensional model, but despite that, sustainability and the environment are not synonymous.

On the one hand, some forms of environmental degradation are relatively easily reversed but are highly harmful in the present, like many forms of air and water pollution, for example. These have a vital aspect of well-being, and, indeed, they are included under both social and environmental dimensions in the European Union guidelines for impact assessment (European Commission, 2005). On the other hand, what we bestow on future generations, except the aforementioned, also encompasses cultural heritage, namely, art and cultural landscapes, as well as technology and institutions.

As Gieryn (1999) mentions, sustainability has become a "boundary term" where science meets politics and politics meets science. The boundary work relevant to the sustainability of building epistemic communities of shared understanding and joint commitment to connecting environmental and economic development subjects has become a central concern all over the world.

Ecologists have been concerned for a long time about how ecosystems react to shocks and stresses. Mathematical ecology thrived through the 1970s and 1980s, with the vital work of scientists like Buzz Holling and Bob May on the stability and resilience properties of both models and actual biological systems (Holling 1973; May 1977). From this point of view, sustainability can be characterized as the ability of a system to bounce back from such shocks and stresses and adopt stable states. (Scoones, 2007)

In an era where social and economic inequalities, climate change, and environmental abasement have become vital challenges worldwide, the international community has committed itself to the 2030 agenda for sustainable development (Banerjee & Duflo, 2011). The sustainable development goals are categorized into five groups, called the "five P's": People, Planet, Prosperity, Peace, and Partnership. Each of those goals constitutes a global challenge that needs a global response. (Brzyska & Szamrej-Baran, 2023) Additionally, the SDGs are common to every UN member state, although each country can adjust them to its own context and specified needs. (Un Summit, 2015)

The European Union is strongly committed to promoting sustainable development while adopting a comprehensive approach to integrating the SDGs into its policies. Furthermore, it financially supports an extensive range of projects and initiatives that support sustainable development goals, regularly reviews its progress on them, reports to its citizens, and makes a contribution to the global review process at the UN's high-level political forum on sustainable development. (Brzyka & Szamrei-baran, 2023)

Sustainability indicators constitute a beneficial tool in people's attempts to measure sustainability and sustainable development. These indicators provide information about the state, dynamics, and underlying drivers of human-environmental systems (Wu & Wu, 2012). Sustainable development indicators must be developed to create stable bases for decision-making at all levels and to contribute to the self-regulating sustainability of integrated environment and development systems (UN, 1992). Moreover, Meadows (1998) mentions that indicators become sustainable or unsustainable when time, limits, or targets are linked with them.

Generally, sustainability indicators are composite indicators (CI). Namely, they are mathematical aggregations of a set of individual indicators that measure multidimensional concepts but usually have no common units of measurement (Nardo et al., 2005). Composite indicators have also been widely established as valuable tools for monitoring performance, benchmarking, policy analysis, and public communication in the field of sustainability. (Zhou et al., 2018)

After analyzing this section, it becomes clear that various definitions of sustainability exist. Although the three-pillar approach (economic, environmental, and social) is widely acknowledged, the existence of diverse approaches can result in different interpretations of sustainability. This fact, in turn, can lead to communication problems. Consequently, it is essential to set a universally accepted definition to guarantee that all parties involved comprehend the term in the same way.

2.2 DEA and Sustainability

Data envelopment analysis (DEA) is a mathematical programming technique for evaluating the performance of a set of peer entities called Decision Making Units (DMUs), which transform multiple inputs into multiple outputs (Cooper et al., 2010). In the original article of Charnes et al. (1978) DEA was described as a "mathematical programming model applied to observational data [that] provides a new way of obtaining empirical estimates of relations – such as the production functions and/or efficient production possibility surfaces – that are cornerstones of modern economics." (Cook & Zhu, 2005, p. 2; Cooper et al., 2010, p. 2; Macedo et al., 2023, p. 39)

According to Emrouznejad et al. (2023), DEA is a mathematical tool for assessing the relative efficiencies of decision-making units (DMUs) with multiple inputs and multiple outputs. Thus, it has been established as an appropriate method to measure sustainability, and as Karadayi and Ekinci (2018) mention, it has become the most widely employed model in the relevant literature.

There are some crucial features of DEA that make it a suitable method for assessing the relative sustainability efficiency of different countries as well as the relative efficiency of different kinds of DMUs generally. Firstly, it can use the data from observed operating units to create other feasible, in principle, operating units even if not observed in practice, and secondly, it is a boundary method that, out of the units observed or created, can identify the most "efficient" ones. (Georgiou et al., 2021)

Adler (2011), as well as Førsund and Sarafoglou (2002), delineate DEA as a non-parametric method that is used for the evaluation of the technical efficiency of Decision Making Units relative to one another, where technical efficiency can be defined as a measure of how well a DMU converts its inputs into outputs (Tsaples & Papathanasiou, 2021).

According to Thanassoulis (2001), technical efficiency can be viewed as the ability of a DMU to convert its inputs into outputs, and it is defined as the ratio of the sum of its weighted outputs over the sum of its weighted inputs as represented in the undermentioned expression:

$$technical\ efficiency = \frac{\Sigma w_{output^*}y}{\Sigma w_{input^*}x},$$

where x = input level and y = output level

DEA was created in the influential papers of Charnes, Cooper and Rhodes (1978) and Banker, Charnes and Cooper (1984), who developed the Constant Returns to Scale model and the Variable Returns to Scale model correspondingly. In the general-purpose linear programming model, it is assumed that there are N DMUs (j= 1…N) using m inputs to secure s outputs. It is also denoted that Xij and Yrj are the level of the ith input and rth output that observed at DMU j.

Now it is possible to compute the technical efficiency k_0 for the input oriented model by solving the following linear program (LP) :

$$\min k_0 - \varepsilon[\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+]$$

Subject to constraints:

$\sum_{J=1}^{N} \lambda_j\, x_{ij} = k_0 x_{ij_0} - S_i^-$ i = 1 . . . m

$\sum_{J=1}^{N} \lambda_j\, y_{rj} = S_r^+ + y_{rj0}$ r = 1 . . . s

$\lambda_j \geq 0, j = 1 \ldots N, S_i^-, S_r^+ \geq 0\ \forall\ i\ and\ r,\ k_0$ free.

ε is a non-Archimedean infinitesimal.

The variable λj is the weight calculated from the above model for DMU j whilst the variables S_r^+, S_i^- are the slack variables that are used in linear programming. These slacks represent any further output increase or input decrease that can be achieved by the DMU.

Once k_0 has been minimized, the model tries to obtain the maximum sum of the slack values S_r^+, S_i^-. If at least one of those values is greater than zero at the optimal solution of the model, it entails that the equivalent input or output of the DMU j_0 can be further improved after its input levels have been contracted to the proportion k_0*, which is the technical efficiency of the model.

If k_0* = 1 and S_r^+ = 0, r = 1…s, S_i^- = 0, i = 1… m then DMU is Pareto-efficient because the model was unable to decrease the input level without decreasing the output level or increasing the output level without correspondingly increasing the input level of the specific DMU.

In summary, the technical efficiency of the above problem for $DMUj_0$ is the variable k_0 and can take values from 0 to 1 (or from 0 to 100%). The mathematical program indicated by the above equations is solved individually for every DMU, and its solution leads to three possible cases:

1. $DMUj_0$ is Pareto-efficient if and only if $k_0* = 1$ and $S_r^+ = 0$, r = 1…s, $S_i^- = 0$, i = 1… m

2. If the value of one of the slacks S_r^+, S_i^- is greater than zero at the optimal solution, the related input (or output) can additionally be improved

3. If none of the aforementioned happens, $DMUj_0$ has technical efficiency equal to k_0*. In this case, the technical efficiency at the optimal solution $k_0* < 1$ reflects the maximum radial contraction of the input levels without deteriorating the level of outputs in order for the $DMUj_0$ to be considered efficient. (Thanassoulis, 2001)

The above program constitutes a classical DEA model. These models have several advantages, so they have been utilized a lot in the recent literature. Nevertheless, classic DEA models do not come without their drawbacks. For instance, they treat the stage of converting inputs into outputs as a "black box" and, as a result, the robustness of the method is decreased.

With the goal of getting over this limitation, a new area of development in the context of DEA has appeared, where the authors develop DEA models with internal structures (or network DEA models). One subcategory of those models are the two-stage DEA models, whose structure is illustrated in the figure below.

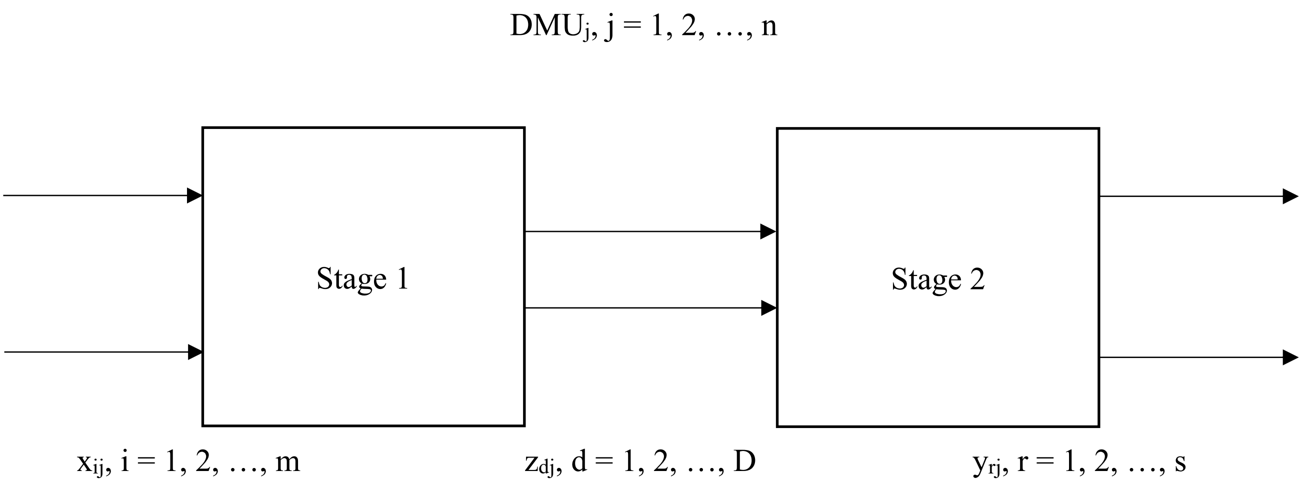

Figure 1 Standard two stage model structure

Utilizing the notation of Chen and Zhou (2004), denote that every DMU_j has m inputs x_{ij}, (i = 1, 2, …, m) and D outputs z_{dj}, (d = 1, 2, …, D) in the first stage. Afterwards, those outputs D become inputs of the second stage, they will be referred to as intermediates, while the outputs of the same stage are y_{rj}, (r = 1, 2, …, s).

Denote also the efficiency of the first stage as e_j^1 and the efficiency of the second stage as e_j^2, for every DMU_j. By utilizing the constant returns to scale (CRS) DEA model of Charnes et al. (1978), the aforementioned efficiencies can be defined as indicated by the following equations:

$$e_j^1 = \frac{\sum_{d=1}^{D} w_d * z_{dj}}{\sum_{i=1}^{m} v_i * x_{ij}}$$

and

$$e_j^2 = \frac{\sum_{d=1}^{D} u_r * y_{rj}}{\sum_{d=1}^{D} \widetilde{w}_d * z_{dj}}$$

where v_i, w_d, $\widetilde{w}_d$, u_r are unknown and non-negative weights. Note also that w_d and $\widetilde{w}_d$ can be equal.

Tsaples and Papathanasiou (2021) contributed to the growing body of literature on Data Envelopment Analysis (DEA) by conducting a comprehensive literature review for the years 2016-2020, with a specific focus on the use of DEA to measure sustainability and related notions of efficiency. Building upon the earlier review by Zhou et al. (2018), the authors aimed to provide an up-to-date and detailed overview of the current state of research in this area.

Both of the studies acknowledged that even in the context of DEA, there is no generally accepted definition of sustainability, while the authors, over the years, used terms like efficiency or eco-efficiency to describe it. They also pointed out in their efforts the lack of use of social indicators as inputs and outputs in the existing DEA models, resulting in an incomplete measure of sustainability that neglects the social dimension.

This omission can lead to incorrect assessments of sustainability performance, as it fails to consider the impact of social factors on overall sustainability. Therefore, there is a need for further research to apply more comprehensive models that incorporate social indicators and provide a more holistic assessment of sustainability.

For the current work, a search in bibliographic databases like Google Scholar was carried out, using key phrases such as “DEA and country/regional efficiency”, “DEA and efficiency indicators” and “DEA and sustainability”. This review includes papers published between 2021 and 2023 and intends to extend the work of the above-mentioned authors.

Summarizing the findings of this section, it becomes evident that the nature of DEA, as well as its merits, make it a valuable method for evaluating complex notions such as sustainability efficiency. Nevertheless, the classic models possess their own demerits, and hence, researchers in the field have developed new models with internal structures (network models) aiming to overcome them. Finally, in the case of sustainability measurement using DEA, it is highlighted that there is a requirement for a widely accepted definition and that the social dimension is underrepresented.

The following tables summarize the findings from the recent literature, grouped by geographical region of application.

Table 1 Summary of the new research about DEA and sustainability – Applications in Europe

Work	Input	Intermediate	Output	Index	DEA variation	Combination with other method	Area of application
Kiani Mavi, Kiani Mavi (2021)	Researchers in R&D (per million people), R&D expenditure(% OF GDP)	Eco patents	High-technology exports, Energy Productivity, Resource productivity	Eco-innovation	dynamic DEA	goal programming	Eu countries
Łącka, Brzezicki (2022)	Labor force, energy consumption	GDP, GHG emissions	Clean water and sanitation, affordable and clean energy	Eco-Efficiency, eco-innovation, SDGs	Dynamic network slack based measure DEA	Malmquist index	EU countries
Lubsanova, Maksanova, Eremko, Bardakhanova, Mikheeva (2022)	Total annual emissions of pollutants into the atmospheric air from stationary sources, Total annual emissions of pollutants into the atmospheric air from mobile sources,Volumes of non-treated or non-sufficiently treated wastewaters that were discharged into water bodies,		Gross regional product in current prices, resident population	Eco-efficiency	Slack based measure DEA		Russian regions

	amount of land field waste, volume of fresh water abstraction						
Moutinho, Madaleno (2021)	Gross fixed capital formation (formerly gross domestic fixed investment), Labor per capita, Energy use/area, electricity/area, Deviations temp		GDP pc/(GHG/area)	Eco-efficiency	Classic dea	FRM	Eu countries
Stanković, Marjanović, Stojković (2021)			Employment rate, Medium equivalized net income, Gdp per capita, people at risk of poverty or social exclusion	Socio-economic Efficiency	Classic DEA		European countries
Tsaples, Papathanasiou, Georgiou (2022)	Gross fixed capital at current prices, Total Labor force, population, gross electricity production, GDP per capita in PPS Index	GDP per capita in PPS Index, final energy consumption, Total expenditure per euro habitant	Median equivalized net income, final consumption expenditure of households, Terrestrial protected area (km2), Share of renewable energy in gross final energy consumption (%), Greenhouse gas emissions	Sustainability index	Two-stage DEA	Machine learning	Eu countries

			(in CO2 equivalent), Patent applications to the European patent office (EPO) by priority year, Overall life satisfaction, Satisfaction with living environment, Percentage of females in total labor population				

Table 2 Summary of the new research about DEA and sustainability – Applications in Asia

Work	Input	Intermediate	Output	Index	DEA variation	Combination with other method	Area of application
Chen, Kourtzidis, Tzeremes, Tzeremes (2022)	Capital, labor, energy	GDP	CO2; SO2	Environmental Sustainability index	multiplicative relational network DEA	Window analysis, Wilcoxon–Mann–Whitney rank-sum test	Chinese regions
Qu, Wang,	Capital, Labor, Energy, environmental investment, health investment	Waste water, solid waste, untreated waste water, untreated solid waste	GDP, mortality rate, survival rate	Sustainability	Super efficiency network DEA		Chinese regions

Liu (2022)							
Shah, Hao, Yan, Yasmeen, Lu (2022)	Capital, labor, energy		CO2 emissions, GDP	Energy efficiency	Slack based measure undesirable output DEA	Malmquist index, Mann Whitney U and Kruskal–Wallis test	South Asian countries
Sun, Wang (2021)	Labor force, capital		GDP, industrial SO2 emissions, industrial soot emissions, total waste water volume	Eco efficiency	DEA super-efficient slack-based measure	Malmquist productivity index, Tobit regression	Chinese cities
Sun, Wang, Ortiz, Huang, Zhao, Wang (2022)	Investment, population, electricity, land, water	GDP, industrial dust production, industrial SO2 production	Number of middle school students, Insurance covered, Social welfare, industrial dust emission, industrial SO2 production	Eco-efficiency	Slack based measure network DEA		Chinese cities
Wang, Nguyen, Chang (2021)	Energy consumption from coal, oil, gas sources Volume of Vehicles		GDP, CO_2 Emissions, CH_4 Emissions	Environmental Efficiency, eco-efficiency	SBM bad-output model		Asian countries

Wang, Wang, Yao (2021)	Labor, capital, energy		Real GDP, SO2, CO2	energy efficiency	Super efficiency DEA	Theil index	Chinese regions
Zhang, Zhao, Zha (2021)	Labor, Investment of fixed assets, energy	Volume of pollutants generated in industrial waste gas, treatment investment of industrial waste gas, capacity of industrial waste gas treatment(previous year)	Industrial value added, capacity of industrial waste gas treatment(current year), volume of pollutants emission in industrial waste gas	Efficiency	Sbm Two stage DEA		Chinese regions
Zuo, Guo, Li, Cheng (2022)	Full time equivalent of r&d personell, R&D internal expenditures	Number of patent application authorizations, number of scientific papers published, Labor, Fixed investment, energy	Total industrial output value of mining industry, solid waste discharge, exhaust emissions, waste water discharge	mining Technological innovation efficiency, mining eco-efficiency, mining comprehensive efficiency	Two stage DEA		Chinese regions

Table 3 Summary of the new research about DEA and sustainability – Applications in various countries or regions

Work	Input	Intermediate	Output	Index	DEA variation	Combination with other method	Area of application
Afzalinejad (2021)	Capital, labor		GDP, life satisfaction, Greenhouse gases: Total emissions excluding LULUCF	Efficiency	modified Radial DEA		various countries
Alidrisi (2021)	CO2, unemployment level		Manufacturing value added to GDP, charges for the use of intellectual property in terms of payments, charges for the use of intellectual property in terms of receipts, Value of high tech exports, number of R&D researchers, number of scientific and technical journal articles published, the number of patent applications filed, the	Efficiency-Based Global Green Manufacturing Innovation Index	Classic DEA		Various countries

			number of trademark applications filed				
Camioto, Pulita (2022)	CO2 emissions, energy use, percentage of unemployed		GDP, life expectancy at birth	SD Efficiency	Slack based measure DEA		Various countries
da Silva, Ribeiro, Rego (2023)	Levelized cost of energy, socio-economic vulnerability index		Number of direct employees, GDP, avoided CO2	Socio-economic efficiency	Non decreasing returns to scale dea		Brazilian regions
Fathi, Ashena, Anisi (2022)	Capital inventory, labor force, energy consumption	GDP, CO2 emissions, CO2 emission intensity index	equity index	E4 Efficiency	two-stage DEA	A nash bargaining game	Various countries
Goto,and Sueyoshi (2023)	Capital, Labor, Energy		GDP, CO2-solid emissions, CO2-gas emissions, CO2-other emissions, CO2-liquid emissions	Efficiency	Dea-ea		Various countries
Lu, Chiu, Chiu, Chang (2022)	Agricultural labor, Agricultural land use, nitrogen fertilizer, industrial employment, industrial energy use	Agricultural output, Industrial output, Agricultural methane CH4, Industry CO2 emissions	Average temperature change, Average rainfall amount, Disaster occurrence	sustainability efficiency	Dynamic parallel three-stage network DEA model		Worldwide countries

Sarpong, Wang, Cobbinah, Makwetta, Chen (2022)	Capital, Labor, Energy		GDP, CO2 emissions	Energy efficiency	Dynamic SBM DEA	Malmquist productivity index, Tobit regression	African countries
Yousefi, Hassanzadeh,Saen, Mousavi Kashi (2021)	Import of goods and services, Fossil fuel consumption		Gdp per capita, Compulsory education duration, CO_2emissions	Sustainability	Inverse RAM DEA		Islamic countries
Zhou & Xu (2022)	Capital, labor, energy	GDP, CO2 emissions	Re-adjusted GDP, Re-adjusted CO2 emissions	Energy efficiency	three-stage Undesirable-SBM-DEA	Malmquist index, SFA	RCEP member states

2.2.1 Sustainability Indicators

The tables above concern papers that were published from 2021 to 2023. One of the most intriguing aspects of the recent literature is the variety of indicators employed by researchers to measure sustainability, sustainable development, and related notions of efficiency. As already mentioned, sustainability includes three interconnected pillars, namely economic, environmental, and social. Therefore, it is imperative for researchers to utilize indicators that are pertinent to each of these dimensions in order to produce the best possible results.

Upon reviewing the extant literature, it becomes apparent that the most commonly employed economic indicators are capital and labor force, serving as inputs or intermediates, and gross domestic product (GDP) as an output or intermediate. Specifically, capital has been identified as an economic variable in 10 out of the 25 analyzed papers (40%), while the labor force indicator has been featured in 14 of them (56%). This fact highlights the significant role that these economic variables play in shaping economic outcomes.

It is unsurprising that capital and labor, as fundamental factors of production in economic theory, are frequently used in the reviewed literature. Moreover, indicators such as GDP, Real GDP, and Gross Regional Product appeared 18 times (72%) in the reviewed papers, highlighting their widespread acceptance among researchers as a means of measuring economic efficiency at the national or regional level. This finding underscores the significance of these indicators in evaluating economic performance.

It is noteworthy that while the classic indicators of labor, capital, and GDP are not always employed directly in the reviewed studies, authors frequently utilize related indicators. Specifically, variables such as investment, gross fixed capital formation, employment rate, and unemployment level are commonly employed. These indicators are closely related to the three fundamental indices and provide insight into economic performance from alternative perspectives (Zhou et al., 2018).

In addition to economic indicators, environmental variables have also been widely employed by researchers in their efforts to measure sustainability, both in older and more recent literature. Specifically, indicators such as energy use or consumption, and industrial energy use, are frequently utilized. Similarly, carbon dioxide (CO_2) emissions or industrial CO_2 emissions are commonly employed. These environmental variables provide insight into the environmental impact of economic activity and are crucial for assessing sustainability.

In particular, in the 25 studies analyzed, these variables have appeared 16 (64%) and 13 (52%) times correspondingly, and as a result, it is inferred that they constitute two of the most significant environmental indicators in order to assess the environmental efficiency of a country or a region.

