ONLINE FIRST ARTICLE
doi.org/10.26398/IJAS.239

Statistica Applicata - Italian Journal of Applied Statistics

Received: 20.02.23/ **Accepted:** 29.11.24
Published online: 06.12.24

MEASURING GENDER EQUALITY IN THE EUROPEAN UNION: SCRUTINIZING THE GENDER EQUALITY INDEX

Milica Maričić, Department of Operational Research and Statistics, University of Belgrade – Faculty of Organizational Sciences, Serbia. ORCID: 0000-0003-0441-9899. milica.maricic@fon.bg.ac.rs. Corresponding author.

Veljko Jeremić, Department of Operational Research and Statistics, University of Belgrade – Faculty of Organizational Sciences, Serbia. ORCID: 0000-0002-0266-5247.

***Abstract**. Over the last few decades, gender (in)equality has become a topic of high interest because it has possible implications for the global economy and the overall level of sustainability. With this in mind, the policymakers needed an accurate, reliable, and multidimensional measure of the level of gender (in)equality. Indicator-based measurement approaches deemed as an obvious solution. Among several composite indicators in the field of gender equality, the Gender Equality Index (GEI), devised by the European Institute for Gender Equality, attracts attention for its methodology, structure, and the number of yearly publications. Considering the possible consequences of GEI results on the EU level, our study aims to analyze the methodological choices of this composite measure, precisely to review its current indicator list and weights assigned to them. To conduct the GEI analysis, we applied a statistical I-distance method, i.e. a methodology that can overcome the issue of subjective weight assignment. The performed twofold I-distance approach gave us an insight into domains and total score dynamics, while the applied Composite I-distance Indicator (CIDI) methodology proposed corrections of domain weights. Finally, through the iterative exclusion of indicators by the level of their relevance, using the post hoc I-distance, we provide an in-depth analysis of the countries' rank consistency depending on the number of remaining framework indicators. The obtained results indicate that the expert-driven weights assigned to domains are supported by the data and are unbiased, but that there is place for reducing the number of framework indicators.*

***Keywords**: Gender Equality Index, European Union, ranking of countries, I-distance method, CIDI methodology*

1. INTRODUCTION

The interest and attention given to the concept of gender equality have increased over the last few decades because of its relationship with economic growth at the macro-level and micro-level (Kabeer and Natali, 2013; Girón and Kazemikhasragh, 2021). One of the first academics to point out this relation and its possible contribution is Klasen (1999). He stated that gender equality in human resource management could impact overall economic growth through *optimal use of human resources* and *family relations*. The perspective of *optimal use of human resources* suggests that gender equality will raise the productivity of human capital. On the other side, *the family relations* perspective anticipates that the next generation's productivity will increase as the positive work experience will be transmitted from mother to children. The precondition for both perspectives to have an effect is education. Growth-related impacts are the results of investments in women's education. In addition, evidence from earlier research indicates that investment in women's social capital has a higher ROI factor than in men regarding non-market return (Leeves and Herbert, 2014). Some investment effects, such as a rise in employment rates and earnings, are visible in the short term. Nevertheless, cumulative and large effects are to be felt nationwide only in the long term, i.e. in 25 years or more (Appiah and McMahon, 2002).

Another relevant aspect of gender equality, which is not visible at first sight, is its relationship with sustainability. The recognition that these two concepts are intertwined has increased in recent decades (UN, 2014a), leading to a better understanding of each. There are several reasons for this relation to appear. First, a sustainable future should be built on moral and ethical standards, which implies gender equality. Secondly, the contribution of women's knowledge can have the crucial potential to help society create a more sustainable environment and economy (Cela et al., 2013; Birindelli et al., 2019).