Apart from the environmental indicators contained in the recent literature of studies that use DEA to measure efficiency, eco-efficiency, or sustainable efficiency, there is a need to recognize the indicators used to express the social dimension of sustainability. As mentioned before, Tsaples and Papathanasiou (2021) pointed out in their literature review the lack of social indicators in the DEA models that were applied in the period from 2016 to 2020 and tried to measure regional or country efficiency in terms of sustainability.

This research gap has prompted several researchers to incorporate more social indicators in their studies, recognizing the importance of addressing all three pillars of sustainability. In both the reviewed articles and older literature, there are indicators that combine social and economic information, such as gross fixed capital at current prices in purchase power standards (PPS), GDP per capita in PPS index, the employment rate, and the medium equivalized net income. However, purely social indicators such as life satisfaction and social welfare remain underrepresented in these studies.

The analysis of DEA models employed for measuring sustainability efficiency during the years 2021-2023 has led to the observation that an increasing number of social indicators have been integrated into these models in an attempt to assess sustainability accurately. To achieve this objective, all researchers in this domain must support this approach and broaden the range of social indicators incorporated into the relevant literature.

Regarding the papers under review, only 40 percent of them, namely 10 out of 25, included social indicators, while the others neglected the social dimension of sustainability by using only economic and/ or environmental variables.

The last group of indicators that appears in the studies under review with the purpose of measuring sustainability are the research and development (R&D) indicators. The emergence of research and development (R&D) indicators as a group of metrics for measuring sustainability in recent literature can be attributed to several factors. One possible explanation is that these indices are increasingly used in contemporary studies due to the belief that R&D can contribute to both economic growth and social welfare while minimizing environmental degradation.

This perspective is grounded in the notion that technological advancements and innovation can facilitate the attainment of sustainable development goals by promoting resource efficiency, reducing waste, and mitigating environmental risks. Consequently, the incorporation of R&D indicators into sustainability assessments is a promising avenue for advancing the understanding of sustainable development pathways and informing policy decisions.

Although the current body of literature on sustainability efficiency measurement primarily employs a limited number of research and development (R&D) indicators, with only four out of 25 analyzed articles incorporating such metrics, this trend is expected to evolve in the coming years. As researchers continue to explore the role of R&D, technology, and innovation in sustainable development, it is likely that an increasing number of variables relevant to these domains will be incorporated into regional and country-level sustainability efficiency assessments. This development is significant as it underscores the growing recognition of the critical role that R&D plays in promoting sustainable development pathways and highlights the need for more comprehensive and nuanced approaches to measuring sustainability efficiency.

In this part of the literature review, each of the 25 selected articles that utilize data envelopment analysis (DEA) for measuring sustainability or similar notions of efficiency will be subjected to a detailed analysis. The objective of this analysis is to identify the specific sustainability indicators employed by the authors to construct their final efficiency index, which is typically used to evaluate the performance of countries or regions.

Starting with the paper of Tsaples et al. (2022), the authors proposed a two-stage DEA model with deviational variables in the objective function. They employed that model to create three sub-indicators, one for each dimension of sustainability. Afterwards, these sub-indicators are aggregated into a final sustainability index, which is eventually used to measure the relative sustainability efficiency of the 28 European Union countries. The authors used as inputs in their model the gross fixed capital at current prices, the total labor force, the population, the gross electricity production, and the GDP per capita in purchasing power standards (PPS) index. The GDP per capita in PPS index, the final energy consumption, and the total expenditure per euro habitant were used as the intermediate variables of the study. Regarding the output variables of the model, the authors used the following:

1. The medium equivalized net income
2. The final consumption expenditure of households
3. The Terrestrial protected area
4. Share of renewable energy in gross final energy consumption
5. The Greenhouse gas emissions (in CO2 equivalent)
6. The Patent applications to the European Patent Office (EPO) by priority year
7. The overall life satisfaction
8. Satisfaction with living environment
9. And the percentage of females in the total labor population.

In addition to the traditional DEA approach, this study also explores the application of machine learning techniques, specifically Classification and Regression Trees (CART) and boosting regression, to the results of DEA computations. The aim of this exercise is to investigate how the sustainability efficiency of countries behaves under different scenarios.

This innovative approach is significant as it enables a more nuanced and dynamic understanding of sustainability efficiency beyond a simple ranking or classification of countries based on their overall efficiency scores. By utilizing machine learning techniques, this study seeks to uncover the underlying relationships and patterns that govern sustainability efficiency and provide insights into the factors that drive or hinder sustainable development pathways.

Moreover, Moutinho and Madaleno (2021) applied a classical constant return to scale (CRS) DEA model to create an eco-efficiency index and concurrently assess the efficiency of 27 European countries. The authors included several economic and social indicators in their study. More specifically, they used Gross fixed capital formation (formerly gross domestic fixed investment) and labor per capita as economic inputs, while their environmental inputs were energy use/area, electricity/area, and the Deviation temp.

Finally, the composite index of the ratio of the value of gross domestic product per capita and the value of the volume of the GHG emissions by area of a given European country (GDP pc/(GHG/area)) was the output variable of their model. The researchers also used a Fractional Regression Model to analyze the relationship between the scores of their

DEA model and some possible influencing factors, in particular, eight different types of pollutants.

Kiani Mavi and Kiani Mavi (2021) developed a dynamic DEA common set of weights model, which includes a novel goal programming technique in order to avoid weight flexibility. This model is used to aggregate various indicators into an eco-innovation index, which aims to evaluate comparatively the efficiency of the EU-27 countries. The number of researchers in R&D (per million people) and the R&D expenditure expressed as a percentage of GDP were the input variables of this study. The patents in environment-related technologies served as a proxy for eco patents and were used as the intermediate variable, while the corresponding outputs of the model were the high-technology exports, the energy productivity, and the resource productivity.

Furthermore, Stanković et al. (2021) applied a classic variable returns to scale (VRS) DEA model with the purpose of creating a composite socio-economic index, which aims to measure the socio-economic efficiency of EU28 countries. The variables contained in the model entailed both economic and social information and were comprised of the following indicators:

1. The employment rate
2. The medium equivalized net income
3. The GDP per capita and
4. The number of people at risk of poverty or social exclusion, which is described as the percentage of people who are either at risk of poverty or severely materially deprived or living in a very low work intensity household.

Lacka and Brzezicki (2022) employed a Dynamic network slack-based measure DEA model, aiming to create an eco-efficiency index, an eco-innovation index, and an SDGs index to assess the corresponding efficiencies of European Union countries. They also employed a Dynamic division Malmquist productivity index (DDMI), which was first proposed by Tone and Tsutsui (2017), with the aim of explaining the productivity changes in divisions. The indicators that were included by the authors in order to produce the indices mentioned above were labor force and energy consumption as input variables, which produce the eco-efficiency index; GDP and GHG emissions as intermediate variables, which produce the eco-innovation index and as output variables the clean water and sanitation alongside affordable and clean energy, which constitute Europe's SDG 6 and SD7 respectively (Eurostat, 2023) and produce the SDGs index.

Lubsanova et al. (2022) applied a non-radial and non-oriented slack-based measure DEA model in order to develop an eco-efficiency index and eventually evaluate the relative eco-efficiency of some north Russian regions. In their attempt to create this index, the authors aggregate numerous indicators representing economic, environmental, and social dimensions. More precisely, the following indicators were utilized as inputs into the model:

1. The total annual emissions of pollutants into the atmospheric air from a stationary source

2. The total annual emissions of pollutants into the atmospheric air from mobile sources

3. The volumes of non-treated or non-sufficiently treated wastewater that were discharged into water bodies

4. The amount of land field waste

5. And the volume of freshwater abstraction

Simultaneously, they included the gross regional product at current prices alongside the resident population as output variables.

Wang et al. (2021) proposed an extended SBM bad-output DEA model with the aim of developing an environmental efficiency/ eco-efficiency index and eventually evaluating the corresponding relative efficiency of the top 20 Asian economies. In order to create that index, the authors included as input variables:

1. The energy consumption from coal
2. The indicator of oil, as well as
3. Gas sources volume of vehicles.

Finally, GDP, CO_2 emissions, and CH_4 emissions were used as output variables in the model.

Moreover, Qu et al. (2022) developed a modified super-efficiency radial network DEA model without infeasibility. The authors applied this model with the purpose of constructing a composite sustainability index in order to assess the regional sustainability performance in China. Capital, labor, energy, environmental investment, and health investment were contained as inputs in the model. Furthermore, wastewater, solid waste, untreated wastewater, and untreated solid waste constitute the intermediate variables of the study, while GDP, mortality rate, and survival rate constitute the respective outputs.

Concerning the work of Sun et al. (2022), the authors aim to measure the urban sustainable development in 284 cities in China. In order to do so, they propose a SBM network DEA model. In the first stage, which is referred to as the production stage, investment, population, electricity, land, and water constitute the input variables, while the last three serve as proxies for the natural resources used in urban development. The output variables of this stage, which are simultaneously the inputs of the second stage, are the GDP as undesirable output and the industrial dust production alongside industrial SO2 production as undesirable outputs.

In addition, the authors used three desirable outputs in the second stage, namely the number of middle school students, the insurance covered, which is calculated as the average of the number of pension and unemployment insurance covered, and the social welfare, which is computed by the aggregation of three sub-indicators, namely, the public transit, the number of doctors and the green area. Eventually, industrial dust emission and industrial SO2 production were the undesirable outputs of the model.

Furthermore, Chen et al. (2022) applied a multiplicative relational network DEA model with the aim of producing a composite sustainability index, which consists of a Production Efficiency index and an Eco-efficiency index, and eventually evaluating the regional efficiency in China for the years between 2000 and 2012. They also employed a window analysis of the multiplicative efficiency decomposition approach as well as a Wilcoxon–Mann–Whitney rank-sum test to discover additional results. The specific variables used in the model in order to create the sustainability index are the indicators of capital, labor, and energy as inputs, the GDP as an intermediate variable, and the CO2 emissions with the SO2 emissions as the undesirable outputs of the final stage.

Sun and Wang (2021) employed a super-efficient DEA model with a slack-based measure aiming to develop an eco-efficiency index and to assess the efficiency of China's Loss Plateau Region relatively. The authors also used, alongside DEA, a Malmquist productivity index to measure productivity change and the entropy-weighted TOPSIS model to weigh the variables objectively.

Furthermore, they employed a Tobit regression model to discover factors outside of those included in the model that influence eco-efficiency. The authors employed the indicators of labor force and capital as input variables in their work, while the respective outputs were the GDP, the industrial SO2 emissions, the industrial soot emissions, and the total wastewater volume of the cities under assessment.

Wang et al. (2021) applied a super-efficiency DEA Model to create an energy efficiency index and assess the level of the corresponding regional efficiency in China. They also used the Theil index for the purpose of dividing the regional energy efficiency differences of the assessed regions into differences within the region and differences between the regions. The authors included as input variables in their model the most common set of indicators among the under-review studies, namely the indicators of labor, capital, and energy, while the real GDP was the desirable output and the SO2 emissions together with the CO2 emissions constitute the undesirable outputs.

Moreover, Zhang et al. (2021) aggregate several indicators into a final efficiency index with the purpose of evaluating comparatively the efficiency of numerous Chinese regions. In order to achieve that, the authors applied a two-stage DEA model with a slack-based measure that consists of several variables. More precisely, labor, investment of fixed assets, and energy are the inputs of the first stage, while industrial value added is the desirable output variable of the same stage.

The volume of pollutants generated in industrial waste gas constitutes the undesirable output and simultaneously the input of the next stage. Furthermore, treatment investment of industrial waste gas and the capacity of industrial waste gas treatment in the previous year are the other inputs of the second stage. Finally, the capacity of industrial waste gas treatment in the current year is the desirable output, and the volume of pollutants emission in industrial waste gas is the undesirable output of the last stage.

In the paper of Zuo et al. (2022), the authors applied a two-stage DEA model in order to create three different indicators to evaluate the efficiency of 30 Chinese provinces. The first stage of the model calculates the mining technological innovation efficiency, and the second stage calculates the mining eco-efficiency. Finally, the mining comprehensive efficiency is defined as the square geometric mean of the efficiency in each stage.

The indicators included in their model were the full-time equivalent of R&D personnel alongside the R&D internal expenditures as input variables, and the number of patent applications authorizations, the number of scientific papers published, labor, energy, and fixed investment were used as intermediates. Ultimately, the following four indicators are used as outputs:

1. The total industrial output value of the mining industry
2. The solid waste discharge
3. The exhaust emissions and
4. The wastewater discharge.

Yousefi et al. (2021) developed an inverse RAM DEA model in order to measure the sustainability efficiency of several Islamic countries and to suggest strategies to improve it. The modified model determines the optimal inputs and outputs under managerial disposability and natural disposability so that the efficiency scores of the DMUs will remain the same. The authors contained in their model the imports of goods and services as managerial input and the fossil fuel consumption as natural input, while GDP per capita, together with compulsory education, were the desirable outputs and the indicator of the CO2 emissions, served as the respective undesirable output.

Additionally, Lu et al. (2022) employed a modified dynamic parallel three-stage network DEA model to construct a sustainability index and evaluate the corresponding efficiency of several countries around the world. The authors included agricultural labor, agricultural land use, nitrogen fertilizer, industrial employment, and energy use as inputs in the first stage of the models. The intermediate variables of the model were the agricultural output, the industrial output, the agricultural methane CH4, and the industry CO2 emissions, while the average temperature change, the average rainfall amount, and the disaster occurrence were the outputs. Finally, the indicator of the agricultural fixed assets constitutes the carry-over variable of t to the t+1 period.

Aldirisi (2021) applied a classic input-oriented constant returns to scale (CRS) DEA model to construct a global green manufacturing innovation index and measure the respective comparative efficiency of the top 15 manufacturing countries worldwide. In order to do so, the author used as input variables in his model the indicators of CO2 emissions alongside the unemployment level, while the outputs contained in his study are the following:

1. The Manufacturing value added to GDP
2. The charges for the use of intellectual property in terms of payments

3. The value of high-tech exports
4. The number of R&D researchers
5. The number of scientific and technical journal articles published
6. The number of patent applications filed and
7. The number of trademark applications filed

Regarding the paper of Camioto and Pulita (2022), the authors created a composite sustainable development index by employing a variant SBM DEA model. They generated relative sustainable development efficiency rankings for BRICS and G7 countries. The CO2 emissions, energy use, and the percentage of unemployed were used as inputs. At the same time, the two output variables of the study consist of the GDP and the life expectancy at birth.

Goto and Sueyoshi (2023) utilized a DEA-EA model to evaluate the relative efficiency of numerous countries worldwide. More specifically, their work measures the degree of unified index and that of unified efficiency, both under managerial disposability and constant damages to scale (DTS). The authors used the indicators of capital, labor, and energy as input variables in their model. At the same time, they used GDP as desirable output alongside four different undesirable outputs, namely CO2 emissions from solid fuel consumption (mainly from coal), CO2 emissions from liquid fuel consumption (mainly from petroleum-derived fuels, CO2 emissions from gaseous fuel consumption (mainly from natural gas) and CO2 emissions from the other sources.

Moreover, Sarpong et al. (2022) applied a dynamic non-oriented SBM DEA model to develop a composite index and simultaneously assess the relative energy efficiency of 9 West African countries. The authors also utilized the Malmquist productivity index in order to measure the technology change improvement from 2007 to 2020, as well as a Tobit regression model to ascertain the relationship between the dependent variable, namely the respective DEA efficiency score, and the independent variables (or the influencing factors), such as the gross national income, the population, the final consumption expenditure, the human development index, the foreign direct investment, and the urbanization population. Finally, the indicators included in this work were capital, labor, and energy as inputs, the GDP as desirable output, and the CO2 emissions as undesirable output.

Da Silva et al. (2023) applied an output-oriented Non-Decreasing Returns to Scale (NDRS) DEA model to measure the socio-economic efficiency of several Brazilian mesoregions where electricity generation facilities are installed. In order to do so, the authors utilized numerous socio-economic and environmental indicators in their paper. More precisely, the levelized cost of energy, which is computed as the sum of a facility's Capital Expenditure (CAPEX), the Operational costs to operate it (OPEX), and decommissioning costs, discounted to present-day value, divided by the electricity supplied to the grid throughout the operational life of the technology, together with the socio-economic vulnerability index constitute the inputs of the DEA model. Finally, the output variables consist of the number of direct employees, the GDP, and the avoided CO2 emissions.

Furthermore, Shah et al. (2022) utilized an undesirable output SBM DEA model to evaluate the comparative energy efficiency of 6 South Asian countries, while they also employed a Malmquist productivity index to calculate the energy efficiency change from 2001 to 2019. In addition, in their attempt to estimate the impact of the energy policy of 2010 over the study period, the statistical significance of the difference in mean scores for energy efficiency and productivity over two time periods (2001–2010 and 2011–2019) and six countries were examined, utilizing the Mann–Whitney U and Kruskal–Wallis tests. Concerning the indicators of this study, capital, labor, and energy constitute the input variables, while the CO2 emissions and the GDP were the undesirable output and the desirable output, respectively.

Afzalinejad (2021) developed in his paper a new modified radial DEA model, which takes into consideration undesirable outputs by separating the assessment of operational and environmental efficiency, aiming to develop a final efficiency index and to measure the relative efficiency of 28 countries around the world in economic, social and environmental dimensions. In order for this composite efficiency index to be constructed, the author used the capital alongside the labor force as input variables, while the indicators of the GDP, life satisfaction, and greenhouse gas emissions excluding emissions from land-use, land-use change and forestry (LULUCF) constitute the set of output variables.

Zhou and Xu (2022) proposed a modified three-stage undesirable output SBM DEA model to measure the comparative energy efficiency of the Regional Economic Comprehensive Partnership (RCEP) members. In the first stage, the undesirable SBM DEA model is employed to estimate the value of energy efficiency and the slack variables in the case of original input and output. Input variables consist of capital stock, labor, and primary energy consumption, while GDP alongside CO2 emissions constitute the desirable and the undesirable output, respectively. In the second stage, a stochastic frontier analysis regression (SFA) model is applied to expunge the effect of exterior differences, while in the third stage, the re-adjusted data that arise from the SFA were used in the radial undesirable SBM DEA model to produce the final efficiency scores. In addition, the authors utilized the Malmquist productivity index model to investigate further the dynamic changes in the energy efficiency of the countries under assessment.

The last of the papers under review is that of Fathi et al. (2022). In this study, the authors applied a network two-stage DEA model to evaluate the E4 efficiency of various countries worldwide, where E4 stands for the connection between energy, environment, economy, and equity. Capital inventory, labor force, and energy consumption constitute the inputs of the first stage, while GDP and CO2 emissions are the output variables of the same stage. Moreover, the CO2 emission intensity index, which equals the CO2 emissions divided by energy consumption, is the input of the second stage, while the equity index, for which economic welfare is used as a proxy, constitutes the respective output variable. Finally, the authors utilized a Nash bargaining game model, based on the work of Nobelist Mathematician John Nash (1950, 1953), to measure the efficiency of the network structure and simultaneously accomplish a fair efficiency decomposition for both stages.

To summarize, the articles reviewed in this analysis utilized indicators to assess sustainable development, which can be categorized into four main groups: economic, social, environmental, and research and development. However, the majority of the studies focused primarily on economic and environmental variables, leading to an inadequate representation of the social dimension of sustainability. Nevertheless, some authors have attempted to address this issue by introducing additional social indicators in recent literature. Lastly, it can be inferred that the number of R&D, technology, and innovation indicators is on the rise as researchers aim to incorporate new perceptions of sustainability.

2.2.2 Notions of Efficiency

In the literature that is currently under review, the primary objective of the studies is to utilize DEA methodology to aggregate various indicators and measure the efficiency of different countries or regions. However, it is noteworthy that each author has a distinct interpretation of efficiency. The table below summarizes the different notions of efficiency that are used in recent studies.

Table 4 Frequency of appearance of different notions of efficiency

Index	Frequency
Sustainability	4
Environmental sustainability	1
SD Efficiency	1
Eco-efficiency	4
E4 Efficiency	1
Eco-Efficiency, eco-innovation, SDGs	1
Environmental Efficiency, eco-efficiency	1
Efficiency	3
Socio-economic Efficiency	2
Eco-innovation	1
Efficiency-Based Global Green Manufacturing Innovation Index	1
Mining Technological innovation efficiency, mining eco-efficiency, mining comprehensive efficiency	1
Energy efficiency	4

Upon examining the table above, it is clearly noticed that the authors use numerous terms to describe sustainability or sustainable development. In comparison with earlier studies, there has been an upward trend in the frequency of appearance of sustainability indices. More specifically, 7 out of the 25 above papers measure efficiency in terms of sustainability, environmental sustainability, Sustainable Development, or SDGs.

Moreover, another index that has appeared multiple times in the recent literature is the eco-efficiency index. In particular, it appears four times on its own and three times alongside other indices such as eco-innovation, environmental efficiency, and mining technological innovation index, among others.

According to Schaltegger & Muller (1996), eco-efficiency is a concept of economic and ecological efficiency, which first appeared in the nineties as a practical approach to the more encompassing notion of sustainability, while Zhang et al. (2008) refer to eco-efficiency as an instrument for sustainability analysis.

The World Business Council for Sustainable Development (WBCSD) described eco-efficiency as: "The delivery of competitively priced goods and services that satisfy human needs and bring quality of life, while progressively reducing ecological impacts and resource intensity throughout the life-cycle to a level at least in line with the Earth's estimated carrying capacity" (WBCSD, 1992, p. 3).

The indicators developed to represent the notion of eco-efficiency are based on ratios that relate the economic value of goods and services produced to the environmental pressures or impacts caused by production processes. The larger the ratio, the higher the level of eco-efficiency. (see Schmidheiny & Zorraquin (1996), Figge & Hahn (2004), Huppes & Ishikawa (2005).

Furthermore, according to Huppes and Ishikawa (2005), the term eco-efficiency can be calculated in real life in four different ways:

1. As the ratio of economic output to environmental pollution, which is called environmental productivity

2. As the ratio of environmental pollution to economic activity, which is called environmental intensity

3. As the ratio of improvement cost to environmental improvement, which is called environmental improvement cost

4. As the ratio of environmental improvement to improvement cost, which is called environmental cost effectiveness.

Although the eco-efficiency concept does not contain the economic dimension of sustainability, its efforts could improve financial, environmental, and social performance (Alves & Dumke De Medeiros, 2015) for businesses and for countries or regions. For that reason, researchers of the older and the newer literature have adopted it as a proxy for the sustainability index.

Another common index in the reviewed papers is energy efficiency, which can be found in 4 out of 25 studies. Energy efficiency constitutes a generic term, and so there is no strictly defined quantitative way to calculate it.