The acknowledgement of the gender equality concept led to a steady rise in the importance of gender (in)equality measurements (Permanyer, 2010; Dilli, Carmichael and Rijpma, 2019). The history of gender (in)equality metrics began with Kersti Yllo. In 1984, she published the article "The Status of Women, Marital Equality, and Violence against Wives" in which she presented her composite index for measuring gender inequality (Bericat, 2012). However, the importance of her ideas on gender equality measurements was recognized more than ten years later at the 1995 United Nations World Conference on

Women (Amici and Stefani, 2013). Since then, gender equality indices have proliferated on a worldwide level. Today, measuring gender (in)equality is a topic of high importance that has relevant policy applications and implications. However, such measures have not received the attention from the academic community and literature they deserve (Permanyer, 2010; Amin and Sabermahani, 2017).

Gender equality rankings can be used to stimulate countries to focus their attention on gender inequality problems and to introduce policies aimed at its reduction (Permanyer, 2010). For this reason, it is crucial to provide a transparent and unbiased ranking. Accordingly, this study will try to address several related topics. First, we will discuss several gender (in)equality indices and their frameworks to observe some of their imperfections. Secondly, we will review the methodological choices of the *Gender Equality Index* (GEI) (EIGE, 2022), a multidimensional composite index of gender equality created to provide a measure across 27 EU member states.

The United Nations Development Program (UNDP) was the first to develop a composite index on gender equality. In 1995, the UNDP published two indices with the idea of capturing gender disparities at the world level: the *Gender Development Index* (GDI) and the *Gender Empowerment Measure* (GEM) (Amici and Stefani, 2013). The GDI is often called Gender-related HDI as it is computed by calculating the score of each HDI dimension of the index for the two genders separately. Countries are later ranked based on the absolute deviation from the gender parity in the HDI (UNDP, 2023). On the other hand, GEM is a complement to the GDI. It encompasses what GDI does not - women's participation in political and economic life (Amici and Stefani, 2013). In 2000, for the 20th anniversary of Human Development Report, the UNDP presented a new measure: *Gender Inequality Index* (GII). The GII is another composite index with a goal to measure women's disadvantage in three dimensions - empowerment, economic activity and reproductive health (Permanyer, 2013). The World Economic Forum created the most recent global composite index in 2006: *the Global Gender Gap Index.* This index aims to capture the scale of gender-based disparities by tracking the country's progress in equal economic participation, educational attainment, political empowerment, and health and survival (WEF, 2022).

All these indices still have a highly complicated measuring system with conceptual and methodological flaws such as subjective weighting process and/or questionable framework indicators (Klasen, 2006; Permanyer, 2013; Elias, 2013). Learning on their limitations, Plantenga and associates (2009) created the *European Union Gender Equality Index* (EUGEI). They wanted to emphasize the importance of measuring and determining gender equality across the EU countries. The substantial improvement this index brought to further gender equality measurement was the inclusion of unpaid time. Taking time into account was an important advancement as equal distribution of unpaid work is a precondition for an equal distribution of paid work (Plantenga et al., 2009).

Many of the indicators mentioned above were created to measure various aspects of gender (in)equality at the global level. Accordingly, they have fallen short of providing the kind of measures that would start a debate and contribute to decisions made at the EU and member-state levels. In addition, the EUGEI also did not have the expected impact on EU policymakers. To answer the lack of effective quantitative measurement of gender equality, the European Institute for Gender Equality (EIGE) created the Gender Equality Index (GEI). GEI is a composite indicator with a three-level structure (indicators – sub-domains – domains) aimed at ranking EU member states based on the achieved level of gender equality.