According to Patterson (1996), energy efficiency refers to using less energy to produce the same amount of services or valuable output. For instance, in the industrial sector, energy efficiency can be computed by the amount of energy required to produce a tonne of product. Thus, it can be defined by the following ratio :

$$\frac{useful\ output\ of\ a\ process}{energy\ input\ into\ process}$$

Although energy efficiency is not exactly used as a proxy for sustainability by the authors, and hence, these studies could be excluded from the review, the two terms are closely related to each other. In their study, Zakari et al. (2022) aimed to investigate the connection between energy efficiency and sustainable development goals for 20 Asian and Pacific countries using DEA. The authors found through the Panel Correction Standard Error (PCSE) estimates that sustainable economic development and sustainable financial development are both closely related to energy efficiency.

Another reason to encompass these studies in the current review is the set of indicators that they use to measure energy efficiency. More specifically, all of these studies include economic, environmental, and socioeconomic indices similar to the rest of the studies, such as labor, capital, GDP, and real GDP, to express the economic as well as the socioeconomic dimension, whereas energy, CO2 emissions, and SO2 emissions are used to express the environmental dimension.

One more notion that has appeared multiple times in recent literature is the eco-innovation. The eco-innovation index appears in one study on its own and one time alongside other indices. In addition, two studies use similar indices, namely mining Technological innovation efficiency and the Efficiency-Based Global Green Manufacturing Innovation Index.

Oltra and Saint Jean (2009) focus on the effect of eco-innovation and define it as innovations that consist of new or modified processes, practices, systems, and products that benefit the environment and thus make a contribution to environmental sustainability. Rennings (2000) gives a definition focused on the motivation of eco-innovation and describes it as the innovation processes towards sustainable development.

European Commission (2007) mentions that eco-innovation is any form of innovation that aims to make significant and provable progress toward the goal of sustainable development through reducing negative impacts on the environment or achieving more efficient and responsible use of natural resources, including energy.

The above definition implies that the goal of eco-innovation is to create new goods, processes, and practices with the purpose of promoting sustainable development. Hence, it is a closely related term to sustainability. Moreover, in order for somebody to innovate, the existence of Research and Development activities is a necessary condition. This fact could be one more explanation of why the authors in the recent literature encompass more and more R&D indicators in their attempt to measure efficiency in terms of sustainability.

Other indices that are found are the more general efficiency index in three studies, the socio-economic efficiency, which takes into account the social and economic dimensions in two studies, and the E4 efficiency in one study. The term E4, as explained by Fathi et al. (2022) in their paper, stands for the relationship of energy, environment, economy, and equity.

Upon examining this section, it becomes clear that the authors utilize several diverse definitions of sustainability in the context of DEA. This phenomenon leads to the use of different sets of indicators in each study and consequently decreases the robustness of the obtained results. Hence, this fact highlights the requirement to set a commonly accepted definition of sustainability (and a widely accepted set of indicators) in the field of DEA, aiming to improve the reliability of the outcomes.

2.2.3 DEA variations employed

After the identification of the various notions of efficiency that are used in the reviewed papers, it would be interesting to explore which different DEA variations are employed by the authors of the recent literature in order to measure sustainability.

The following table establishes the frequency of appearance of every DEA variation in the under-review studies.

Table 5 Frequency of appearance of different DEA variations employed

DEA VARIATION	FREQUENCY
CLASSIC DEA	3
TWO-STAGE DEA	3
SBM TWO STAGE/NETWORK DEA	2
SBM BAD/UNDESIRABLE OUTPUT MODEL	2
SBM DEA	2
DYNAMIC DEA	1
DYNAMIC NETWORK SBM DEA	1
SUPER EFFICIENCY NETWORK DEA	1
A MULTIPLICATIVE RELATIONAL NETWORK DEA	1
DEA SUPER-EFFICIENT SBM	1
SUPER EFFICIENCY DEA	1
INVERSE RAM DEA	1
PARALLEL THREE-STAGE NETWORK DEA MODEL	1
DEA-EA	1
DYNAMIC SBM DEA	1
NDRS DEA	1
MODIFIED RADIAL DEA	1
THREE-STAGE UNDESIRABLE-SBM-DEA	1

From the above table, it is easily observed that the authors use a lot of different DEA variations. The most-employed variations are the classic DEA model, the two-stage DEA model, the slack-based measure (SBM) DEA model, the SBM two-stage/network DEA, and the SBM undesirable/BAD output DEA model from which the first two appear three times, while the others appear two times each, in the under-review literature. In addition, some variations in the reviewed papers are similar to the above but slightly modified, such as super efficiency DEA, super efficiency network DEA, SBM two-stage DEA, and dynamic SBM DEA.

Even though classic DEA models have been found appropriate for measuring efficiency by both older and recent literature, an effort is made from the newer studies to employ different, more sophisticated models. This phenomenon has led to the increasing use of models, such as slack-based measure models and network models, in the last few years.

Network models are trying to explain the stage where inputs convert into outputs, which is considered as a black box for the classic DEA models. For instance, in the current review, Tsaples et al. (2022) proposed a two-stage DEA model, which first calculates three sub-indicators to express the three dimensions of sustainability and then uses the benefit of the doubt model to calculate the overall sustainability index.

Another network model that has been employed is the parallel three-stage network DEA model by Lu et al. (2022). In that model, every DMU is composed of two sub-DMUs in parallel, which constitute the outputs of the initial stage and the inputs of the following stage simultaneously.

Regarding the slack-based measure models, Lubsanova et al. (2022) applied a non-radial and non-oriented SBM model developed by Tone (2001), aiming to measure the eco-efficiency of northern Russian regions. The authors support that, unlike the classical CCR and BBC models, the slack variables of the SBM model can better solve the relationship between inputs and undesirable outputs. Hence, the SBM model is particularly useful in measuring sustainability, sustainable development, or similar notions of efficiency that deal with undesirable outputs.

Moreover, Wang et al. (2021) in their study applied an undesirable output model proposed by Cooper et al. (2006), which modified SBM to account for undesirable outputs. The authors employed this model, which constitutes a modified version of Tone's model that was mentioned above, in order to measure the environmental efficiency and the eco-efficiency of the top 20 Asian economies.

Regarding the abovementioned, it is inferred that the authors of the recent literature use numerous variations of DEA aiming to measure sustainability, and as a result, there is a lack of a unified methodological framework. This phenomenon could decrease the reliability of the obtained results and consequently affect the policymaking that arises from them.

2.3 Information entailed from the literature

In this section, a synthesis of the reviewed literature on sustainability and Data Envelopment Analysis (DEA) will be presented, with a focus on highlighting the key findings and identifying any research gaps that emerged.

In the initial stages of this investigation, it became apparent that the concept of sustainability, or sustainable development, is subject to varying interpretations among scholars. The three-dimensional approach, which emphasizes the economic, environmental, and social pillars, has gained widespread acceptance. However, this lack of a universally accepted definition poses a challenge, as it can lead to confusion and miscommunication. Therefore, it is imperative to set a well-established and widely accepted definition of sustainability to ensure that all the parties involved understand the term in the same way.

In addition, even sustainability in the context of data envelopment analysis appears to have different definitions. The majority of studies focus on economic indicators such as labor force, capital, and gross domestic product (GDP), as well as environmental indicators like energy consumption and carbon dioxide (CO_2) emissions. This phenomenon has led to the underrepresentation of the social dimension and inaccurate outcomes in DEA studies. Recently, researchers have attempted to address this shortcoming by incorporating social indicators such as social welfare and life satisfaction into their models in order to achieve more comprehensive results.

Furthermore, another trend in the recent literature is the growing number of indicators related to R&D, technology, and innovation. This fact may be attributed to two possible reasons. Firstly, some researchers may argue that these factors play a crucial role in promoting sustainable development. This perspective is grounded in the notion that technological advancements and innovative solutions can facilitate the transition towards more sustainable practices and reduce environmental degradation.

Secondly, others may have attempted to incorporate new perceptions into the topic of sustainability by incorporating R&D, technology, and innovation indicators. The increasing number of those indices in the literature highlights the growing recognition of their significance in gauging sustainable development.

Moreover, in the current study, it was observed that a multitude of terms were used in various investigations in order to describe sustainability or sustainable development. While numerous studies utilize the terms sustainability or sustainable development, others utilize proxies such as eco-efficiency or eco-innovation. This fact has resulted in the use of disparate sets of input and output variables in these studies, which ultimately impact their final outcomes. This finding underscores the need for a widely acknowledged definition of sustainability and sustainable development in academic research and public policy fields.

Ultimately, in the under-review papers, numerous different DEA variations were identified. On the one hand, this phenomenon adds new perceptions to the subject, while it also helps to overcome some limitations of the classic DEA models, such as the treatment of undesirable outputs or the "black box" of the conversion process. On the other hand, it reveals the lack of existence of a standardized, widely accepted DEA framework for gauging sustainability efficiency. These differences in the DEA variations employed decrease the validity of the research outcomes and simultaneously affect the policymaking that is based on them.

3. Theoretical framework

3.1 Data Envelopment Analysis

Data envelopment analysis (DEA) is a data-oriented, non-parametric, mathematical programming technique that is used for the evaluation of the relative efficiency of a group of homogeneous entities. These entities, which are called Decision Making Units (DMUs), transform multiple inputs into multiple outputs. All the input-output correspondences that can be achieved by a DMU (regardless of whether these correspondences are observed in practice or not) constitute the Production Possibility Set (PPS). The term input stands for what every DMU consumes in order to produce the corresponding output, as can be seen in the figure below:

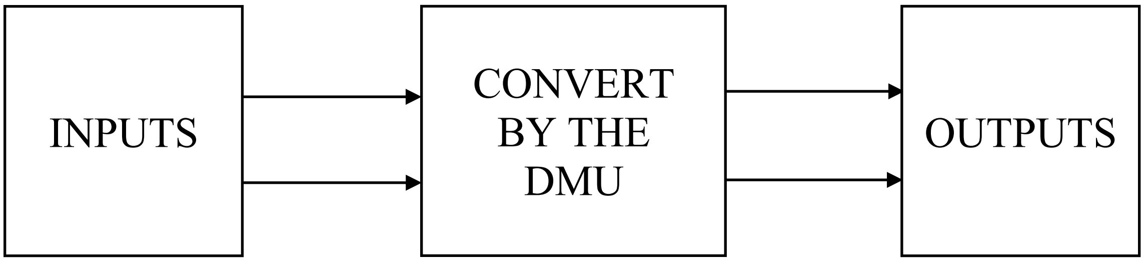

Figure 2 Classic DEA structure

(Thanassoulis, 2001)

The corresponding technical efficiency of any DMU is computed by forming the ratio of a weighted sum of outputs to a weighted sum of inputs, as indicated in the following expression:

$$technical\ efficiency = \frac{\Sigma w_{output} * y}{\Sigma w_{input} * x}$$

The weights (multipliers) of the outputs, as well as the inputs, have to be chosen in a way that computes the Pareto efficiency measure of each homogeneous entity subject to the constraint that none of these DMUs can have a relative efficiency rating bigger than 1 (technical efficiency ≤ 1). (Charnes et al., 2010)

There are two definitions that describe Pareto efficiency. One of them is considered output orientation, and it is suitable when outputs are controllable, while the other is labelled as input orientation, and it is suitable when inputs are controllable.

Let a set of homogeneous DMUs use one or more outputs to secure one or more inputs then:

Output orientation definition: A DMU is considered Pareto efficient if it is not feasible to increase any of its observed output levels without simultaneously decreasing at least another one of its output levels and/or without raising at least one of its input levels.

Input orientation definition: A DMU is considered Pareto efficient if it is not feasible to decrease any of its observed input levels without simultaneously increasing at least another one of its input levels and/or without lowering at least one of its output levels.

There are two more notions of efficiency relevant to the DEA method, and these are the input overall efficiency and the input allocative efficiency.

Input allocative efficiency: For given input prices, let C_{min} represent the minimum cost at which a DMU could secure its outputs and C_{te} the cost of its technically efficient input levels for its input mix. Then, the input allocative efficiency of the DMU is C_{min} / C_{te}

Input overall efficiency: With C_{min} being the same as above and C_{ob} representing the cost of the DMU's observed input levels, the input overall efficiency of the DMU is C_{min} / C_{ob}

If the ratio input overall efficiency / input allocative efficiency will be formed, then from the above definitions:

$(C_{min} / C_{ob}) / (C_{min} / C_{te}) = C_{te} / C_{ob}$ that equals Technical Input Efficiency.

Thus: Input Overall Efficiency / Input Allocative Efficiency = Technical Input Efficiency

or

Input overall efficiency = Technical Input Efficiency x Input Allocative Efficiency (Thanassoulis, 2001)

DEA as a methodology offers some serious advantages versus parametric methods, like regression analysis or stochastic frontier analysis:

1. Firstly, it does not need to identify the process of how inputs convert to outputs. It can calculate the technical efficiency of a DMU with multiple inputs and outputs, requiring only information about output and input quantities (not prices).

2. It requires less information than parametric methods; for example, it does not require a specific functional form that relates the independent variables to the dependent variable(s) or specific assumptions about the distribution of the error terms.

3. It can showcase why an entity (DMU) is inefficient and suggest plans to improve its efficiency. (He et al., 2016; Shi et al., 2010)

4. Possible sources of inefficiency can be determined, as well as efficiency levels. It gives a means of "dividing" economic inefficiency into technical and allocative inefficiency. Moreover, it also lets technical inefficiency be

decomposed into scale effects, the effects of unwanted inputs that the agency cannot dispose of, and a residual component.

5. By discovering the DMUs (in the current thesis, the countries or regions) that are not observed to be efficient, it provides a set of possible role models that a country or region can look to, in the first instance, for ways of improving its operations. This fact makes DEA a potentially helpful tool for benchmarking among those homogeneous entities.

The above Advantages of DEA are primarily based on the works of Papathanasiou et al. (2021), SCRCSSP (1997), Nunamaker (1985), and Tsaples (2022)

Nevertheless, DEA has its own disadvantages. Just like other empirical techniques, DEA is based on several simplifying assumptions that need to be taken into consideration in order to interpret its results precisely. Some of its main limitations, which are mainly based on Tsaples (2022), SCRCSSP (1997), and Nunamaker (1985), comprise the following:

1. Being more a deterministic rather than statistical technique, DEA generates scores that are especially sensitive to measurement error. If the inputs of a DMU are understated or its corresponding outputs are overstated, then this specific unit can become an outlier, which notably twists the shape of the frontier and simultaneously deteriorates the efficiency score of nearby entities.
2. DEA only measures efficiency compared to the best practice within the particular set of DMUs under assessment. Hence, the final outcomes concern only this particular study, and they cannot be generalized or compared with scores that have arisen from other studies. For instance, in the current thesis, a DEA study that contains 20 EU countries as DMUs cannot tell how these countries are compared with other countries around the world.
3. DEA results are sensitive both to the size of the sample as well as to the total number of input and output variables. Increasing the number of countries in the current study’s case will lead to the reduction of the average efficiency score because encompassing more DMUs provides greater scope for DEA to find similar comparison partners. In contrast, containing too few DMUs compared to the number of outputs and inputs can artificially inflate the efficiency scores. Finally, increasing the total number of outputs and inputs included without simultaneously increasing the number of DMUs will lead to the raising of the

efficiency scores on average. A general rule about the number of DMUs, outputs, and inputs in DEA is that the number of DMUs in the sample should be at least three times greater than the sum of outputs and inputs used in the model.

4. Efficient DMUs cannot be ranked absolutely since all of them have an efficiency score equal to unity
5. The classical DEA methods cannot aggregate different dimensions of efficiency.

Furthermore, as it was mentioned before, the influential papers that well-established the basic DEA models are those of Charnes, Cooper and Rhodes (1978) and Banker, Charnes and Cooper (1984). The first study assumes that the homogeneous entities operate under Constant Returns to Scale (CRS) in a perfectly competitive environment, while the second makes an attempt to relax those assumptions and assumes that the DMUs operate under Variable Returns to Scale (VRS).

The term returns to scale originates from the field of microeconomics and explains what happens in production when inputs are increased by a factor α, where $\alpha > 0$. Formally, according to economic theory:

1. If a radial increase in the level of factors of production (inputs) leads simultaneously to a larger proportionate radial increase in the level of production (outputs), then a production correspondence is said to exhibit Increasing Returns to Scale (IRS).
2. If a radial increase in the level of factors of production (inputs) leads simultaneously to a smaller proportionate radial increase in the level of production (outputs), then a production correspondence is said to exhibit Decreasing Returns to Scale (DRS).
3. Finally, If a radial increase in the level of factors of production (inputs) leads simultaneously to the same proportionate radial increase in the level of production (outputs), then a production correspondence is said to exhibit Constant Returns to Scale (CRS).

In the case of a DEA model with multiple inputs and outputs, the definitions of IRS, DRS, and CRS in terms of the connection between the percentage radial changes in input and output levels can be generalized as follows.

Let a DMU j have input levels x = {x_{ij}, i = 1…m} and output level y = {y_{rj}, r = 1…s}. Scale its input levels to αx = {αx_{ij}, i = 1…m} and its output levels βy = {βy_{rj}, r = 1…s}. Ultimately, let:

$$\rho = \lim_{\alpha \to 1} (\frac{\mathrm{b} - 1}{a - 1})$$

Then, the following three scenarios are possible:

1. ρ > 1. In this scenario, an increase in the level of inputs by a small percentage leads the level of output to expand by a larger percentage, and hence IRS are observed at (x, y)
2. ρ < 1. In this scenario, an increase in the level of inputs will lead the level of outputs to expand by a smaller percentage, and hence DRS are observed at (x, y)
3. ρ = 1. In this scenario, an increase in the level of outputs will lead to an increase in the level of outputs by the same percentage, and hence CRS are observed at (x, y)

Finally, note that in scaling input levels by α and output levels by β, the mix of inputs and outputs of a DMU j needs to be kept constant. (Thanassoulis, 2001)

Moreover, DEA models can be either input-oriented or output-oriented. An input-oriented model is a model that makes an attempt to minimize its input levels while simultaneously trying to produce the given output levels. Thus, it shows how a DMU can reduce its level of inputs and still produce the same level of outputs. In contrast, an output-oriented model is a model that tries to maximize its output levels for a given level of inputs. Thus, it shows how much a DMU can increase its output levels while it keeps its input levels constant.

In every case, DEA identifies an efficient boundary or a boundary of ‘‘best practices’’ that envelops all the DMUs. All of those DMUs that recline on the boundary are said to be the most efficient or ‘‘best practice units’’, while those that are underneath the boundary are characterized as inefficient. In the inefficient units, a number is given that represents their radial distance from the efficient boundary. Its difference from the maximum value points out:

1. The level of input(s) decrease that the DMU needs to achieve for its given level of outputs (input-oriented DEA model) in order to be characterized as efficient.

2. The level of output(s) increase that the DMU needs to achieve for its given level of inputs (output-oriented DEA model) in order to be characterized as efficient.

3.1.1 DEA and Constant Returns to Scale

One of the most basic DEA models is the CCR or CRS model, which was originally developed by Charnes, Cooper and Rhodes in 1978, as mentioned earlier. There, for every DMU, the virtual input and output by weights (v_i) and (u_r), which are yet unknown, are formed.

$$virtual\ input = v_1 x_{1o} + \cdots + v_m x_{mo}$$

$$virtual\ output = u_1 y_{1o} + \cdots + u_s x_{so}$$

Then, an attempt is made to determine the weights using Linear programming so as to maximize the ratio that is indicated below:

$$\frac{virtual\ output}{virtual\ input}$$

The optimal weights that will be given to the DMU may (and generally will) differ from one DMU to another, and hence, the weights in DEA arise from the data instead of being fixed in advance.

Furthermore, the evaluation of the performance of a DMU by DEA contains two basic steps. More specific:

1. The construction of the Production Probability set (PPS), which is described as was mentioned before, as a set that encompasses all the input-output correspondences that are feasible in principle, containing those observed in the DMUs that are under evaluation
2. The estimation of the size of the maximum expansion of the output levels or the respective maximum contraction of the input levels of the DMUs within the boundaries that are determined by the PPS

In order to formulate the CRS model, an assumption is made that there are N DMUs, which are using m inputs to secure s outputs. It is also stated that x_{ij} (x = 1…m, j = 1…N) and y_{rj} (y = 1…s, j = 1…N) are the ith input and rth output correspondingly.

Then, the technical efficiency, concerning the input-oriented fractional model, can be computed by solving the following fractional program (FP).

$$h_{j_0} = Max \frac{\sum_{r=1}^{s} U_r * y_{rj_0}}{\sum_{i=1}^{m} V_i * x_{ij_0}}$$

Subject to:

$\frac{\sum_{r=1}^{s} U_r * y_{rj}}{\sum_{i=1}^{m} V_i * x_{ij}} \leq 1$ $\quad$ j = 1…j_0…N

$U_r \geq \varepsilon$ $\quad$ r = 1…s

$V_i \geq \varepsilon$ $\quad$ i = 1…m

ε is a non-Archimedean infinitesimal.

The goal of the program is to determine the weights V_i and U_r that maximize the ratio of the DMU under assessment. As it entails from the first constraint, the optimal value cannot be greater than unity. Thus, the technical efficiency which is calculated by the ratio $\frac{\sum_{r=1}^{s} U_r * y_{rj_0}}{\sum_{i=1}^{m} V_i * x_{ij_0}}$ cannot take values greater than unity. Finally, if $\frac{\sum_{r=1}^{s} U_r * y_{rj_0}}{\sum_{i=1}^{m} V_i * x_{ij_0}} = 1$, then the DMUs under evaluation is said to be efficient.

Now, the above FP can be converted to the following linear program (LP) using the Charnes and Cooper transformation (1962):

$$Max \sum_{r=1}^{s} u_r * y_{rj_0}$$

Subject to constraints:

$$\sum_{i=1}^{m} v_i * x_{ij_0} = 1$$

$\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} \leq 0$, j = 1…$j_0$…N

$u_r \geq \varepsilon$ $\quad$ r = 1…s

$v_i \geq \varepsilon$ $\quad$ i = 1…m

ε is a non-Archimedean infinitesimal.

The first constraint $\sum_{i=1}^{m} v_i * x_{ij_0} = 1$ is called normalization constraint, while together with the constraint $\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} \leq 0$, they denote that the technical efficiency, which is estimated by the $\sum_{r=1}^{s} u_r * y_{rj_0}$, cannot be greater than one. Of course, if $\sum_{r=1}^{s} u_r * y_{rj_0}$ equals 1, then the equivalent DMU is considered efficient.

Note that the technical efficiency of the above program, which is called value-based DEA model, is approximately but not exactly equal to the true technical efficiency, which can be computed by its dual envelopment model. Its equivalent envelopment model can be solved as follows:

$$\min k_0 - \varepsilon[\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+]$$

Subject to constraints:

$\sum_{J=1}^{N} \lambda_j * x_{ij} = k_0 x_{ij_0} - S_i^-$ i = 1 . . . m

$\sum_{J=1}^{N} \lambda_j * y_{rj} = S_r^+ + y_{rj0}$ r = 1 . . . s

$\lambda_j \geq 0, j = 1 \ldots N, S_i^-, S_r^+ \geq 0 \ \forall$ *i and r*, k_0 free.