Considering the importance of the GEI, its results, and the policy implications it might have, it is important to evaluate this composite indicator's structure and methodological choices. The process of composite index creation and development encompasses ten steps as defined in the OECD's handbook (OECD, 2005): Theoretical framework, Data selection, Imputation of missing data, Multivariate analysis, Normalisation, Weighting and aggregation, Uncertainty and sensitivity analysis, Back to the data, Links to other indices, and Visualization of the results. In the presented study, we focused on exploring the GEI methodological choices for the 6th step: weighting and aggregation. We question if the chosen weighting scheme is appropriate and whether the index structure could be simplified by assigning some indicators the zero-weight. Among many statistical methods and methods of operational research available, we decided to use a statistical, multivariate, data-driven, distance-based analysis, the I-distance method (Ivanovic, 1977; Maričić et al., 2019). This method has been extensively used in composite indicator creation and evaluation (for example Jeremic et al., 2011; Maricic and Kostić-Stanković, 2014). I-distance stands out among many methods because it allows for the ranking of entities, without the need of the decision-maker to provide inputs on weights. Namely, the method can, based on the data, suggest data-driven weights. What is also convenient is the fact that the method can be used as post hoc analysis for exploring the composite indicator structure.

In this study, we applied the I-distance method to the GEI to evaluate its weighting scheme and structure. The analysis was threefold: first, we performed the two-fold I-distance method to aggregate the sub-domain values to domains and domains to overall I-distance values; second, we proposed new domain weights; and finally, we performed the post hoc I-distance to explore the GEI structure.

The paper is conceptualized as follows. The next section features the GEI and its structure. Next, we give an overview of the statistical multivariate method used to perform the analysis – the I-distance method and related analyses. In the section that comes after, we present and highlight the results obtained after scrutinizing the official data set for 27 member states for the year 2022. In the last two sections, we provide discussion and concluding remarks.

2. GENDER EQUALITY INDEX (GEI)

Equality between women and men is one of the EU's fundamental values incorporated in its Treaties, including the Charter of Fundamental Rights of the European Union (EUR-Lex, 2012). The European Commission concluded that the EU needed a composite index that would measure the level of gender equality in its member states with high precision. Accordingly, the first task put in front of the newly formed and long-awaited European Institute for Gender Equality (EIGE) was the creation of a gender equality index. After three years of devoted work, the Gender Equality Index (GEI) was created in 2013.

Through multiple dimensions, the GEI aims to picture how close the EU and member states have come towards achieving gender equality. Besides its complex assignment, the GEI is easy to understand and to communicate its idea as a mean of gender equality promotion. The GEI is also able to measure the progress of each member state over time (EIGE, 2022).

The index comprises six domains, which make the core index and an additional satellite domain. The satellite domain *Violence* is not included in the core framework as it focuses statistically on violence against women (EIGE, 2022). The EIGE stated that this domain is expected to be a part of the GEI from the edition 2024 after a comprehensive EU Gender-based violence survey (EU-GBV) is completed. Table 1 shows 31 indicators that make the GEI framework, divided into six domains and 14 sub-domains. The raw indicator data has been collected from the Eurostat, Gender Statistics Database, and EIGE.

Tab. 1: Domains, sub-domains, indicators and their codes used for determining countries' level of gender equality

Domains	Sub-domains	Indicators
Work (A)	Participation in work (A1)	FTE employment (A1.1)
		Duration of working life (A1.2)
	Segregation and quality of work (A2)	Sectoral segregation (A2.1)
		Flexibility of working time (A2.2)
		Career prospects index (A2.3)
Money (B)	Financial resources (B1)	Earnings (B1.1)
		Income (B1.2)
	Economic situation (B2)	Not at-risk-of-poverty (B2.1)
		Income distribution (B2.2)
Knowledge (C)	Attainment and segregation (C1)	Tertiary (C1.1) education
		People employed in education, human health and social work activities (C1.2)
	Segregation (C2)	Tertiary students in the fields of education, health and welfare, humanities and arts (C2)
Time (D)	Care activities (D1)	Childcare activities (D1.1)
		Domestic activities (D1.2)
	Social activities (D2)	Sport, culture and leisure activities (D2.1)
		Volunteering and charitable activities (D2.2)
Power (E)	Political (E1)	Ministerial representation (E1.1)
		Parliamentary representation (E1.2)
		Regional assemblies representation (E1.3)
	Economic (E2)	Members of boards (E2.1)
		Members of Central Bank (E2.2)
	Social (E3)	Share of board members of research funding organizations (E3.1)
		Share of board members in publicly owned broadcasting organizations (E3.2)
		Share of members of highest decision-making body of the national Olympic sport organizations (E3.3)
Health (F)	Status (F1)	Self-perceived health (F1.1)