ε is a non-Archimedean infinitesimal.

As it is already mentioned, the variable λj is the weight calculated from the above model for DMUj whilst the variables S_r^+, S_i^- are the slack variables that are used in linear programming. These slacks represent any further output expansion or input contraction that is feasible to be achieved by the evaluated DMU.

This linear program is solved in two stages:

1. Firstly, the model makes an attempt to minimize the technical efficiency k_0, without taking into consideration the slack variables S_r^+, S_i^-. The minimized k_0 constitutes the optimal value k^*_0
2. The technical efficiency k_0 is replaced in the LP, and it is solved with the purpose of maximizing the slacks $\sum_{i=1}^{m} S_i^-$, $\sum_{r=1}^{s} S_r^+$. The outcomes of the second stage

constitute the optimal values of the slack variables when it is secured that the optimal technical efficiency has been used in the computation.

If at least one of those values is greater than zero at the optimal solution of the model, it entails that the respective input or output of the DMU j_0 can be further improved after its input levels have been contracted to the proportion k_0*, which is the technical efficiency of the model.

If k_0* = 1 and $S_r^+ = 0$, r = 1…s, $S_i^- = 0$, i = 1… m then DMU is Pareto-efficient because the model was unable to decrease the input level without decreasing the output level or increasing the output level without correspondingly increasing the input level of the specific DMU.

In summary, the technical efficiency of the above problem for DMUj_0 is the variable k_0 and can take values from 0 to 1 (or from 0 to 100%). The mathematical program indicated with the above equations is solved individually for every DMU, and its solution can lead to three different scenarios:

1. DMUj_0 is Pareto-efficient if and only if k_0* = 1 and $S_r^+ = 0$, r = 1…s, $S_i^- = 0$, i = 1… m
2. If the value of one of the slacks S_r^+, S_i^- is greater than zero at the optimal solution, the related input (or output) can additionally be improved
3. If none of the above mentioned happens, DMUj_0 has technical efficiency equal to k_0*. In this case, the technical efficiency at the optimal solution k_0* < 1 represents the maximum radial contraction of the input levels without deteriorating the level of outputs in order for the DMUj_0 to be considered efficient.

Every DEA model that has been described so far has been formulated under input orientation. Nevertheless, there are cases where the equivalent output-oriented DEA models need to be utilized. In order to formulate the respective output-oriented envelopment DEA model, assume that there are N DMUs, which use m inputs to produce s outputs. Assume also that x_{ij} (x = 1…m, j = 1…N) constitutes the level of ith input of DMU j and y_{rj} (y = 1…s, j = 1…N) the level of rth output of the DMU j correspondingly.

The technical efficiency under output orientation of DMU j_0 equals $1/h^*_{j_0}$, where $h^*_{j_0}$ constitute the optimal value of h_{j_0} in:

$$\text{Max}\, h_{j0} - \varepsilon[\sum_{i=1}^{m} I_i + \sum_{r=1}^{s} O_r]$$

Subject to constraints:

$$\sum_{j=1}^{N} a_j * x_{ij} = x_{ij_0} - I_i \qquad i = 1 \ldots m$$

$$\sum_{j=1}^{N} a_j * y_{rj} = O_r + h_{j_0} * y_{rj_0} \quad r = 1 \ldots m$$

$a_j \geq 0, j=1 \ldots N, I_i, O_r \geq 0\ \forall\ i$ *and r*, h_{j_0} *free.*

ε is a non-Archimedean infinitesimal

The first objective of the above model is to maximize the h_{j_0}. The model recognizes a point within the boundaries of the PPS that offers output levels that reflect the maximum feasible radial expansion of the output levels of the DMU j_0 without any increase to its input levels. Hence, by definition, the variable 1 / $h^*_{j_0}$ is the technical output efficiency of the DMU j_0.

Concerning the slack variables I_i and O_r, they are interpreted in a similar way as the slack variables S_i^- and S_i^+ of the equivalent input-oriented model. More specifically, they mirror any additional output augmentation and/or input reduction that may be indispensable in order for the DMU j_0 to be regarded as efficient.

If $h^*_{j_0}$ = 1 and O^*_r = 0, r = 1…s, I^*_i = 0, i=1…m then the DMU j_0 is considered to be pareto efficient. Hence, the model has not been able to recognize a point within the feasible area that can improve on any of the output levels of the DMU j_0 without simultaneously increasing some other input levels. The above LP model is solved in two stages in a similar way as the DEA CRS model under input orientation. More precisely:

1. In the first stage, the model seeks to maximize the variable h_{j_0}, ignoring the slack variables I_i and O_r. This yields the maximum value $h^*_{j_0}$ of h_{j_0}
2. In the second stage, h_{j0} has been replaced by h^*_{j0}, and the model is solved with the aim of maximizing $\sum_{i=1}^{m} I_i + \sum_{r=1}^{s} O_r$.

Summarizing, the solution of the above model shares the following information:

1. A DMU j_0 is considered Pareto-efficient if $h^*_{j_0}$, = 1 and $O^*_r = 0$, r = 1…s, $I^*_i = 0$, i =1…m
2. The technical output efficiency of the DMU j_0 is equal to 1 / $h^*_{j_0}$. It is necessary to note that the technical input efficiency and the technical output efficiency are equal when it is assumed that the DMUs operate under Constant Returns to Scale. (Cooper et al, 2010)

Ultimately, the above envelopment model has its dual, in terms of linear programming, value-based model under output orientation, which can be computed by the equations below:

$$Min \sum_{i=1}^{m} v_i * x_{ij_0}$$

Subject to constraints:

$$\sum_{i}^{s} u_r * y_{rj_0} = 1$$

$\sum_{i}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} \leq 0$, j = 1…$j_0$…N

$u_r \geq \varepsilon$, r = 1…s

$v_i \geq \varepsilon$, i = 1…m

ε is non-Archimedean infinitesimal.

The technical efficiency of the above model equals $\frac{1}{\sum_{i=1}^{m} v_i^{*} * x_{ij_0}}$, which is approximately, but not exactly, equal to the true technical efficiency $\frac{1}{h_{j_0}^{*}}$ of the equivalent envelopment model. Of course, if $\frac{1}{\sum_{i=1}^{m} v_i^{*} * x_{ij_0}}$ equals one, the equivalent DMU j_0 regarded as Pareto efficient.

The solutions of the above basic CRS and VRS DEA models can share directly some useful information. More precisely, the DEA envelopment model yields the following information:

1. It notes a measure of the efficiency of each DMU
2. When a DMU is regarded as Pareto-efficient, it can be utilized as a role model for inefficient DMUs
3. When a DMU is not pareto-efficient: Firstly, it can identify efficient peers whose practices may attempt to imitate in order to enhance its performance, and secondly, it can estimate target input-output levels, which the DMU should, in principle, be able to reach under efficient operation.

Regarding its dual value-based DEA model, outside of the above, it can also take a view on the robustness of the efficiency measure of a DMU. (Thanassoulis, 2001)

3.1.2 DEA and Variable Returns to Scale

In many real-life situations, the assumption that the DMUs operate under Constant return to Scales is not proper. In those cases, it is preferable to utilize the model that was first developed by Banker, Charnes and Cooper in 1984, who made an attempt to relax those assumptions by assuming that the DMUs under evaluation operate under Variable Returns to Scale.

It is quite simple to reform the DEA models that were constructed with the aim of assessing DMUs under CRS so that they can be utilized to evaluate efficiency under VRS. More precisely, assume that there are N DMUs (j=1…N) that use m inputs to generate s outputs. Let also state that x_{ij} (i=1…m, j =1…N) is the level of ith input of the DMU j and y_{rj} (r=1…s, j =1…N) is the level of rth output of the DMU j.

Then, the technical efficiency of the input-oriented envelopment model under VRS (which is called pure technical efficiency) can be measured by solving the following LP:

$$\min h_0 - \varepsilon[\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+]$$

Subject to constraints:

$\sum_{J=1}^{N} \lambda_j * x_{ij} = k_0 x_{ij_0} - S_i^-$ i = 1 . . . m

$\sum_{J=1}^{N} \lambda_j * y_{rj} = S_r^+ + y_{rj0}$ r = 1 . . . s

$$\sum_{J=1}^{N} \lambda_j = 1$$

$\lambda_j \geq 0, j = 1 \ldots N, S_i^-, S_r^+ \geq 0 \; \forall \; i \; and \; r, k0$ free.

ε is a non-Archimedean infinitesimal.

As it is easily observed, the only difference between the above LP model and its equivalent CRS model is the addition of the third constraint $\sum_{J=1}^{N} \lambda_j = 1$, which is called convexity constraint, and it is useless under CRS.

Similar to the respective CRS envelopment model, the solution of the above model shares the following information:

1. If h_0^* equals 1 and the slack variables S_i^- = 0, i =1…m, S_r^+ = 0, r = 1…s, then the DMU j_0 is considered to be Pareto-efficient.

2. The pure technical efficiency of the DMU j_0 equals h_0^*

Note that the pure technical input efficiency of the DMU j_0 is always greater than its respective technical input efficiency.

The equivalent input oriented value-based DEA model under Variable Returns to Scale is:

$$Max \sum_{r=1}^{s} u_r * y_{rj_0} + \omega$$

Subject to constraints:

$$\sum_{i=1}^{m} v_i * x_{ij_0} = 1$$

$\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} + \omega \leq 0$, j = 1…$j_0$…N

$u_r \geq \varepsilon$ r = 1…s

$v_i \geq \varepsilon$ i = 1…m

ω free.

For the above model, if the pure technical efficiency $\sum_{r=1}^{s} u_r * y_{rj_0} + \omega = 1$, then the DMU j_0 is regarded to be efficient.

Likewise, the VRS envelopment model under output orientation can be solved as follows:

$$\text{Max} \quad z - \varepsilon[\sum_{i=1}^{m} I_i + \sum_{r=1}^{s} O_r]$$

Subject to constraints:

$$\sum_{j=1}^{N} a_j * x_{ij} = x_{ij_0} - I_i \qquad i = 1 \ldots m$$

$$\sum_{j=1}^{N} a_j * y_{rj} = O_r + z * y_{rj_0} \quad r = 1 \ldots s$$

$a_j \geq 0$, *j=1…N*, I_i, $O_r \geq 0 \; \forall$ *i and r*, h_{j_0} *free.*

ε is a non-Archimedean infinitesimal.

The solution of the above model shares the following information:

1. If $\frac{1}{z^*_{j_0}}$ equals 1 and the slack variables S_i^- = 0, i =1…m, S_r^+ = 0, r = 1…s, then the DMU j_0 is considered to be Pareto-efficient.

2. The pure technical efficiency of the DMU j_0 equals $\frac{1}{z^*_{j_0}}$

Note that the pure technical output efficiency of the DMU j_0 is always greater than its respective technical output efficiency

Finally, the respective VRS value-based model under output orientation can be solved as follows:

$$Min \sum_{i=1}^{m} v_i * x_{ij_0} + \omega$$

Subject to constraints:

$$\sum_{r=1}^{s} u_r * y_{rj_0} = 1$$

$\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} - \omega \leq 0,$ j = 1…j_0…N

$u_r \geq \varepsilon$ r = 1…s

$v_i \geq \varepsilon$ i = 1…m

ω free.

The pure technical efficiency equals $\frac{1}{\sum_{i=1}^{s} v_i^* * x_{ij_0} + \omega}$. If $\frac{1}{\sum_{i=1}^{s} v_i^* * x_{ij_0} + \omega} = 1$, then the DMU j_0 is regarded to be efficient.

As it is already stated, the efficiency as computed by the CRS models is referred to as technical efficiency. This term can further be analyzed to pure technical efficiency (emerges as a result of the VRS-BCC models) and scale efficiency. Denote:

$$Technical\ Efficiency = TE$$

$$Pure\ Technical\ Efficiency = PTE$$

$$Scale\ Efficiency\ = SE$$

Then, the different terms of efficiency are related by the equation below:

$$TE = PTE * SE$$

4. Data, Models' formulation and Application

4.1 Research's Data

As it is already mentioned, the complex notion of sustainability consists of three different pillars: the economic, the environmental, and the social. Hence, in order to create a composite sustainability index, it is imperative to encompass and aggregate indicators that express all of the abovementioned dimensions.

In order to select the appropriate data for this research's scenarios, the frequency of occurrence of each indicator used in the years 2021 to 2023 is taken into consideration. The indicators utilized in previous studies on DEA and sustainability by other scholars are listed below in a tabular format.

Table 6 Frequency of appearance of different indicators contained in the studies about DEA and sustainability for the years 2021-2023

Inputs	**Frequency**	**Intermediates**	**Frequency**	**Outputs**	**Frequency**
Labor force	13	GDP	6	GDP	9
Energy consumption/use	11	GHG emmissions	1	CO2 emissions	6
Capital	10	Eco patents	1	Medium equivalised net income	2
Population	2	GDP per capita in PPS Index	1	SO2 emissions	2
CO2 emissions	2	Final energy consumption	1	Gdp per capita	2
Unemployment level/percentage of unemployed	2	Total expenditure per euro habitant	1	Clean water and sanitation	1
Researchers in R&D	1	Waste water	1	Affordable and clean energy	1
R&D expenditure	1	Solid waste	1	Resident population	1

Total annual emissions of pollutants into the atmospheric air from stationary sources	1	Untreated waste water	1	GDP pc/(GHG/area)	1
Total annual emissions of pollutants into the atmospheric air from mobile sources	1	Untreated solid waste	1	Employment rate	1
Volumes of non-treated or non-sufficiently treated wastewaters that were discharged into water bodies	1	Industrial dust production	1	High-technology exports	1
Amount of land field waste	1	Industrial SO2 production	1	People at risk of poverty or social exclusion	1
Volume of fresh water abstraction	1	Volume of pollutants generated in industrial waste gas	1	Final consumption expenditure of households	1
Gross fixed capital formation	1	Treatment investment of industrial waste gas	1	Terrestrial protected area	1
Labor per capita	1	Capacity of industrial waste gas treatment(previous year)	1	Share of renewable energy in gross final energy consumption	1
Energy use/area	1	Number of patent application authorizations	1	Greenhouse gas emissions (in CO2 equivalent)	1
Electricity/area	1	Number of scientific papers published	1	Patent applications to the European patent office (EPO) by priority year	1
Deviations temp	1	Labor force	1	Overall life satisfaction	1
Gross fixed capital at current prices	1	Fixed investment	1	Satisfaction with living environment	1
GDP per capita in PPS Index	1	Energy	1	Percentage of females in total labor population	1
Environmental investment	1	CO2 emissions	2	Energy Productivity	1
Health investment	1	CO2 emission intensity index	1	Resource productivity	1
Investment	1	Agricultural output	1	Gross regional product in current prices	1
Electricity	1	Industrial output	1	Mortality rate	1
Land	1	Agricultural methane CH4	1	Survival rate	1
Water	1	Industry CO2 emissions	1	Industrial SO2 emissions	1

Energy consumption from coal	1			Industrial soot emissions	1
Oil	1			Total waste water volume	1
Gas sources Volume of Vehicles	1			Number of middle school students	1
Investment of fixed assets	1			Insurance covered	1
Full time equivalent of R&D personnel	1			Social welfare	1
R&D internal expenditures	1			Industrial dust emmisions	1
Levelized cost of energy	1			Industrial SO2 production	1
Socio-economic vulnerability index	1			CH4 emissions	1
Capital inventory	1			REAL GDP	1
Agricultural labor	1			Industrial value added, capacity of industrial waste gas treatment(current year)	1
Agricultural land use	1			Volume of pollutants emission in industrial waste gas	1
Nitrogen fertilizer	1			Total industrial output value of mining industry	1
Industrial employment	1			Solid waste discharge	1
Industrial energy use	1			Exhaust emissions	1
Import of goods and services	1			Waste water discharge	1
Fossil fuel consumption	1			Life satisfaction	1
				Greenhouse gases: Total emissions excluding LULUCF	1
				Manufacturing value added to GDP	1
				Charges for the use of intellectual property in terms of payments	1
				Charges for the use of intellectual property in terms of receipts	1
				Value of high tech exports	1
				Number of R&D researchers	1

				Number of scientific and technical journal articles published	1
				The number of patent applications filed	1
				The number of trademark applications filed	1
				Life expectancy at birth	1
				Number of direct employees	1
				Avoided CO2	1
				Equity index	1
				CO2-solid emissions	1
				CO2-gas emissions	1
				CO2-liquid emissions	1
				CO2-other emissions	1
				Average temperature change	1
				Average rainfall amount	1
				Disaster occurrence	1
				Compulsory education duration	1

The analysis of the literature reveals a prevalent reliance on certain indicators as inputs in empirical studies. Specifically, the labor force, energy consumption, and capital were utilized in 13, 11, and 10 studies, respectively. The GDP has emerged as the most frequently employed intermediate variable, appearing in six instances. Finally, both the GDP and CO2 emissions were utilized as output variables in nine and six investigations, respectively. This pattern highlights the significance of these indicators in examining environmental sustainability and economic growth.

In accordance with the aforementioned considerations and data availability, the following indicators were selected for this study, as presented in the ensuing table.

Table 7 Indicators contained in the study

Indicator	Sustainability dimension	Source	Year
Final energy consumption	Environmental	Eurostat	2021
Total labor force	Economic	World Bank	2021
Gross fixed capital formation (Investments)	Economic	Eurostat	2021
Overall life satisfaction	Social	Eurostat	2021
GDP	Economic	World Bank	2021
GHG emissions	Environmental	Eurostat	2021

Note that Gross fixed capital formation (GFCF) will be used as a proxy for capital and GHG emissions as a proxy for CO2 emissions. Additionally, all data were retrieved from either Eurostat or World Bank for the year 2021, which was the last common year with all of the above data available for all countries under assessment. Ultimately, the DMUs of the research's models will be 20 countries from the Eurozone. More specifically, the countries which use the euro as their official currency.

4.2 Models' formulation

4.2.1 Scenarios 1a, 1b

Aiming to select the indicators that will be employed in the first scenario, the frequency of appearance of the indicators used in relevant studies, as can be seen in table 6, will be taken into account. Based on this criterion, the inputs will include the labor force, energy consumption, and gross fixed capital formation (GFCF), which serves as a proxy for capital. Finally, the single output variable will be the gross domestic product (GDP).

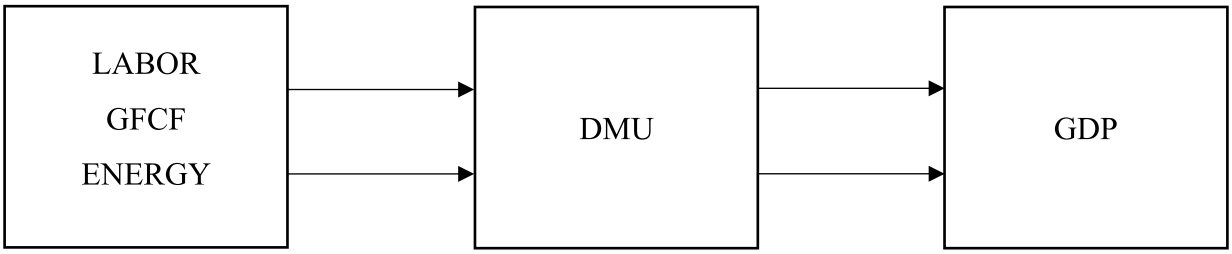

Figure 3 Indicators used in scenario 1a and scenario 1b

With the purpose of constructing a composite index, the output-oriented value-based CRS DEA model, which is represented by the following linear program, will be employed:

$$Min \sum_{i=1}^{m} v_i * x_{ij_0}$$

Subject to constraints:

$$\sum_{r}^{s} u_r * y_{rj_0} = 1$$

$\sum_{r}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} \leq 0$, j = 1…j0…N

$u_r \geq \varepsilon$, r = 1…s

$v_i \geq \varepsilon$, i = 1…m

ε is non-Archimedean infinitesimal.

In addition, except the CRS model, the output oriented value-based VRS DEA model will be utilized too. The VRS model is expressed as follows:

$$Min \sum_{i=1}^{m} v_i * x_{ij_0} + \omega$$

Subject to constraints:

$$\sum_{r=1}^{s} u_r * y_{rj_0} = 1$$

$\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} - \omega \leq 0,$ j = 1…j0…N

$u_r \geq \varepsilon \qquad r = 1 \ldots s$

$v_i \geq \varepsilon \qquad i = 1 \ldots m$

ω free.

The output-oriented models aim to maximize the output variables without increasing any of the inputs. In the current case, the goal of the model is to maximize the GDP of the countries under assessment without simultaneously increasing any of the input variables, namely the labor force, energy consumption, and capital.

In contrast, the target of the input-oriented models is to decrease the input variables as much as possible without simultaneously decreasing the output. In the current case, those models aim to minimize the labor force, energy consumption, and capital without decreasing the GDP.

In practical applications, governing bodies and policymakers strive to attain a dual objective of maximizing gross domestic product (GDP) while minimizing environmental pollution by reducing energy consumption. This pursuit is driven by the belief that economic growth, as measured by GDP, is a primary indicator of societal well-being, and that minimizing pollution is essential for preserving the environment.

Regarding capital, governments aim to increase it as a means of fostering economic development and prosperity. Hence, the goal of the input-oriented model, which is to minimize it, is considered to be the opposite of the real goal. In contrast to capital and GDP, the labor force is a relatively stable measure that cannot be changed dramatically by policy implications. Therefore, the goal of the input-oriented models to minimize the labor force is not either meaningful or realistic.

In light of the aforementioned analysis, the output-oriented variation was selected for this scenario, as well as for the remaining applications presented in this research, due to its greater realism compared to the input-oriented approach that aims to minimize inputs such as capital and labor. This variation is deemed more realistic because maximizing economic output without increasing inputs is a more practical objective than decreasing inputs without simultaneously decreasing output.

With the goal of solving the above models, the software tool deaR - shiny was utilized. DeaR - shiny constitutes an interactive, user-friendly software (Benítez et al., 2021), that serves as a frontend for the functions that exist in the deaR package of R (Coll-Serrano et al., 2023). This software tool provides a convenient and intuitive interface for implementing and solving various DEA models.

The solution yielded the following results.