		Life expectancy (F1.2)
		Healthy life years (F1.3)
	Behaviour (F2)	Non-smoking and non-drinking (F2.1)
		Doing physical activities (F2.2)
	Access (F3)	Unmet medical needs (F3.1)
		Unmet dental needs (F3.2)

Source: EIGE (2022)

EIGE states that women's and men’s participation in paid work is key to paving the way for further progress in gender equality (EIGE, 2022). Therefore, the first domain *Work* aims to measure women’s *Participation* (A1) in the labour market and *Segregation and quality of work* (A2) that women encounter. The aspect of participation is measured through FTE employment and duration of working life, while the segregation and quality of work have been measured with three indicators. One of them is sectoral segregation, i.e. the proneness of men and women to work in different occupational fields. Sectoral segregation has been widely marked as a source of the gender pay gap (Bergmann et al., 2019).

A society striving to achieve gender equality should be based on the principle of both men and women being paid equally for their work (Plantenga et al., 2013). At the same time, personal earnings allow women to be financially independent, automatically granting them equal rights (Hendriks, 2019). For this reason, financial indicators were included in the GEI framework. *Money* domain covers the difference in earnings (*Financial resources,* B1*)* and earning allocation (*Economic situation,* B2*).*

The third domain, *Knowledge,* examines the gaps between women and men regarding educational attainment and training. Today, women tend to reach or even exceed men’s educational attainment. Such changes oppose the traditional gender ideology seen by older generations (EIGE, 2022). This is why, even today, access to education for women is difficult in some more traditional societies. Without proper education, a woman’s probability of getting into the labour market decreases, leading her to economic dependency on men, especially during COVID and post-COVID pandemic (Reichelt et al., 2021). In a way, this puts *Knowledge* at the core of gender inequality.

Including *Time* in a framework for measuring gender equality is a major step forward. This domain tries to break the typical division of activities into productive and reproductive ones, which is the core of gender inequality

(Crompton, 2006). Although women are more present in the labour market than in the past decades, the responsibility and burden of managing the household is still on them (Aassve et al., 2014). The idea of equal sharing of time refers to the time men and women spend on household and family activities (*Care activities,* D1*)* and leisure and community work (*Social Activities,* D2).

Domain *Power* focuses on the gap between women's and men's level of representation in the decision-making positions in the political, economic, and social spheres (EIGE, 2022). This domain complies with previous gender inequality indices such as GEM. In addition, it measures the EU's development goal of achieving a balanced participation of men and women in decision-making processes (Plantenga et al., 2009). Despite the increase in female representatives in decision-making bodies, the struggle for equity in this sphere exists (Celis and Lovenduski, 2018). The *Power* domain is seen through *Political* (E1)*, Economic* (E2)*,* and *Social* (E3) representation in public and private institutions.

The last domain, *Health,* measures the impact of gender on one's health. All individuals should have the same access to public and private goods and services (EIGE, 2022), but access to women can be difficult. Weaker labour market attachment, lower socio-economic position, and lesser participation in the public sphere are reasons for such occurrences (Heise et al., 2019). Besides the health *Status* (F1)*,* which measures self-perceived health and life expectancy, this domain measures *Access* (F3) to basic medical and dental needs, as well as healthy life *Behaviour* (F2).