Table 8 Results of the efficiency scores - scenario 1a and scenario 1b

Country	Efficiency CRS	Rank CRS	Efficiency VRS	Rank VRS
Estonia	0,633489	20	0,66028	20
Latvia	0,682505	19	0,791653	15
Finland	0,697725	18	0,821558	13
Slovakia	0,72372	17	0,723736	19
Austria	0,726322	16	0,817488	14
Slovenia	0,735073	15	0,760237	17
Croatia	0,737463	14	0,783098	16
Lithuania	0,746536	13	0,755219	18
Belgium	0,750593	12	0,877362	11
France	0,751755	11	0,965326	10
Malta	0,796039	10	1	1
Italy	0,807885	9	0,975239	9
Portugal	0,817809	8	0,843946	12
Cyprus	0,824552	7	1	1
Germany	0,82715	6	1	1
Spain	0,850058	5	1	1
Netherlands	0,898061	4	1	1
Greece	1	1	1	1
Ireland	1	1	1	1
Luxembourg	1	1	1	1

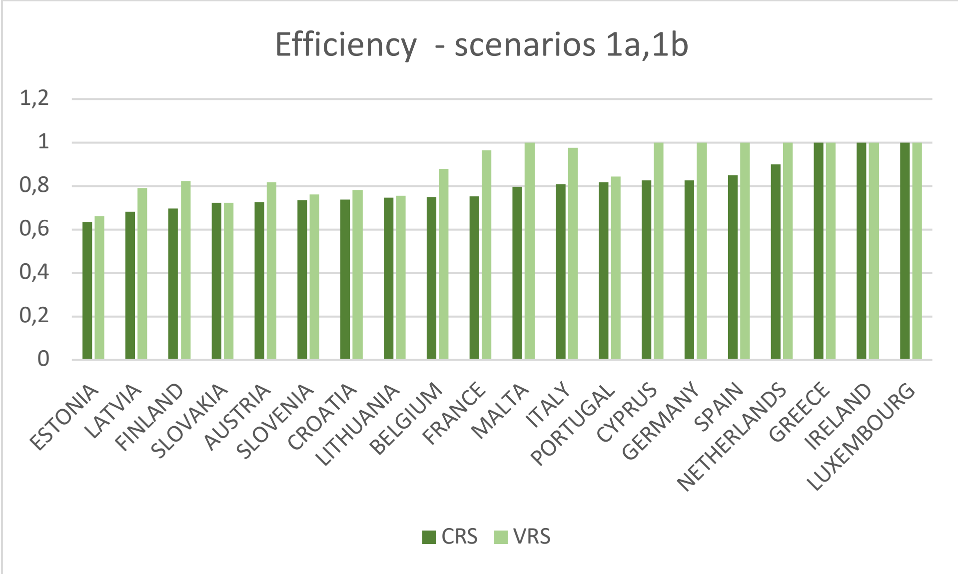

Figure 4 Efficiency scores - scenario 1a and scenario 1b

Moreover, some descriptive statistics of the efficiency scores were computed.

Table 9 Descriptive statistics of the efficiency scores - scenario 1a and scenario 1b

Mean		Standard Deviation		Minimum		Maximum	
CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
0,80033675	0,8887571	0,10584152	0,11663061	0,633489	0,66028	1	1

In the context of the classical Constant Returns to Scale (CRS) model, three countries, namely Greece, Ireland, and Luxembourg, achieved maximum comparative efficiency. The remaining seventeen countries in the sample of 20 Eurozone countries obtained efficiency scores below the maximum level. Specifically, Cyprus, Germany, Italy, Netherlands, Portugal, and Spain have attained efficiency levels lower than one but greater than the overall mean value of 0.80033675. The remaining countries have obtained efficiency scores below this average value. The countries with the lower efficiency scores are Estonia, Latvia, and Finland, which have obtained scores of 0.633489, 0.682505, and 0.697725, respectively.

In the context of the classical Variable Returns to Scale (VRS) model, it is noteworthy that all the countries under assessment achieved higher efficiency scores compared to those obtained through the Constant Returns to Scale (CRS) model, as evidenced by the results presented in the third column. This outcome was anticipated since pure technical efficiency is always equal to or greater than equivalent technical efficiency. Specifically, Cyprus, Germany, Malta, Netherlands, and Spain reached the efficient frontier, joining Greece, Ireland, and Luxembourg in the group of relatively efficient countries.

Moreover, Italy and France were the only inefficient countries with a score greater than the significantly high average of 0,8887571. The remaining countries attained scores greater than 0,70, except Estonia, which obtained a score of 0,66028 and exhibited the worst efficiency score among the 20 Eurozone countries under assessment. It is also worth noting that Slovakia had almost identical scores in both DEA variations, suggesting that it exhibits Constant Returns to Scale even under the assumption of Variable Returns to Scale.

Of course, it must be stated that the above results are not robust on their own since DEA is an empirical technique, which is sensitive in the selection of the variables used in the model. Hence, in order to increase the validity of the obtained outcomes, different sets of indicators must be taken into consideration.

4.2.2 Scenario 2

As already stated, the indicators used in the first application are the most common in the relevant literature. Nevertheless, it seems that this group of indicators pays more attention to the economic and the environmental dimensions without taking into account the third pillar of sustainability, namely the social. Thus, in order to overcome this problem and simultaneously apply a form of sensitivity analysis to the model's results, it is imperative to add a social variable to the existing set of indicators contained in the study.

Upon reviewing Table 5, it becomes apparent that none of the social indicators employed in recent literature on DEA and sustainability have been consistently utilized. While indicators such as social welfare, overall life satisfaction, and people at risk of poverty or social exclusion have appeared in recent studies, none of them have been repeatedly selected. In light of this, the current research will incorporate the indicator of overall life satisfaction as an output variable in order to represent the social dimension of sustainability. Thus, the measures that are contained in scenario b as inputs are the following:

1. Labor force
2. Energy consumption and

3. Gross Fixed Capital Formation

The measures that are used in scenario b as outputs are the following:

1. GDP
2. Overall life satisfaction

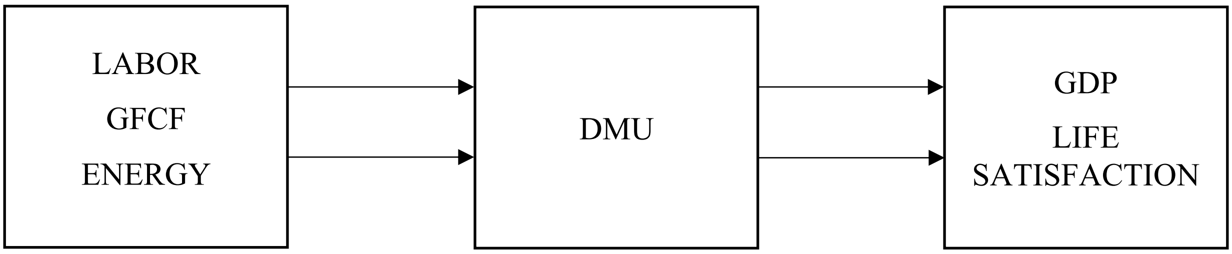

Figure 5 Indicators used in scenario 2

In order to aggregate the above indicators into a composite sustainability index, the classical output-oriented value-based CRS DEA model will be applied again. It must be noted that the VRS model will not be considered since its results are not meaningful, given that it identifies an excessive number of countries as efficient.

The results of the model are illustrated below.

Table 10 Results of the efficiency scores - scenario 2

Country	Efficiency	Rank
Finland	0,697725	20
Estonia	0,708602	19
Austria	0,726322	18
Slovakia	0,745746	17
Belgium	0,750593	16
France	0,751755	15
Lithuania	0,77403	14
Slovenia	0,804195	13
Croatia	0,807591	12
Italy	0,807885	11
Portugal	0,817809	10
Latvia	0,825989	9
Germany	0,82715	8
Spain	0,850058	7

Netherlands	0,898061	6
Cyprus	0,990776	5
Greece	1	1
Ireland	1	1
Luxembourg	1	1
Malta	1	1

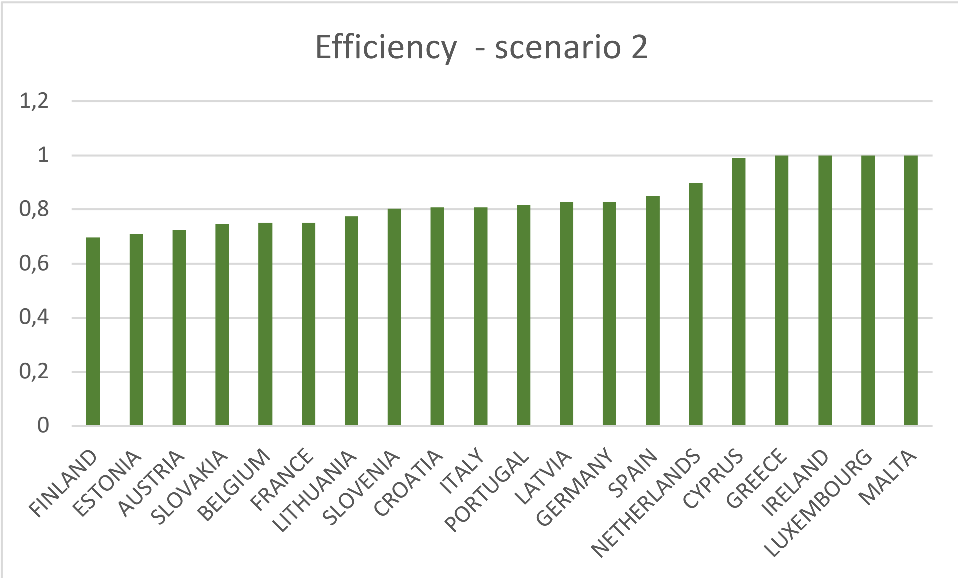

Figure 6 Efficiency results – scenario 2

Furthermore, some descriptive statistics of the efficiency scores were calculated.

Table 11 Descriptive statistics of the efficiency scores - scenario 2

Mean	Standard Deviation	Minimum	Maximum
0,839214	0,105695208	0,697725	1

Upon analyzing the results table of scenario 2, it is evident that four countries are considered to be comparatively efficient. More specifically, Malta joined Greece, Ireland, and Luxembourg, which were efficient considering the previous scenarios, in the group of relatively efficient countries. Moreover, Cyprus almost reached the efficient frontier by obtaining a score of 0,990776, while Spain was the last country to attain a sustainability score above the mean value of 0,839214.

The 12 remaining countries had a score between the average value and 0,70, except Finland, which obtained a score of 0,697725. Finally, Estonia and Austria, with scores of

0,708602 and 0,726322 correspondingly, complete the group of the worst three performances, according to scenario 2.

It is noteworthy that the efficiency results obtained under scenario 2 are either equivalent to or exceed those obtained under scenario 1a. Specifically, the results of twelve countries have remained unchanged, while the remaining eight have experienced some degree of improvement. The most significant advancements were observed in Malta and Cyprus, with increases in performance scores of approximately 20 and 16 percentage units, respectively.

4.2.3 Scenarios 3a, 3b, 3c

In scenarios 1a and 1b, the efficiency of the Eurozone countries was measured, taking into account the most common set of indicators in the literature, while in scenario 2, the overall life satisfaction was added to the set with the goal of expressing the social dimension of sustainability. Nevertheless, both scenarios ignored a measure that has appeared many times in recent studies, namely CO_2 emissions.

The problem with this measure is that it constitutes an undesirable output variable, while the goal of the classical output-oriented models is to maximize those variables. Hence, those models cannot manage that kind of indicators without the appropriate modifications.

In order to face this problem, the researchers of the topic applied several approaches to treat the undesirable outputs. The most common ways to overcome this limitation, according to the works of Halkos and Petrou (2018) and Halkos and Petrou (2019), are summarized below:

1. Neglecting them: Not encompassing undesirable outputs in the production function
2. Employ them as input variables
3. Utilization of non-linear models
4. Transformation in the data of the undesirable output such as:
    a. Set (U) = -U, which was proposed by Koopmans (1951)
    b. Set (U) = -U+β (Ali and Seiford, 1990; Scheel, 2001; Seiford and Zhu, 2001)
    c. Set f(U) = 1/U (Golany and Roll, 1989; Lovell et al., 1995).

5. Developing new models

It needs to be stated that the assumption of variable returns to scale (VRS) is taken into consideration in all cases where the above transformations are utilized (Wojcik et al., 2017).

Considering the abovementioned and the data availability, the set of indicators that were used in scenario 3a are the following:

The input variables of scenario 3a are:

1. The labor force and
2. The energy consumption

The output variables of scenario 3a are:

1. The GDP and
2. The GHG emissions

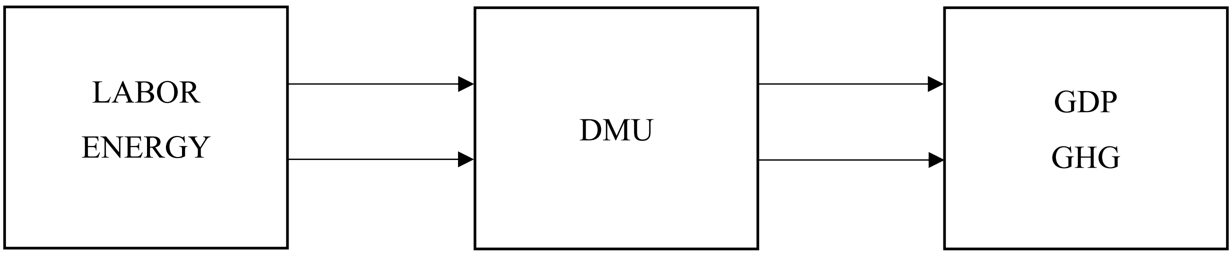

Figure 7 Indicators contained in scenario 3a

It is worth noting that the measures of the GFCF and the overall life satisfaction were excluded in scenario 3a because, together with the above variables, they resulted in most of the countries under assessment being considered comparatively efficient. Hence, the results of the model weren't particularly useful or meaningful. Ultimately, in order to treat GHG emissions, which is an undesirable output, the transformation proposed by Koopmans (1951) will be applied, while VRS are assumed.

The results of the output-oriented value-based VRS DEA model are the following:

Table 12 Efficiency results - scenario 3a

Country	Efficiency	Rank
Croatia	0,238111	20
Slovakia	0,241326	19
Latvia	0,247806	18
Lithuania	0,305483	17
Slovenia	0,315937	16
Estonia	0,336965	15
Greece	0,357106	14
Portugal	0,41132	13
Cyprus	0,428519	12
Finland	0,520592	11
Austria	0,775964	10
Spain	0,839955	9
Italy	0,889838	8
Belgium	0,910548	7
Netherlands	0,923591	6
France	1	1
Germany	1	1
Ireland	1	1
Luxembourg	1	1
Malta	1	1

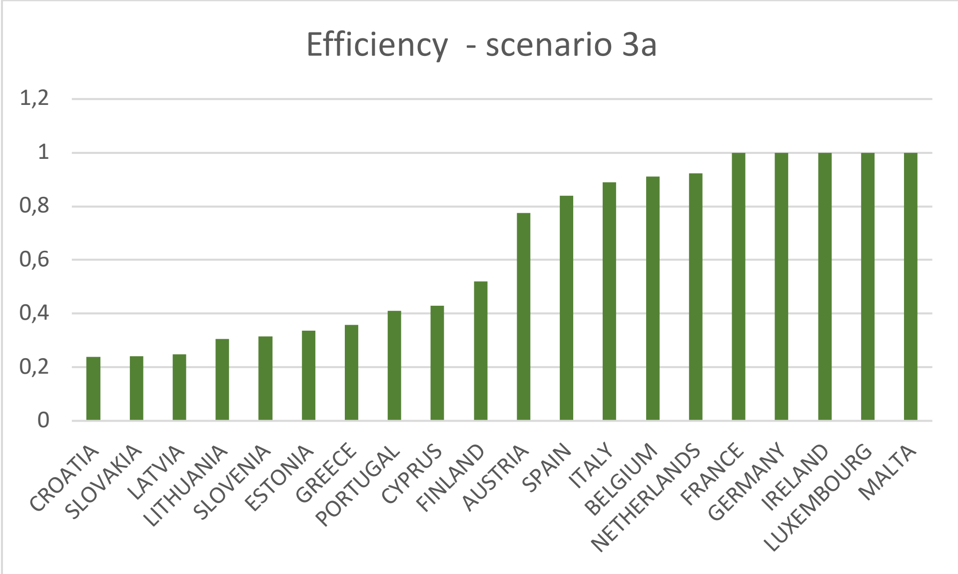

Figure 8 Efficiency results - scenario 3a

Some descriptive statistics of the efficiency scores are presented below.

Table 13 Descriptive statistics of the efficiency scores – scenario 3a

Mean	Standard Deviation	Minimum	Maximum
0,637153	0,315951	0,238111	1

As it is clearly noticed, the range of the efficiency scores has been significantly increased compared to the previous scenarios. Their average value is 0,637153, while their minimum value is 0,238111. Furthermore, five countries obtained an efficiency score equal to 1. Ireland and Luxembourg are the only countries that remain efficient in all scenarios so far, while Malta, which was efficient in scenario 2, together with France and Germany, complete the group of countries that reached the efficient frontier according to scenario 3.

Worth mentioning is the difference in the scores of the 10th and the 11th countries, which are Austria and Finland, respectively. More precisely, Austria attained a score of 0,775964, while Finland had a score of 0,520592. In addition, the three worst performances belong to Latvia, Slovakia, and Croatia, which obtained scores of 0,247806, 0,241326, and 0,238111, correspondingly. Ultimately, another interesting finding of the above results is the worsening of Greece's performance, which was efficient in the previous scenarios, while in scenario 3, it attained a score of 0,357106.

As already mentioned, DEA is highly sensitive to the selection of variables included in the model. This fact becomes apparent, considering the results of the first three scenarios, where countries like Greece and Cyprus exhibited high scores in the first two, while their performances worsened a lot with the change of the variables that were utilized in scenario 3 (even under the assumption of variable returns to scale). On the contrary, with the new set of variables, countries such as France and Germany improved their performance compared to the previous scenarios and became efficient.

Moreover, the GHG emissions will be treated as an input variable to explore how the scores of scenario 3a will react. Therefore, the CRS model and the VRS model under output orientation will be employed. The following measures are used as inputs in scenarios 3b and 3c:

1. Labor force
2. Energy consumption
3. GHG emissions

Lastly, Gross domestic product constitutes the only output variable.

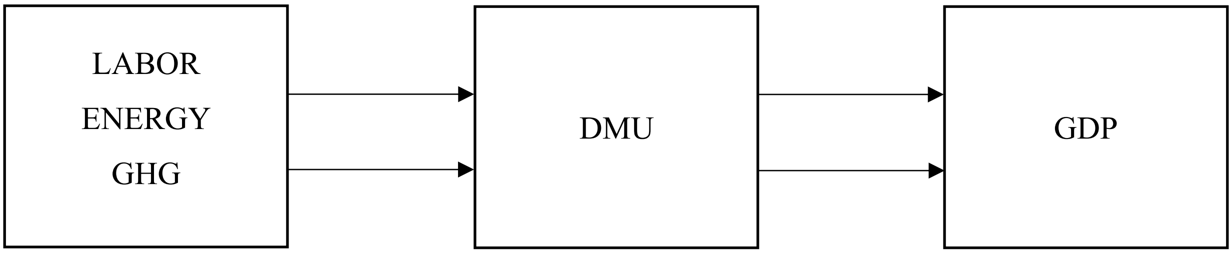

Figure 9 Indicators used in scenarios 3b and 3c

The solutions of the models under the scenarios 3b and 3c yielded the following results.

Table 14 Results - scenario 3b, 3c

Country	Efficiency CRS	Rank CRS	Efficiency VRS	Rank VRS
Latvia	0,230326	20	0,291663	20
Croatia	0,274901	19	0,39653	15
Slovenia	0,287642	18	0,327298	18
Lithuania	0,295834	17	0,305483	19
Slovakia	0,304173	16	0,359073	16
Estonia	0,306278	15	0,336965	17
Cyprus	0,365764	14	0,428519	14

Greece	0,423747	13	0,46017	13
Portugal	0,567479	12	0,678017	11
Malta	0,582819	11	1	1
Finland	0,612625	10	0,674463	12
Austria	0,795982	9	0,8878	9
Spain	0,841092	8	0,858539	10
Belgium	0,881018	7	0,953598	6
Italy	0,895023	6	0,900479	8
Netherlands	0,940875	5	0,948911	7
France	1	1	1	1
Germany	1	1	1	1
Ireland	1	1	1	1
Luxembourg	1	1	1	1

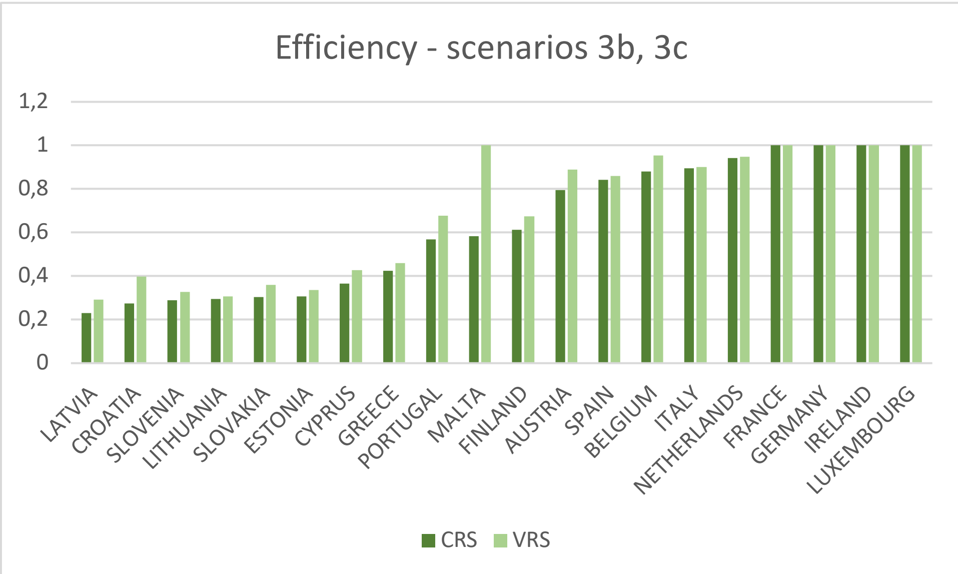

Figure 10 Results - scenarios 3b, 3c

In addition, some descriptive statistics of the efficiency scores were computed.

Table 15 Descriptive statistics of the scores - scenarios 3b,3c

Mean		Standard Deviation		Minimum		Maximum	
CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
0,630279	0,690375	0,298599	0,290985	0,230326	0,291663	1	1

Upon analyzing the results of scenarios 3b and 3c, it is evident that they are highly similar to those of scenario 3a. This conclusion is also corroborated by the correlation table below, which indicates a very strong positive correlation between their scores.