Two important aspects of the GEI should be elaborated upon: the weights assigned and the aggregation method. Regarding weights assigned to framework indicators, sub-domains, and domains, they depend on the framework level. Namely, weights are equal (indicator and sub-domain level) or based on experts' opinions (domain level). On the other hand, the aggregation method is arithmetic (indicators to sub-domains) or geometric mean (sub-domains to domains and domains to overall GEI). Namely, indicators are aggregated into sub-domains by arithmetic mean and equal weights, while sub-domains are aggregated into domains by geometric mean and equal weights, and finally, the overall result of GEI is computed as the geometric mean and experts' weights obtained by Analytic Hierarchy Process (AHP) (Papadimitriou et al., 2020). Table 2 shows the GEI aggregation method and weights assigned to the indicators, sub-domains, and domains.

Tab. 2: Aggregation method and the weights assigned to the indicators, sub-domains, and domains

	Indicator level	Sub-domain level	Domain level
Weighting	Equal	Equal	AHP
Aggregation	Arithmetic	Geometric	Geometric

Source: Papadimitriou et al. (2020)

There are no specific information in any official EIGE report on the choice of the aggregation methods used and weighting schemes chosen. The 2017 report states that 3,636 formulas were considered and therefore, 3,636 indices were computed (EIGE, 2017). They mention considering four methods for assigning weights (equal weights, a modified version of equal weights, weights retrieved from statistical analysis, and finally, weights derived from expert opinions) and two aggregation methods (arithmetic and geometric). However, for the purposes of index scrutinization, it would have been better if more information was provided by the index creators.

The performance ranking of member states based on their level of gender equality was first announced biannually but is now announced annually. Our research is based on the data retrieved from the official GEI 2022 data set, depicting the situation in the year 2020. Within the research, we have applied a twofold I-distance approach, defined and calculated the *Composite I-distance Indicator* (CIDI) methodology and performed post hoc analysis. Our analysis is believed to provide additional confirmation that the GEI is a stable, coherent, and reliable metric or will point out future directions of GEI alterations.

3. METHODOLOGY

In this section, we will outline the fundamentals of the I-distance method (Section 3.1), the related Composite I-distance indicator (CIDI) methodology (Section 3.2), as well as the post hoc I-distance approach (Section 3.3).

3.1 I-DISTANCE METHOD

Composite indices have widely been criticized for their subjectivity in the indicators' selection process, weighting system and later aggregation method (Booysen, 2002; Greco et al., 2019). The weighting of indicators plays a major role in the development of a composite index, and as such, it raises uncertainty and debate along the process (Tarantola and Saltelli, 2007; Becker et al., 2017). For that, additional attention should be given to this process when creating a composite index. Some weights can be based on statistical methods, while others might depend on expert opinion to denote the policy priorities better and/or theoretical factors (OECD, 2005). Such as weighting methods, the aggregation methods also vary. Namely, the linear method is preferable when indicators have a measurement unit, while the geometric is more appropriate when no compensability between indicators should be allowed (Munda, 2008). Whichever weighting or aggregation method is used, an adequate and unbiased one cannot be easily resolved (Saisana et al., 2005).

In the 1960's a need emerged for a composite index that will rank countries based on their socio-economic development. A new index should have been created which would be able to use various indicators, to maximize the amount of information gathered, and most importantly, it had to be unbiased. Ivanovic (1977) devised a statistical method, the I-distance method, that was capable of answering all the requirements. Namely, the I-distance method is based on calculating the mutual distances between the entities being processed, whereupon they are compared to one another to create a rank (Jeremic et al., 2011). In order to calculate the distance and rank countries, it is necessary to fix one entity as a reference in the observed set using the I-distance method. The ranking of entities in the set is based on the calculated distance from the referent entity (Maricic et al., 2019).

The referent entity can be an actual or a fictive entity and the choice on that is made by the analyst. The referent entity can be a particular entity (for example Italy) or an entity which has the minimal measured values of each indicator or an entity which has the maximum measured values or even an entity which has values predefined by the analyst. Based on the choice of the referent entity, the obtained values are interpreted. If the referent entity is a particular benchmark country, the values obtained provide information on whether other entities perform better or worse than it. In practice, so far, most commonly, the referent entity was the entity which had the minimal value (Jeremic et al., 2011). The values of the obtained I-distance then provide information on how far away an entity is from the worst-case scenario. In the performed analysis, the authors used as the referent a fictive entity with minimal measured values of each indicator.