Table 16 Correlation coefficients - scenarios 3a, 3b, 3c

	Scenario 3a	Scenario 3b	Scenario 3c
Scenario 3a	1		
Scenario 3b	0,93928802	1	
Scenario 3c	0,97326406	0,9514532	1

Furthermore, Luxembourg, Ireland, Germany, and France reached the efficient boundary in all three scenarios, while Malta joined this group in scenarios 3a and 3c under the assumption of variable returns to scale. In addition, despite being inefficient, the Netherlands, Italy, Belgium, Spain, and Austria demonstrate high scores compared to the remaining countries. Lastly, Latvia, Slovenia, Slovakia, Croatia, and Lithuania exhibited the five worst performances across all three variations.

It is noteworthy that Malta achieved a score of 0.582819 in scenario 3b under the assumption of constant returns to scale, while according to the VRS variations of scenarios 3a and 3c, it was deemed efficient. Finally, the similarity in scores and rankings across all three scenarios increases the robustness of the obtained results.

4.2.4 Scenario 4

An attempt was made to apply a form of sensitivity analysis to the results of scenario 1 by changing the variables that are contained in the DEA model, using models under scenarios 2, 3a, 3b, and 3c. In order to add even more perspectives and test the reliability of the obtained outcomes, a change in the variation of the DEA model would be valuable.

The classical models that were used above are radial. These models deal with proportional changes in inputs and outputs, meaning that they reflect the proportional maximum output (input) extension (depletion) rate, which is mutual to all outputs (inputs). Nevertheless, in real-life situations, not all inputs and outputs behave in a proportional manner. (Tone, 2017)

For instance, some of the variables that are included in the current study are substitutional and do not change proportionally. Finally, another demerit of the classical radial DEA models is that they neglect the slacks. Hence, if the remaining non-radial slacks have a significant role in the estimation of efficiency, the classical approaches may lead to poor and inaccurate results.

With the aim of overcoming those limitations, non-radial approaches will be utilized in the current study. Non-radial approaches in DEA are the slack-based measure (SBM) models, first proposed by Tone (2001), which deal directly with slacks. The SBM model has three different variations, namely input, output, and non-oriented, while the current study will utilize the non-oriented approach, which combines the other two models (Zhou et al., 2014; Chang et al., 2013; Tone, 2010).

The non-oriented SBM DEA model is represented by the following program:

$$\rho_0^{\min}(x_0, y_0) = min_{\lambda, s^-, s^+} \frac{1 - (\frac{1}{m}) \sum_{i=1}^{m} (s_i^- / x_{i0})}{1 - (\frac{1}{s}) \sum_{r=1}^{s} (s_r^+ / y_{r0})}$$

Subject to constraints

$$x_{i0} = \sum_{j=1}^{n} x_{ij} * \lambda_j + s_i^- \qquad \text{i = 1, …, m}$$

$$y_{r0} = \sum_{j=1}^{n} y_{rj} * \lambda_j - s_r^+ \qquad \text{r = 1, …, s}$$

$$\lambda_j \geq 0 \; \forall j, s_i^- \geq 0 \; \forall i, s_r^+ \geq 0 \; \forall r.$$

The indicators that are used in scenario 4, are the same as those of scenario 1:

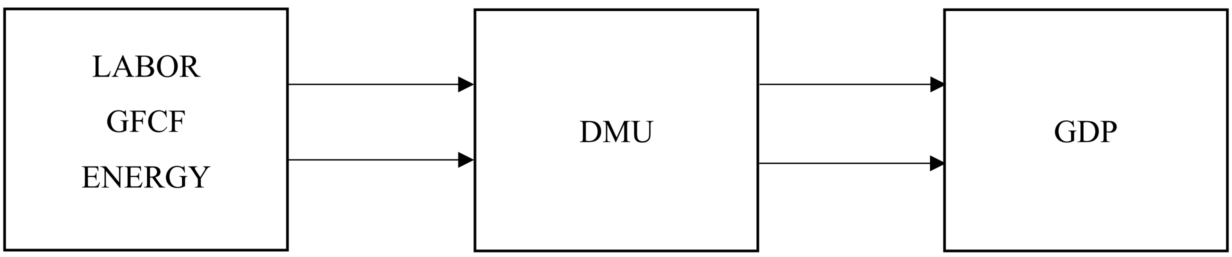

Figure 11 Indicators contained in scenario 4

The results of the model are presented below.

Table 17 Efficiency results - scenario 4

Country	Efficiency	Rank
Estonia	0,45302	20
Latvia	0,47972	19
Slovakia	0,50844	18
Croatia	0,51139	17
Lithuania	0,51598	16
Slovenia	0,52508	15
Finland	0,55833	14
Cyprus	0,56238	13
Portugal	0,56746	12
Malta	0,58371	11
France	0,58521	10
Austria	0,58607	9
Italy	0,60165	8
Spain	0,60823	7
Belgium	0,61674	6
Germany	0,63278	5
Netherlands	0,70781	4
Greece	1	1
Ireland	1	1
Luxembourg	1	1

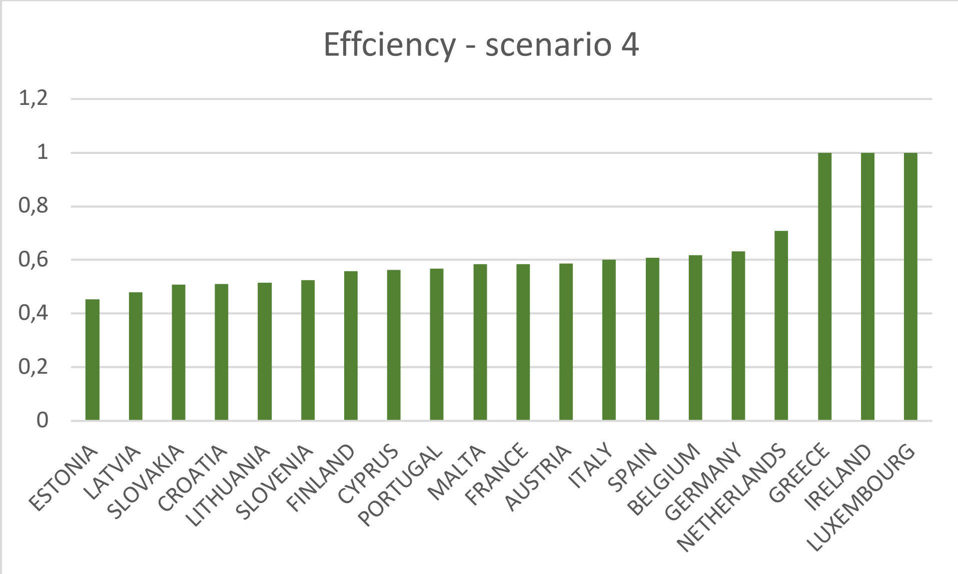

Figure 12 Efficiency results - scenario 4

In addition, some descriptive statistics of the efficiency scores were calculated.

Table 18 Descriptive statistics of the efficiency scores - scenario 4

Mean	Standard deviation	Minimum	Maximum
0,6302	0,169313	0,45302	1

The correlation coefficient between the efficiency scores obtained in scenarios 1a and 4 is 0.9278097, indicating a strong positive correlation between their outcomes. This finding implies that the SBM DEA model, applied in scenario 4, provides similar efficiency rankings to the CCR DEA model, applied in scenario 1a. This similarity in ranking may provide policymakers and governing bodies with greater confidence in the accuracy and reliability of the efficiency rankings obtained from either model.

In both scenarios 1a and 4, Greece, Luxembourg, and Ireland were identified as efficient countries, indicating that they are utilizing their resources optimally with no identifiable inefficiencies. Conversely, Latvia and Estonia were found to have the two worst performances in both scenarios. This consistency in efficiency rankings between the two scenarios for these countries suggests that they are consistently underperforming in terms of resource utilization. Additionally, other countries, such as the Netherlands and Slovenia, have identical efficiency rankings in both models.

Nevertheless, of course, there are significant differences between those scenarios. For example, comparing the efficiency scores obtained in scenarios 1a and 4 reveals that scenario 4 exhibits a greater range of scores. Furthermore, the mean and minimum values of scenario 4 are significantly lower than the corresponding values of scenario 1a. This finding suggests that the SBM DEA model, employed in scenario 4, generates a broader distribution of efficiency scores and lower overall efficiency scores compared to the CCR DEA model, employed by scenario 1a.

Moreover, as already mentioned, the SBM DEA model deals directly with slacks, namely the input excess and output shortfall. Hence, a DMU is considered efficient if and only if its ρ value is equal to 1. This condition is equivalent to the slack variables associated with both inputs and outputs being equal to zero in the optimal solution of the program. If a DMU has remaining non-radial slacks in its optimal solution, then it is deemed SBM inefficient.

In the context of the current case study, the countries of Greece, Luxemburg, and Ireland lie on the efficient frontier, indicating that these nations are optimally utilizing their "resources". In contrast, the remaining countries exhibit input and/or output inefficiencies. This fact implies that these inefficient countries may potentially achieve lower GDP than what is theoretically possible or attain a GDP value with greater energy consumption, GFCF, or labor force than necessary.

4.2.5 Scenario 5

Given that only scenario 2 encompasses all three pillars of sustainability indicators, there arises a necessity to consider an additional scenario that incorporates the social dimension of sustainability. The primary objective is to assess the reliability and robustness of the results. To conduct a sensitivity analysis on the outcomes of scenario 2, a specific input variable, precisely the labor force, will be excluded. Consequently, the inputs for scenario 5 are the subsequent measures:

1. Energy consumption and
2. Gross Fixed Capital Formation

The outputs for scenario 5 include the subsequent measures:

1. Gross domestic product and
2. Overall life satisfaction

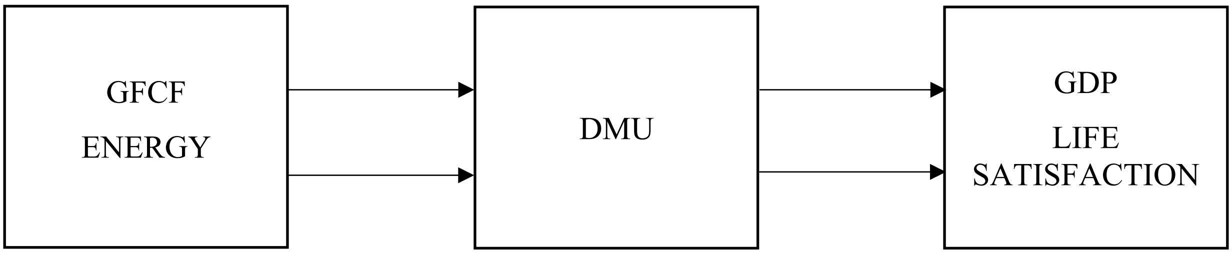

Figure 13 Indicators employed in scenario 5

The same value-based CRS DEA model utilized in scenario 2 will be employed in scenario 5 to maintain consistency and facilitate a direct comparison of the results. It must be stated that the VRS variation will not be employed because it leads to many countries achieving maximum efficiencies. Hence, its results are meaningless.

The obtained results of the model are the following:

Table 19 Results - scenario 5

Country	Efficiency	Rank
Finland	0,669537	20
Estonia	0,705811	19
Austria	0,726322	18
Slovakia	0,738531	17
Belgium	0,750593	16
France	0,751755	15
Lithuania	0,77403	14
Slovenia	0,785867	13
Croatia	0,802118	12
Italy	0,807885	11
Portugal	0,817809	10
Latvia	0,825989	9
Germany	0,82715	8
Spain	0,850058	7
Netherlands	0,898061	6
Cyprus	0,990776	5
Greece	1	1
Ireland	1	1
Luxembourg	1	1
Malta	1	1

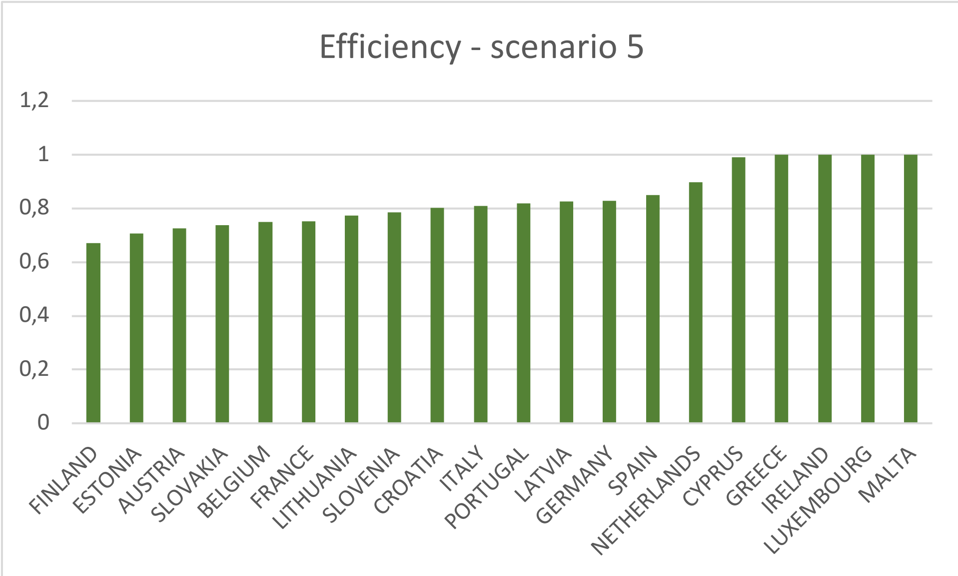

Figure 14 Results - scenario 5

Moreover, some descriptive statistics of the efficiency scores were computed

Table 20 Descriptive statistics - scenario 5

Mean	Standard deviation	Minimum	Maximum
0,836115	0,108814	0,669537	1

Upon examination of the results table and descriptive statistics table of scenario 5 and in comparison to those of scenario 2, it becomes apparent that the two scenarios exhibit highly similar results. This observation is further substantiated by calculating the correlation coefficient between scenarios 2 and 5, which yields a value of 0.99807049. Consequently, it can be inferred that although scenario 5 omits the indicator of labor force, it still produces almost identical efficiency scores to those obtained in scenario 2. This finding indicates that the variable of labor could be unnecessary for constructing the index of scenario 2.

More precisely, 14 countries obtained the same outcomes in both scenarios. Furthermore, Croatia, Estonia, Finland, Slovakia, and Slovenia experienced only minor reductions in their scores when the indicator of the labor force was omitted from scenario 2. The rankings derived from both scenarios are also identical, as demonstrated in Table 21 below. This similarity in results and rankings between scenarios 2 and 5 is indicative of the robustness and reliability of the obtained sustainability scores.

Table 21 Rankings - scenario 2 and scenario 5

Country	Rank 2	Rank 5
Austria	18	18
Belgium	16	16
Croatia	12	12
Cyprus	5	5
Estonia	19	19
Finland	20	20
France	15	15
Germany	8	8
Greece	1	1
Ireland	1	1
Italy	11	11
Latvia	9	9
Lithuania	14	14
Luxembourg	1	1
Malta	1	1
Netherlands	6	6
Portugal	10	10
Slovakia	17	17
Slovenia	13	13
Spain	7	7

4.2.6 Scenario 6

Building upon the concept introduced in scenario 5, scenario 6 aims to investigate the potential impact of a second transformation of variables on the outcomes of scenario 2. The objective is to examine whether altering the underlying structure of the data will significantly alter the results obtained in scenario 2. This analysis will provide insights into the sensitivity of the model employed in scenario 2 to changes in variables and will help in assessing its findings' robustness.

Within the context of scenario 6, the following set of input variables will be used:

1. The Energy consumption and

2. The labor force

As outputs, the following set of indicators will be used.

1. The gross domestic product and
2. The overall life satisfaction

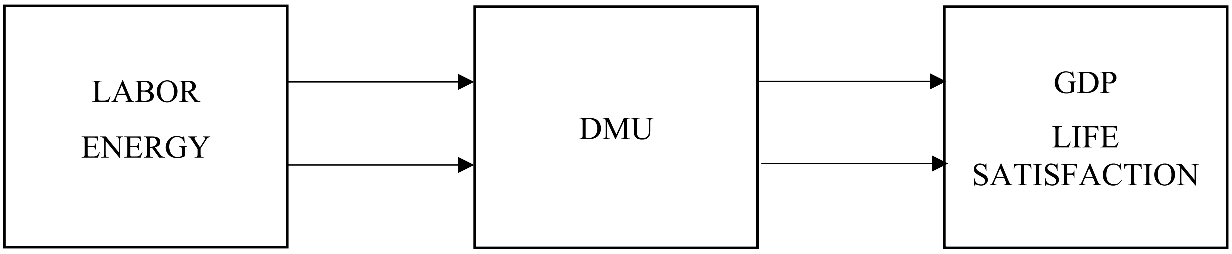

Figure 15 Indicators utilized in scenario 6

Furthermore, the value-based CRS DEA model was employed once more, and the ensuing results were acquired.

Table 22 Efficiency results - scenario 6

Country	Efficiency	Rank
Slovakia	0,263626	20
Croatia	0,293382	19
Greece	0,314298	18
Lithuania	0,351167	17
Portugal	0,352899	16
Latvia	0,358055	15
Spain	0,393006	14
Italy	0,411926	13
Slovenia	0,419678	12
Finland	0,4492	11
France	0,452294	10
Germany	0,464197	9
Austria	0,482323	8
Estonia	0,483124	7
Cyprus	0,51729	6
Belgium	0,519915	5
Netherlands	0,523494	4
Ireland	1	1

Luxembourg	1	1
Malta	1	1

It must be noted that the VRS variation will not be applied because it leads to many countries attaining very high or maximum efficiency scores. Hence, its results are meaningless.

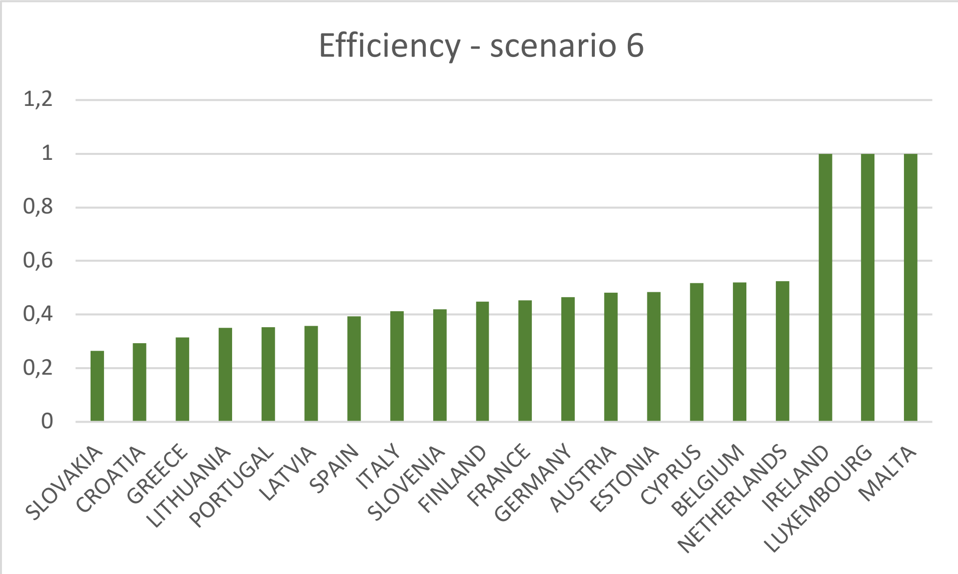

Figure 16 Efficiency rankings - scenario 6

Moreover, some descriptive statistics of the efficiency scores were computed.

Table 23 Descriptive statistics of the efficiency scores - scenario 6

Mean	Standard deviation	Minimum	Maximum
0,502494	0,227168	0,263626	1

Upon examination of the results of scenario 6, it is evident that its efficiency scores differ significantly from those of scenarios 2 and 5, which are identical. This fact highlights the sensitivity of efficiency scores in response to variable changes. Specifically, 17 countries exhibited notably lower efficiency scores when compared to scenarios 2 and 5.

The above outcome suggests that although Eurozone countries may achieve favorable outputs in terms of the level of their available gross fixed capital formation (GFCF), they do not get close to their potential output in terms of their available labor force. This fact implies a potential misallocation of resources or inefficiencies in the utilization of labor, which could hinder countries' economic growth and competitiveness. Further analysis is required to determine the underlying causes and potential policy interventions to address this issue.

Although the majority of countries achieved lower efficiencies in scenario 6, Ireland, Luxembourg, and Malta remained at the efficient frontier. This finding increases the robustness of the results of the previous scenarios and suggests that these countries could employed as role models for the rest of the countries.

In the context of the disparity in rankings between scenarios 2, 5, and scenario 6, a redistribution is evident, which can be attributed to the fact that certain countries experienced larger decreases than others.

4.2.7 Scenario 7

In an attempt to further evaluate the sensitivity of scenario 2 results, the input variables will be altered again in scenario 7. More specifically, the following indicators will be incorporated as inputs:

1. Labor force and
2. Gross fixed capital formation

The following measures were used as outputs in scenario 7:

1. Gross domestic product and
2. Overall life satisfaction

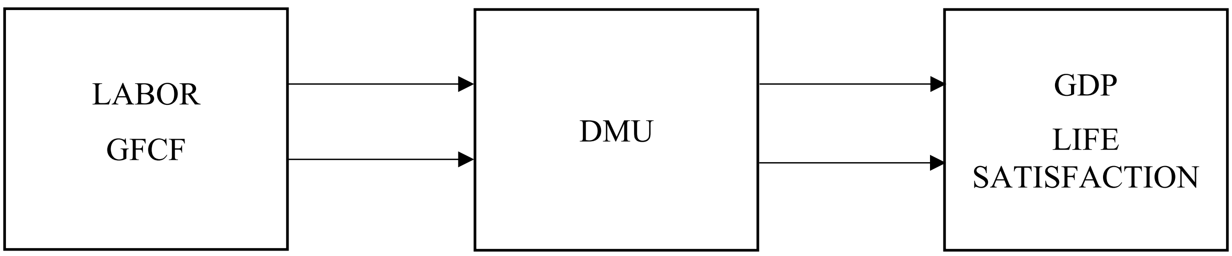

Figure 17 Indicators used in scenario 7

It is noteworthy that scenario 7 does not incorporate environmental variables following the subtraction of energy consumption, thereby indicating that this index primarily measures socio-economic efficiency rather than sustainability among the members of the Eurozone under assessment.

Furthermore, the same output-oriented classic CRS was employed and yielded the following results.

Table 24 Efficiency results - scenario 7

Country	Efficiency	Rank
France	0,664814	20
Austria	0,669169	19
Finland	0,697725	18
Estonia	0,708602	17
Belgium	0,710676	16
Slovakia	0,745746	15
Italy	0,746079	14
Portugal	0,746269	13
Germany	0,752451	12
Lithuania	0,763435	11
Spain	0,774323	10
Netherlands	0,800506	9
Slovenia	0,804195	8
Croatia	0,807591	7
Ireland	0,817882	6
Latvia	0,825989	5
Cyprus	0,990776	4
Greece	1	1
Luxembourg	1	1
Malta	1	1

The VRS variation was not applied because it resulted in many countries being deemed efficient. Hence, its results are not meaningful.

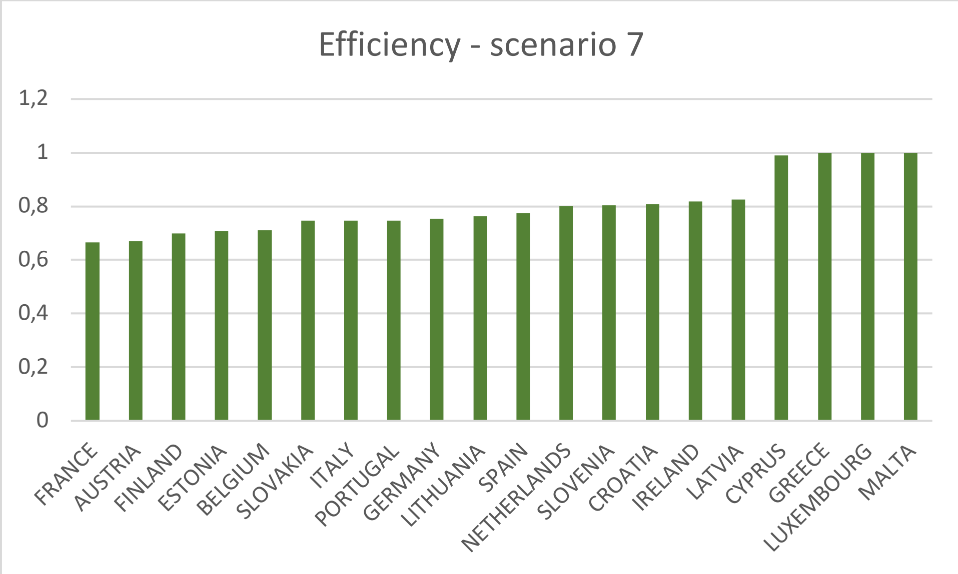

Figure 18 Efficiency results of scenario 7

Moreover, some descriptive statistics of the efficiency scores were computed.

Table 25 descriptive statistics - scenario 7

Mean	Standard deviation	Minimum	Maximum
0,801311	0,110647	1	0,664814

Upon analyzing the results of scenario 7, it is evident that there is a similarity with those of scenarios 2 and 5. The high correlation coefficients between their scores support this similarity. To be more specific, the results of scenario 7 exhibit correlation coefficients of 0,89588353 and 0,88546982 with those of scenarios 2 and 5, respectively, revealing a strong positive correlation between them. Despite that, as already mentioned, scenario 7 measures socio-economic efficiency and not sustainability. Hence, the results of those scenarios must not be compared further.

Regarding the scenario's 7 scores, Malta, Luxembourg, and Greece reached the efficient boundary, while Cyprus was the best inefficient country with a score of 0,990776. Furthermore, Latvia, Ireland, Croatia, and Slovenia exhibited scores higher than the average value of 0,801311. The rest of the countries obtained scores above 0,70, except Finland, Austria, and France, which attained scores of 0,697725, 0,669169, and 0,664814, respectively.

Finally, it is worth noting that despite Ireland being efficient in all variations so far, it is deemed inefficient according to the scores of scenario 7

5. Discussion

In the preceding section, two DEA variations under 10 different scenarios were implemented, aiming to produce composite efficiency indicators and conduct a comparative assessment of the performances of Eurozone countries. These scenarios can be classified into three distinct groups. The first group, consisting of scenarios 1a, 1b, 3a, 3b, 3c, and 4, includes economic and environmental variables, whereas scenarios 2, 5, and 6 integrate variables representing all three dimensions of sustainability. Finally, scenario 7 uses economic and social indicators.

Consequently, the composite indices constructed under the scenarios of the first group serve to quantify the eco-efficiency of 20 members of the Eurozone. In contrast, the scenarios belonging to the second group provide a gauge of the sustainability efficiency of those members. Of course, scenario 7 measures the socio-economic efficiency of the countries under assessment.

All scenarios of the study are summarized in table 26 below.

Table 26 Scenarios of the study

Scenario	Inputs	Outputs	Index	RTS	Orientation	DEA variation
1a	labor, energy, GFCF	GDP	Eco-efficiency	CRS	output	classic
1b	labor, energy, GFCF	GDP	Eco-efficiency	VRS	output	classic
2	labor, energy, GFCF	GDP, life satisfaction	Sustainability	CRS	output	classic
3a	labor, energy	GDP, GHG	Eco-efficiency	VRS	output	classic
3b	labor, energy, GHG	GDP	Eco-efficiency	CRS	output	classic
3c	labor, energy, GHG	GDP	Eco-efficiency	VRS	output	classic
4	labor, energy, GFCF	GDP	Eco-efficiency	CRS	output	SBM
5	energy, GFCF	GDP, life satisfaction	Sustainability	CRS	output	classic
6	labor, energy	GDP, life satisfaction	Sustainability	CRS	output	classic
7	labor , GFCF	GDP, life satisfaction	Socio-economic	CRS	output	classic

A thorough and holistic analysis of each category of indices is necessary to determine the efficient and inefficient countries and to acquire further information. More specifically, the scores obtained in each scenario will be scrutinized and correlated to identify any significant relationships. The following tables summarize the results of the scores of the eco-efficiency scenarios and their respective correlation coefficients.

Table 27 Scores obtained in the eco-efficiency scenarios

Country	Scenario 1a	Scenario 1b	Scenario 3a	Scenario 3b	Scenario 3c	Scenario 4
Austria	0,726322	0,817488	0,775964	0,795982	0,8878	0,58607
Belgium	0,750593	0,877362	0,910548	0,881018	0,953598	0,61674
Croatia	0,737463	0,783098	0,238111	0,274901	0,39653	0,51139
Cyprus	0,824552	1	0,428519	0,365764	0,428519	0,56238
Estonia	0,633489	0,66028	0,336965	0,306278	0,336965	0,45302
Finland	0,697725	0,821558	0,520592	0,612625	0,674463	0,55833
France	0,751755	0,965326	1	1	1	0,58521
Germany	0,82715	1	1	1	1	0,63278
Greece	1	1	0,357106	0,423747	0,46017	1
Ireland	1	1	1	1	1	1
Italy	0,807885	0,975239	0,889838	0,895023	0,900479	0,60165
Latvia	0,682505	0,791653	0,247806	0,230326	0,291663	0,47972
Lithuania	0,746536	0,755219	0,305483	0,295834	0,305483	0,51598
Luxembourg	1	1	1	1	1	1
Malta	0,796039	1	1	0,582819	1	0,58371
Netherlands	0,898061	1	0,923591	0,940875	0,948911	0,70781
Portugal	0,817809	0,843946	0,41132	0,567479	0,678017	0,56746
Slovakia	0,72372	0,723736	0,241326	0,304173	0,359073	0,50844
Slovenia	0,735073	0,760237	0,315937	0,287642	0,327298	0,52508
Spain	0,850058	1	0,839955	0,841092	0,858539	0,60823

Table 28 Correlation coefficients - eco-efficiency scenarios

	Scenario 1a	Scenario 1b	Scenario 3a	Scenario 3b	Scenario 3c	Scenario 4
Scenario 1a	1					
Scenario 1b	0,77402841	1				
Scenario 3a	0,45450475	0,73634803	1			
Scenario 3b	0,4993233	0,69091547	0,93928802	1		
Scenario 3c	0,45524996	0,71365375	0,97326406	0,9514532	1	
Scenario 4	0,92780972	0,63620642	0,44178086	0,50148007	0,44240911	1

In addition, the rankings of every eco-efficiency scenario are summarized in table 29 below.

Table 29 Rankings - eco-efficiency scenarios

Country	Rank 1a	Rank 1b	Rank 3a	Rank 3b	Rank 3c	Rank 4
Austria	16	14	10	9	9	9
Belgium	12	11	7	7	6	6
Croatia	14	16	20	19	15	17
Cyprus	7	1	12	14	14	13
Estonia	20	20	15	15	17	20
Finland	18	13	11	10	12	14
France	11	10	1	1	1	10
Germany	6	1	1	1	1	5
Greece	1	1	14	13	13	1
Ireland	1	1	1	1	1	1
Italy	9	9	8	6	8	8
Latvia	19	15	18	20	20	19
Lithuania	13	18	17	17	19	16
Luxembourg	1	1	1	1	1	1
Malta	10	1	1	11	1	11
Netherlands	4	1	6	5	7	4
Portugal	8	12	13	12	11	12
Slovakia	17	19	19	16	16	18
Slovenia	15	17	16	18	18	15
Spain	5	1	9	8	10	7

In summary, the results of the correlation analysis of the eco-efficiency scenarios reveal that their scores exhibit either moderate or high correlation with one another. Hence, it could be concluded that their results are robust. Nevertheless, aiming to increase the validity and reliability of the findings, the scores obtained by all scenarios will be taken into account simultaneously.

In light of the tables above, it is observed that the countries of Luxembourg and Ireland attained the maximum efficiency score across all scenarios despite the changes in indicator sets and DEA variations employed. This fact suggests that these countries may be considered relatively efficient in terms of eco-efficiency. As a result, they could serve as exemplary models for other Eurozone countries seeking to enhance their eco-efficiency performance.

The implications of this finding are significant, as it highlights the potential benefits that can be derived from adopting best practices and strategies employed by these countries in their pursuit of sustainable development. Policymakers and governing bodies in other Eurozone countries may benefit from studying the experiences of Luxembourg and Ireland in order to identify opportunities for improvement and optimization in their own policies and operations.

Furthermore, countries such as Germany and the Netherlands obtained relatively high scores across all scenarios. Hence, even though they are not efficient, they still exhibit decent sustainability performance compared to the rest of the Eurozone members. Ultimately, countries such as Croatia, Finland, Latvia, Lithuania, Slovakia, and Slovenia exhibited the poorest performances across all eco-efficiency scenarios. These results suggest that these countries require significant improvements in their policies to enhance their overall eco-efficiency.

Following the assessment of the eco-efficiency performance of the Eurozone members under evaluation, it is also imperative to evaluate their sustainability performance, which was the initial goal of the current research. Hence, in order to do so, the sustainability indicators generated in scenarios 2, 5, and 6 must be analyzed thoroughly. The following tables present a summary of the scores obtained from the sustainability indicators and their respective correlation coefficients.

Table 30 Scores obtained by the sustainability scenarios

Country	Scenario 2	Scenario 5	Scenario 6
Austria	0,726322	0,726321906	0,482322867
Belgium	0,750593	0,750592968	0,519915358
Croatia	0,807591	0,80211759	0,293382465
Cyprus	0,990776	0,990775877	0,517290433
Estonia	0,708602	0,705810941	0,483124463
Finland	0,697725	0,669536748	0,449199975
France	0,751755	0,751755349	0,452294263
Germany	0,82715	0,827150384	0,464196522
Greece	1	1	0,314298376
Ireland	1	1	1
Italy	0,807885	0,807884957	0,411926084
Latvia	0,825989	0,825988915	0,358054618

Lithuania	0,77403	0,774029754	0,351166752
Luxembourg	1	1	1
Malta	1	1	1
Netherlands	0,898061	0,898061086	0,52349443
Portugal	0,817809	0,8178086	0,352899244
Slovakia	0,745746	0,738530619	0,263626178
Slovenia	0,804195	0,785866968	0,419677855
Spain	0,850058	0,850058229	0,393006064

Table 31 Correlations coefficients between the scores of the sustainability scenarios.

	Scenario 2	Scenario 5	Scenario 6
Scenario 2	1		
Scenario 5	0,99807049	1	
Scenario 6	0,60788387	0,60311134	1

In addition, the rankings of the sustainability scenarios are presented below.

Table 32 Rankings of the sustainability indices.

Country	Rank 2	Rank 5	Rank 6
Austria	18	18	8
Belgium	16	16	5
Croatia	12	12	19
Cyprus	5	5	6
Estonia	19	19	7
Finland	20	20	11
France	15	15	10
Germany	8	8	9
Greece	1	1	18
Ireland	1	1	1
Italy	11	11	13
Latvia	9	9	15
Lithuania	14	14	17
Luxembourg	1	1	1
Malta	1	1	1
Netherlands	6	6	4
Portugal	10	10	16

Slovakia	17	17	20
Slovenia	13	13	12
Spain	7	7	14

The correlation analysis reveals that the results of the different sustainability scenarios exhibit either medium or high correlation, thus enhancing the robustness of the outcomes. As already mentioned, the divergence observed in the outcomes of scenario 6 compared to those of scenarios 2 and 5 could be attributed to the inability of certain countries to get close to their potential output levels with regard to their available labor force.

Moreover, to further enhance the validity of the current analysis, the Eurozone members will be evaluated, taking into account the results of all scenarios simultaneously. The countries that consistently attain maximum scores across all scenarios can be considered exemplary in terms of sustainability. Based on this criterion, Ireland, Luxembourg, and Malta emerged as the most efficient countries in the current study.

Therefore, these nations could serve as role models for other Eurozone members striving to improve their sustainability performance. This fact highlights the potential for knowledge transfer and best practice sharing among Eurozone countries to promote sustainable development. These results suggest that these countries' experiences and strategies could provide valuable insights for other Eurozone members seeking to enhance their sustainability performance.

In addition, it is noteworthy that Greece reached the efficient frontier in the first two variations but ranked third worst according to scenario 6. This finding could indicate that Greece does not use its labor force resources effectively. Furthermore, countries such as Cyprus and the Netherlands, although they were deemed inefficient, attained high performances across all scenarios compared to the rest of the Eurozone members.

Finally, it is evident that certain countries, such as Austria, Estonia, Finland, and Slovakia, exhibited poor results across all variations. This fact suggests that these nations require a comprehensive review of their policies in order to improve their sustainability outcomes.

6. Conclusions

As was already mentioned in the introduction, the twofold purpose of the current study is:

1. To propose the use of a mathematical programming technique, namely data envelopment analysis, with the aim of aggregating and weighting a multitude of economic, environmental, and social indicators into composite sustainability indices in such a way that reduces the potential biases that may arise from assigning weights to them, considering different sets of sub-indicators and different DEA variations

2. To utilize the acquired results of the DEA models to evaluate the sustainability efficiency of 20 Eurozone members comparatively and eventually share valuable information for analysts, policymakers, and governing bodies that aim to enhance the sustainability performance of those countries

In order to accomplish this goal, a series of steps was developed and implemented. More specifically, a literature review about DEA and sustainability for the years 2021 to 2023 was performed, aiming to identify the key indicators employed by the authors to gauge sustainability performance. This review discovers several significant findings from the recent literature. Firstly, it highlighted the absence of a unified definition of sustainability, which decreases the robustness of the obtained results and consequently affects the policymaking that arises from them. Nevertheless, the three-dimensional approach (economic, environmental, and social), which the authors have widely accepted, was adopted in the current study.

Despite the acceptance of the three-pillar approach, a multitude of different sustainability indicators were found in the literature. This phenomenon can be attributed to the fact that each author incorporates individual perceptions about economic, environmental, and social variables. In addition, several authors tried to encompass R&D, technology, and innovation indicators in their models, increasing further the variety of indicators employed in the relevant literature. All of the above eventually decrease the robustness of the obtained results since every author measures sustainability or sustainable development in a slightly different way.

The next step was the implementation of DEA to construct composite sustainability indices. In order to choose the variables of the study, the frequency of appearance of the indicators used in the literature about DEA and sustainability for the years 2021 to 2023 was taken into consideration. Eventually, 2 DEA models (under 10 different scenarios) that generated 10 composite indices were implemented. Afterwards, those indices were divided into six eco-efficiency, three sustainability indicators, and one socio-economic indicator.

After the presentation and interpretation of the acquired results, all of the sustainability scenarios were taken into account simultaneously to assess the relative efficiency of Eurozone members. The main results yielded from this analysis are summarized below:

1. Ireland, Luxembourg, and Malta reached the efficient boundary across all three sustainability scenarios despite the changes in indicators. Thus, it could be said that these countries are comparatively efficient, and so they could act as exemplary models to the rest of the Eurozone members.
2. In light of the worsening in sustainability scores of 17 countries in scenario 6 compared to the previous sustainability scenarios, it could be concluded that most of the Eurozone members do not get close to their potential output levels (GDP and Overall Life Satisfaction) proportional to their available labor force.
3. Finally, countries such as Austria, Estonia, Finland, and Slovakia exhibited low sustainability efficiency scores across all variations. This finding indicates an urgent need for them to reconsider their policies aiming to enhance their sustainability performance.

Considering and analyzing the six eco-efficiency scenarios in a similar manner, the following main results are acquired:

1. Ireland and Luxembourg obtained the maximum efficiency across all six variations that used different sets of indicators and different DEA variations (classic and slack-based measure approaches). Thus, it could be concluded that these countries are relatively efficient, and so they could act as exemplary models to the rest of the Eurozone members.
2. Moreover, the countries of Croatia, Finland, Latvia, Lithuania, Slovakia, and Slovenia obtained low eco-efficiency scores across all scenarios. This finding suggests that there is an imperative need for these countries to reevaluate, replan, and eventually improve their eco-efficiency policies.

Although efforts were made to increase the conclusions' robustness, this study possesses certain limitations. To be more specific, DEA weighs the indicators objectively. However, under the variations used, it does so in a way that assigns different weights to each country. In addition, the number of parameters used in this study's models is constrained by the number of countries under assessment. If the number of inputs and outputs were increased further, it would result in more DMUs being deemed as efficient, and the obtained scores would not be meaningful. This constraint makes it challenging to include indicators that represent all dimensions of sustainability, resulting in using only one social variable (overall life satisfaction) and zero R&D variables.

Furthermore, treating the transformation process as a black box is another demerit of the current research. Moreover, the study concerns data for the year 2021. It would be beneficial to apply a dynamic DEA model that takes into account several years to provide a comprehensive evaluation of the countries under assessment. All of the above eventually decrease the validity of the research's outcomes.

One potential direction for further research that would help to overcome some of the above limitations is the implementation of two-stage DEA models. This approach would not only provide a more detailed explanation of the transformation process but also allow for the inclusion of additional parameters without artificially inflating the scores of the DMUs. Moreover, the integration of machine learning, game theory or other quantitative techniques alongside DEA could enhance the results' validity and simultaneously provide additional essential information.

In summary, the current thesis proposes a methodological framework for developing composite sustainability indicators, which can have several practical applications in various contexts. For example, this framework could serve as a supplementary instrument for evaluating the sustainability performance of countries that are going to adopt the euro currency by incorporating them into the set of DMUs. In addition, the outcomes of this approach could also identify efficient countries and enable less efficient countries to analyze their policies with the aim of improving their performances. Finally, in light of the above, DEA and the indicators that it generates constitute beneficial quantitative tools in the arsenal of policymakers and governing bodies that can assist them in making more informed decisions and creating more sustainable policies.

References

1. Adler, M. (2011). Well-being and fair distribution beyond cost-benefit analysis. *Oxford University Press*. https://doi.org/10.1093/acprof:oso/9780195384994.001.0001

2. Afzalinejad, M. (2021). Evaluating radial efficiency considering environmental factors: A generalization of classical DEA. *Measurement*, *179*, 109497. https://doi.org/10.1016/j.measurement.2021.109497

3. Ali, A., & Seiford, L. M. (1990). Translation invariance in data envelopment analysis. *Operations Research Letters*, *9*(6), 403–405. https://doi.org/10.1016/0167-6377(90)90061-9

4. Alidrisi, H. (2021). The development of an efficiency-based Global Green Manufacturing Innovation index: An input-oriented DEA approach. *Sustainability*, *13*(22), 12697. https://doi.org/10.3390/su132212697

5. Alves, J. L., & Dumke de Medeiros, D. (2015). Eco-efficiency in micro-enterprises and small firms: A case study in the Automotive Services Sector. *Journal of Cleaner Production*, *108*, 595–602. https://doi.org/10.1016/j.jclepro.2015.07.063

6. Banerjee, A. V., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. PublicAffairs.

7. Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078–1092. Doi: 10.1287/mnsc.30.9.1078

8. Beliakov, G., Pradera, A., & Calvo, T. (2007). Aggregation functions: A guide for practitioners. *Springer*.

9. Benítez, R., Coll-Serrano, V., & Bolós, V. J. (2021). Dear-shiny: An interactive web app for data envelopment analysis. *Sustainability*, *13*(12), 6774. https://doi.org/10.3390/su13126774

10. Brzyska, J., & Szamrej-Baran, I. (2023). The COVID-19 pandemic and the implementation of sustainable development goals: The EU perspective. *Sustainability*, *15*(18), 13503. https://doi.org/10.3390/su151813503

11. Camioto, F. de, & Pulita, A. C. (2022). Efficiency evaluation of sustainable development in BRICS and G7 countries: A Data Envelopment Analysis Approach. *Gestão & Produção*, *29*. https://doi.org/10.1590/1806-9649-2022v29e022

12. Chang, Y.-T., Zhang, N., Danao, D., & Zhang, N. (2013). Environmental efficiency analysis of transportation system in China: A non-radial DEA approach. *Energy Policy*, *58*, 277–283. https://doi.org/10.1016/j.enpol.2013.03.011

13. Charnes, A., & Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, *9*(3–4), 181–186. https://doi.org/10.1002/nav.3800090303

14. Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of Decision Making Units. *European Journal of Operational Research*, *2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

15. Chen, Y., & Zhu, J. (2004). Measuring Information Technology's indirect impact on firm performance. *Information Technology and Management*, *5*(1/2), 9–22. https://doi.org/10.1023/b:item.0000008075.43543.97

16. Chen, Z. *et al.* (2022) 'A robust network DEA model for Sustainability Assessment: An application to Chinese provinces', *Operational Research*, 22(1), pp. 235–262. doi:10.1007/s12351-020-00553-x.

17. Coll-Serrano, V., Bolós, V., & Benítez Suárez, R. (2023). *Dear: Conventional and fuzzy data envelopment analysis*. R package version 1.4.1. https://cran.r-project.org/web//packages/deaR/deaR.pdf

18. Cook, W. D., & Zhu, J. (2005). *Modeling Performance Measurement: Applications and Implementation Issues in DEA*. https://doi.org/10.1007/b104529

19. Cooper, W. W. (2010). *Handbook on Data Envelopment Analysis*. Springer.

20. Cooper, W. W., Seiford, L. M., & Tone, K. (2006). *Introduction to data envelopment analysis and its uses with DEA-Solver Software and references*. Springer.

21. da Silva, A. V., de Oliveira Ribeiro, C., & Rego, E. E. (2023). A data envelopment analysis approach to measuring socio-economic efficiency due to renewable energy sources in Brazilian regions. *Brazilian Journal of Chemical Engineering*. https://doi.org/10.1007/s43153-023-00356-5

22. Debreu, G. (1951). The coefficient of resource utilization. *Econometrica*, *19*(3), 273. https://doi.org/10.2307/1906814

23. Diewert, W. E. (1973). Functional forms for profit and transformation functions. *Journal of Economic Theory*, *6*(3), 284–316.https://doi.org/10.1016/0022-0531(73)90051-3

24. Elkington, J. (1998). Accounting for the triple bottom line. *Measuring Business Excellence*, *2*(3), 18–22. https://doi.org/10.1108/eb025539

25. Emrouznejad, A., Marra, M., Yang, G., & Michali, M. (2023). Eco-efficiency considering NetZero and Data Envelopment Analysis: A critical literature review. *IMA Journal of Management Mathematics*, *34*(4), 599–632. https://doi.org/10.1093/imaman/dpad002

26. European Commission. (2005). Impact assessment guidelines . https://www.orgapet.orgap.org/annexes/annex_A2-1.pdf

27. European Commission. (2007). Competitiveness and Innovation Framework Programme (CIP). Brussels. https://ec.europa.eu/cip/files/docs/factsheets_en.pdf

28. Eurostat. (2023). SDG 6 - Clean water and sanitation: Ensure availability and sustainable management of water and sanitation for all. Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=SDG_6_-_Clean_water_and_sanitation

29. Eurostat. (2023). SDG 7 - Affordable and clean energy: Ensure access to affordable, reliable, sustainable and modern energy for all. Statistics Explained.

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=SDG_7_-_Affordable_and_clean_energy

30. Eurostat. (2023). Statistics in development cooperation - statistics in the European Commission project cycle. Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php/Statistics_in_development_cooperation_-_statistics_in_the_European_Commission_project_cycle

31. Fathi, B., Ashena, M., & Anisi, M. (2022). Efficiency evaluation of sustainability indicators in a two-stage network structure: A Nash Bargaining Game Approach. *Environment, Development and Sustainability*, *25*(2), 1832-1851. https://doi.org/10.1007/s10668-022-02325-3

32. Figge, F., & Hahn, T. (2004). Sustainable value added—measuring corporate contributions to sustainability beyond eco-efficiency. *Ecological Economics*, *48*(2), 173–187. https://doi.org/10.1016/j.ecolecon.2003.08.005

33. Førsund, F. R., & Sarafoglou, N. (2002). On the origins of data envelopment analysis. *Journal of Productivity Analysis*, *17*(1/2), 23–40. https://doi.org/10.1023/a:1013519902012

34. Gan, X., Fernandez, I. C., Guo, J., Wilson, M., Zhao, Y., Zhou, B., & Wu, J. (2017). When to use what: Methods for weighting and aggregating sustainability indicators. *Ecological Indicators*, *81*, 491–502. https://doi.org/10.1016/j.ecolind.2017.05.068

35. Georgiou, A. C., Thanassoulis, E., & Papadopoulou, A. (2021). Using data envelopment analysis in Markovian Decision making. *European Journal of Operational Research*, *298*(1), 276–292. https://doi.org/10.1016/j.ejor.2021.06.050

36. Gibson, R. B. (2001). *Specification of sustainability-based Environmental Assessment Decision Criteria and implications for determining "significance" in environmental assessment*. Canadian Environmental Assessment Agency.

37. Gieryn, T. (1999). *Cultural boundaries of science*. Chicago University Press.

38. Godfrey, L., & Todd, C. (2001). *Defining thresholds for freshwater sustainability indicators within the context of South African water resource management*. In 2nd WARFA/Waternet Symposium: Integrated Water Resource Management: Theory, Practice, Cases (pp. 1-10). Retrieved from http://www.waternetonline.ihe.nl/aboutWN/pdf/godfrey.pdf

39. 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

40. Goto, M., & Sueyoshi, T. (2023). Sustainable development and convergence under energy sector transition in industrial nations: An application of DEA Environmental Assessment. *Socio-Economic Planning Sciences*, *87*, 101316. https://doi.org/10.1016/j.seps.2022.101316

41. Grzebyk, M., & Stec, M. (2015). Sustainable development in EU countries: Concept and rating of levels of development. *Sustainable Development*, *23*(2), 110–123. https://doi.org/10.1002/sd.1577

42. Guitouni, A., & Martel, J.-M. (1998). Tentative guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research*, *109*(2), 501–521. https://doi.org/10.1016/s0377-2217(98)00073-3

43. Hajiagha, S. H., Hashemi, S. S., & Mahdiraji, H. A. (2016). Fuzzy C-means based data envelopment analysis for mitigating the impact of Units' heterogeneity. *Kybernetes*, *45*(3), 536–551. https://doi.org/10.1108/k-07-2015-0176

44. Halkos, G., & Petrou, K. N. (2018). *A critical review of the main methods to treat undesirable outputs in DEA. Multidisciplinary Digital Publishing Institute. https://mpra.ub.uni-muenchen.de/90374/*

45. Halkos, G., & Petrou, K. N. (2019). Treating undesirable outputs in DEA: A critical review. *Economic Analysis and Policy*, *62*, 97–104. https://doi.org/10.1016/j.eap.2019.01.005

46. He, J., Wan, Y., Feng, L., Ai, J., & Wang, Y. (2016). An integrated data envelopment analysis and Emergy-based ecological footprint methodology in evaluating sustainable development, a case study of Jiangsu Province,

China. *Ecological Indicators*, *70*, 23–34. https://doi.org/10.1016/j.ecolind.2016.05.042

47. Hermans, E., Van den Bossche, F., & Wets, G. (2008). Combining road safety information in a performance index. *Accident Analysis & Prevention*, *40*(4), 1337–1344. https://doi.org/10.1016/j.aap.2008.02.004

48. Hezri, A. A., & Dovers, S. R. (2006). Sustainability Indicators, policy and governance: Issues for ecological economics. *Ecological Economics*, *60*(1), 86–99. https://doi.org/10.1016/j.ecolecon.2005.11.019

49. Holling, C. S. (1973). Resilience and stability of Ecological Systems. *Annual Review of Ecology and Systematics*, *4*(1), 1–23. https://doi.org/10.1146/annurev.es.04.110173.000245

50. Huppes, G., & Ishikawa, M. (2005). A framework for Quantified Eco-Efficiency Analysis. *Journal of Industrial Ecology*, *9*(4), 25–41. https://doi.org/10.1162/108819805775247882

51. Ishizaka, A., & Nemery, P. (2013). Multi-criteria decision analysis: methods and software. *John Wiley & Sons*. https://doi.org/10.1002/9781118644898

52. Karadayi, M. A., & Ekinci, Y. (2018). Evaluating R&D performance of EU countries using categorical DEA. *Technology Analysis and Strategic Management*, *31*(2), 227–238. https://doi.org/10.1080/09537325.2018.1493191

53. Keeney, R. L. (1973). A decision analysis with multiple objectives: The mexico city airport. *The Bell Journal of Economics and Management Science*, *4*(1), 101. https://doi.org/10.2307/3003141

54. Keeney, R. L. (1974). Multiplicative utility functions. *Operations Research*, *22*(1), 22–34. https://doi.org/10.1287/opre.22.1.22

55. KEI. (2005). Knowledge Economy Indicators. Work Package 7, State of the Art Report on Simulation and Indicators.

56. Kiani Mavi, R., & Kiani Mavi, N. (2021). National eco-innovation analysis with Big Data: A common-weights model for dynamic DEA. *Technological*

Forecasting and Social Change, *162*, 120369. https://doi.org/10.1016/j.techfore.2020.120369

57. Koopmans, T. (1951). Analysis of production as an efficient combination of activities. In Activity “Analysis of Production and Allocation” (pp. 33-97). John Wiley and Sons.

58. Kuhlman, T., & Farrington, J. (2010). What is sustainability? *Sustainability*, *2*(11), 3436–3448. https://doi.org/10.3390/su2113436

59. Kuosmanen, T., & Kortelainen, M. (2005). Measuring eco-efficiency of production with data envelopment analysis. *Journal of Industrial Ecology*, *9*(4), 59–72. https://doi.org/10.1162/108819805775247846

60. Lovell, C. A. K., Pastor, J. T., & Turner, J. A. (1995). Measuring macroeconomic performance in the OECD: A comparison of european and non-European countries. *European Journal of Operational Research*, *87*(3), 507–518. https://doi.org/10.1016/0377-2217(95)00226-x

61. Lu, L.-C., Chiu, S.-Y., Chiu, Y., & Chang, T.-H. (2022). Sustainability efficiency of climate change and global disasters based on greenhouse gas emissions from the parallel production sectors – a modified dynamic parallel three-stage network DEA model. *Journal of Environmental Management*, *317*, 115401. https://doi.org/10.1016/j.jenvman.2022.115401

62. Lubsanova, N. B., Maksanova, L. B.-Z., Eremko, Z. S., Bardakhanova, T. B., & Mikheeva, A. S. (2022). The eco-efficiency of Russian regions in North Asia: Their green direction of regional development. *Sustainability*, *14*(19), 12776. https://doi.org/10.3390/su141912776

63. Macedo, P., Moutinho, V., & Madaleno, M. (2023). *Advanced mathematical methods for economic efficiency analysis theory and empirical applications*. Springer.

64. May, R. M. (1977). Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature*, *269*(5628), 471–477. https://doi.org/10.1038/269471a0

65. Mayer, A. L. (2008). Strengths and weaknesses of common sustainability indices for Multidimensional Systems. *Environment International*, *34*(2), 277–291. https://doi.org/10.1016/j.envint.2007.09.004

66. Meadows, D. (1998). *Indicators and information systems for Sustainable Development*. Sustainability Institute.

67. Mikulić, J., Kožić, I., & Krešić, D. (2015). Weighting indicators of Tourism Sustainability: A Critical Note. *Ecological Indicators*, *48*, 312–314. https://doi.org/10.1016/j.ecolind.2014.08.026

68. Morse, S., & Fraser, E. D. G. (2005). Making 'dirty' nations look clean? the nation state and the problem of selecting and weighting indices as tools for measuring progress towards Sustainability. *Geoforum*, *36*(5), 625–640. https://doi.org/10.1016/j.geoforum.2004.10.005

69. Moutinho, V., & Madaleno, M. (2021). A two-stage DEA model to evaluate the technical eco-efficiency indicator in the EU countries. *International Journal of Environmental Research and Public Health*, *18*(6), 3038. https://doi.org/10.3390/ijerph18063038

70. Munda, G. (2005). "Measuring sustainability": A multi-criterion framework. *Environment, Development and Sustainability*, *7*(1), 117–134. https://doi.org/10.1007/s10668-003-4713-0

71. Munda, G., & Nardo, M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: A defensible setting. *Applied Economics*, *41*(12), 1513–1523. https://doi.org/10.1080/00036840601019364

72. Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005). Handbook on Constructing Composite Indicators. *OECD Statistics Working Papers*. https://doi.org/10.1787/533411815016

73. Nash, J. (1953). Two-person cooperative games. *Econometrica*, *21*(1), 128. https://doi.org/10.2307/1906951

74. Nash, J. F. (1950). The bargaining problem. *Econometrica*, *18*(2), 155. https://doi.org/10.2307/1907266

75. Negri, M., Cagno, E., Colicchia, C., & Sarkis, J. (2021). Integrating sustainability and resilience in the supply chain: A systematic literature review and a research agenda. *Business Strategy and the Environment*, *30*(7), 2858–2886. https://doi.org/10.1002/bse.2776

76. Ness, B., Urbel-Piirsalu, E., Anderberg, S., & Olsson, L. (2007). Categorising tools for Sustainability Assessment. *Ecological Economics*, *60*(3), 498–508. https://doi.org/10.1016/j.ecolecon.2006.07.023

77. Nikolaou, I. E., Jones, N., & Stefanakis, A. (2021). Circular economy and Sustainability: The past, the present and the future directions. *Circular Economy and Sustainability*, *1*(1), 1–20. https://doi.org/10.1007/s43615-021-00030-3

78. Nunamaker, T. R. (1985). Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation. *Managerial and Decision Economics*, *6*(1), 50–58. https://doi.org/10.1002/mde.4090060109

79. OECD. (2008). Handbook on Constructing Composite Indicators: Methodology and User Guide. *OECD Library*. https://doi.org/10.1787/9789264043466-en

80. Oltra, V., & Saint Jean, M. (2009). Sectoral Systems of Environmental Innovation: An application to the French automotive industry. *Technological Forecasting and Social Change*, *76*(4), 567–583. https://doi.org/10.1016/j.techfore.2008.03.025

81. Oras, K. (2005). Which policy frameworks matter and how to describe them: indicators linking the Lisbon strategy, sustainable development and the MDGs. New York: Statistical Commission and Economic Commission for Europe, UN Economic and Social Council.

82. Papathanasiou, J., Tsaples, G., & Blouchoutzi, A. (2021). *Urban Sustainability: A game-based approach*. Cham, Switzerland: Springer.

83. Parris, T. M., & Kates, R. W. (2003). Characterizing and measuring sustainable development. *Annual Review of Environment and Resources*, *28*(1), 559–586. https://doi.org/10.1146/annurev.energy.28.050302.105551

84. Patterson, M. G. (1996). What is energy efficiency? *Energy Policy*, *24*(5), 377–390. https://doi.org/10.1016/0301-4215(96)00017-1

85. Pollesch, N., & Dale, V. H. (2015). Applications of aggregation theory to Sustainability Assessment. *Ecological Economics*, *114*, 117–127. https://doi.org/10.1016/j.ecolecon.2015.03.011

86. Qu, J., Wang, B., & Liu, X. (2022). A modified super-efficiency Network Data Envelopment Analysis: Assessing Regional Sustainability Performance in China. *Socio-Economic Planning Sciences*, *82*, 101262. https://doi.org/10.1016/j.seps.2022.101262

87. Rennings, K. (2000). Redefining innovation — eco-innovation research and the contribution from Ecological Economics. *Ecological Economics*, *32*(2), 319–332. https://doi.org/10.1016/s0921-8009(99)00112-3

88. Sachs, J. D. (2015). *The Age of Sustainable Development*. Columbia University Press, New York https://doi.org/10.7312/sach17314

89. Sarpong, F. A., Wang, J., Cobbinah, B. B., Makwetta, J. J., & Chen, J. (2022). The drivers of Energy Efficiency Improvement among nine selected West African countries: A two-stage DEA methodology. *Energy Strategy Reviews*, *43*, 100910. https://doi.org/10.1016/j.esr.2022.100910

90. Schaltegger, S., & Müller, J. (1996). *Corporate Environmental Accounting*. John Wiley & Sons.

91. Scheel, H. (2001). Undesirable outputs in efficiency valuations. *European Journal of Operational Research*, *132*(2), 400–410. https://doi.org/10.1016/s0377-2217(00)00160-0

92. Schmidheiny, S., & Zorraquin, F. J. L. (1996). *Financing change, the financial community, eco-efficiency, and sustainable development. MIT Press, Cambridge*

93. Scoones, I. (2007). Sustainability. *Development in Practice*, *17*(4–5), 589–596. https://doi.org/10.1080/09614520701469609

94. SCRCSSP. (1997). Data Envelopment Analysis: A technique for measuring the efficiency of government service delivery. AGPS, Canberra. https://www.pc.gov.au/research/supporting/data-envelopment-analysis

95. Seiford, L. M., & Zhu, J. (2001). Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research*, *142*(1), 16–20. https://doi.org/10.1016/s0377-2217(01)00293-4

96. Stanković, J. J., Marjanović, I., & Stojković, N. (2021). DEA assessment of Socio-Economic Development of European countries. *Management:Journal of Sustainable Business and Management Solutions in Emerging Economies*. https://doi.org/10.7595/management.fon.2020.0012

97. Shah, W. U., Hao, G., Yan, H., Yasmeen, R., & Lu, Y. (2022). Energy Efficiency Evaluation, changing trends and determinants of energy productivity growth across South Asian countries: SBM-DEA and Malmquist approach. *Environmental Science and Pollution Research*, *30*(8), 19890–19906. https://doi.org/10.1007/s11356-022-23484-w

98. Shi, G.-M., Bi, J., & Wang, J.-N. (2010). Chinese regional industrial energy efficiency evaluation based on a DEA model of fixing non-energy inputs. *Energy Policy*, *38*(10), 6172–6179. https://doi.org/10.1016/j.enpol.2010.06.003

99. Sun, B., Wang, H., Ortiz, J., Huang, J., Zhao, C., & Wang, Z. (2022). A decomposed data analysis approach to Assessing City Sustainable Development Performance: A network DEA model with a slack-based measure. *Sustainability*, *14*(17), 11037. https://doi.org/10.3390/su141711037

100. Sun, Y., & Wang, N. (2021). Eco-efficiency in China's Loess Plateau region and its influencing factors: A data envelopment analysis from both static and dynamic perspectives. *Environmental Science and Pollution Research*, *29*(1), 483–497. https://doi.org/10.1007/s11356-021-15278-3

101. Thanassoulis, E. (2001). *Introduction to the theory and application of data envelopment analysis*. Springer. https://doi.org/10.1007/978-1-4615-1407-7

102. Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, *130*(3), 498–509. https://doi.org/10.1016/s0377-2217(99)00407-5

103. Tone, K. (2010). Variations on the theme of slacks-based measure of efficiency in DEA. *European Journal of Operational Research*, *200*(3), 901–907. https://doi.org/10.1016/j.ejor.2009.01.027

104. Tone, K. (2017). On the consistency of slacks-based measure-max model and super-slacks-based measure model. *Universal Journal of Management*, *5*(3), 160–165. https://doi.org/10.13189/ujm.2017.050307

105. Tone, K., & Tsutsui, M. (2017). The dynamic network DEA model. *Advances in DEA Theory and Applications*, 74–84. https://doi.org/10.1002/9781118946688.ch9

106. Trianni, A., Cagno, E., & Neri, A. (2017). Modelling barriers to the adoption of industrial sustainability measures. *Journal of Cleaner Production*, *168*, 1482–1504. https://doi.org/10.1016/j.jclepro.2017.07.244

107. Tsaples, G., & Papathanasiou, J. (2021). Data envelopment analysis and the concept of Sustainability: A review and analysis of the literature. *Renewable and Sustainable Energy Reviews*, *138*, 110664. https://doi.org/10.1016/j.rser.2020.110664

108. Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An exploratory DEA and Machine Learning Framework for the evaluation and analysis of sustainability composite indicators in the EU. Mathematics, 10(13), 2277. https://doi.org/10.3390/math10132277

109. United Nations. (1972). United Nations Conference on the Human Environment: Stockholm, Sweden, June 5-16, 1972.

110. United Nations. (1992). United Nations Conference on Environment and Development, Rio de Janeiro, June 3-14, 1992

111. United Nations. (2015). Transforming our World: The 2030 Agenda for Sustainable Development. https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

112. Wang, R., Wang, Q., & Yao, S. (2021). Evaluation and difference analysis of regional energy efficiency in China under the carbon neutrality

targets: Insights from DEA and Theil models. *Journal of Environmental Management*, *293*, 112958. https://doi.org/10.1016/j.jenvman.2021.112958

113. Wang, C.-N., Nguyen, H.-P., & Chang, C.-W. (2021). Environmental efficiency evaluation in the top Asian economies: An application of DEA. *Mathematics*, *9*(8), 889. https://doi.org/10.3390/math9080889

114. Warhurst, A. (2002). Sustainability Indicators and Sustainability Performance Management. https://www.iied.org/sites/default/files/pdfs/migrate/G01026.pdf

115. WCED. (1987). *Report of the world commission on environment and development: Our common future*. Oxford University Press.

116. Wojcik, V., Dyckhoff, H., & Gutgesell, S. (2017). The desirable input of undesirable factors in data envelopment analysis. *Annals of Operations Research*, *259*(1–2), 461–484. https://doi.org/10.1007/s10479-017-2523-2

117. World Business Council for Sustainable Development . (1992). Eco-efficiency learning module. https://www.wbcsd.org/Projects/Education/Resources/Eco-efficiency-Learning-Module

118. Wu, J. (2013). Landscape Sustainability Science: Ecosystem Services and human well-being in changing landscapes. *Landscape Ecology*, *28*(6), 999–1023. https://doi.org/10.1007/s10980-013-9894-9

119. Wu, J., & Wu, T. (2012). Sustainability Indicators and indices: An overview. *Handbook of Sustainability Management*, 65–86. https://doi.org/10.1142/9789814354820_0004

120. Yousefi, S., Hassanzadeh, A., Saen, R. F., & Mousavi Kashi, Z. (2021). Assessing sustainability of Islamic countries via data envelopment analysis (DEA). *Clean Technologies and Environmental Policy*, *24*(4), 1129–1143. https://doi.org/10.1007/s10098-020-02002-x

121. Zakari, A., Khan, I., Tan, D., Alvarado, R., & Dagar, V. (2022). Energy Efficiency and Sustainable Development Goals (sdgs). *Energy*, *239*, 122365. https://doi.org/10.1016/j.energy.2021.122365

122. Zhang, B., Bi, J., Fan, Z., Yuan, Z., & Ge, J. (2008). Eco-efficiency analysis of industrial system in China: A Data Envelopment Analysis Approach. *Ecological Economics*, *68*(1–2), 306–316. https://doi.org/10.1016/j.ecolecon.2008.03.009

123. Zhang, L., Zhao, L., & Zha, Y. (2021). Efficiency evaluation of Chinese regional industrial systems using a dynamic two-stage DEA approach. *Socio-Economic Planning Sciences*, *77*, 101031. https://doi.org/10.1016/j.seps.2021.101031

124. Zhou, G., Chung, W., & Zhang, Y. (2014). Measuring Energy Efficiency Performance of China's transport sector: A Data Envelopment Analysis Approach. *Expert Systems with Applications*, *41*(2), 709–722. https://doi.org/10.1016/j.eswa.2013.07.095

125. Zhou, H., Yang, Y., Chen, Y., & Zhu, J. (2018). Data Envelopment Analysis Application in sustainability: The origins, development and Future Directions. *European Journal of Operational Research*, *264*(1), 1–16. https://doi.org/10.1016/j.ejor.2017.06.023

126. Zhou, S., & Xu, Z. (2022). Energy Efficiency Assessment of RCEP member states: A three-stage slack based measurement DEA with undesirable outputs. *Energy*, *253*, 124170. https://doi.org/10.1016/j.energy.2022.124170

127. Zuo, Z., Guo, H., Li, Y., & Cheng, J. (2022). A two-stage DEA evaluation of Chinese mining industry technological innovation efficiency and eco-efficiency. *Environmental Impact Assessment Review*, *94*, 106762. https://doi.org/10.1016/j.eiar.2022.106762

128. Łącka, I., & Brzezicki, Ł. (2022). Joint Analysis of National ECO-EFFICIENCY, eco-innovation and SDGS in Europe: Dea Approach. *Technological and Economic Development of Economy*, *28*(6), 1739–1767. https://doi.org/10.3846/tede.2022.17702

129. Tsaples, G. (2022). A novel Multi-level Data Envelopment Analysis variation for the construction of composite indicators and the measurement of sustainability "PhD thesis". University of Macedonia, Thessaloniki.