For a selected number of variables (indicators), denoted with k, $X^T = (X_1, X_2, \ldots X_k)$, chosen to characterize the entities, the I-distance between the entity $e_r = (x_{1r}, x_{2r}, \ldots x_{kr})$ and the fictive entity $e_s = (x_{1s}, x_{2s}, \ldots x_{ks})$ is defined as:

$$D(r,s) = \sum_{i=1}^{k} \frac{|d_i(r,s)|}{\sigma_i} \prod_{j=1}^{i-1} \left(1 - r_{ji.12\ldots j-1}\right) \tag{1}$$

where $e_r = (x_{1r}, x_{2r}, \ldots x_{kr})$ and $e_f = (x_{1f}, x_{2f}, \ldots, x_{kf})$ are values of indicators I, $I = \{1, \ldots, k\}, i \in I$ of the observed entity e_r and fictive entity e_f;

$d_i(r,s)$ is the distance between the values of the indicator X_i for entities e_r and e_s e.g. the discriminate effect:

$$d_i(r,s) = x_{ir} - x_{is} \quad i \in \{1, \ldots k\} \tag{2}$$

σ_i is the standard deviation of indicator i, $i \in I$ and

$r_{ji.12\ldots j-1}$ is a partial coefficient of the correlation between indicators i and j where $j<i$, $i \in I$, $j \in I$, while the effects of all other indicators 1,2,…, j-1 are eliminated (Jeremic et al., 2011; Maricic and Jeremic, 2023). Partial coefficient of correlation describes the strength of a linear relationship between two variables, holding constant a number of other variables. The partial coefficient of correlation eliminated the effect of the other confounding variable(s) that is numerically related to both variables of interest (Baba et al., 2004).

The calculation of the I-distance is an iterative process, consisting of several steps. First, the value of the discriminate effect of the first variable (the most valuable variable, which provides the largest amount of information on the phenomena upon which the entities will be ranked) is calculated. Then, the value of the discriminate effect of the second variable that is not covered by the first one is calculated. This procedure is repeated for the all observed variables in the data set (Radojicic et al., 2019).

To overcome the problem of negative coefficient of partial correlation, which can occur when it is not possible to achieve the same direction of variables, it is suitable to use the square I-distance (Maricic and Kostic-Stankovic, 2016; Maricic and Jeremic, 2023). It is given as:

$$D^2(r,s)=\sum_{i=1}^{k}\frac{d_i^2(r,s)}{\sigma_i^2}\prod_{j=1}^{i-1}\left(1-r_{ji.12\ldots j-1}^2\right) \tag{3}$$

Instead of using the standard deviation of indicator i (σ_i) and partial coefficient of the correlation ($r_{ji.12\ldots j-1}$) between indicators i and j, the square I-distance uses the variance of indicator i (σ_i^2) and coefficient of partial determination ($r_{ji.12\ldots i-1}^2$) between indicators i and j, where $j<i$.

Square I-distance can be used even if the sign of the coefficients of correlation is positive. Also, when there is a large number of variables used the application of the square I-distance is recommended. The order of variables by which they are entered in the I-distance is of high importance. The first entered variable is the variable which is the most correlated with the rest (Jeremic et al., 2011). It is expected that this variable has the largest explanatory effect. The other variables are entered in the algorithm following the same procedure. When there is a large number of indicators, it can happen that the information carried by indicators which enter last in the algorithm is consumed by the indicators which entered the algorithm prior. Therefore, applying the square I-distance method is advisable as it can minimize the amount of lost information by using the partial coefficient of determination. As the presented framework has 14 sub-domains, the square method was used.

3.2 COMPOSITE I-DISTANCE INDICATOR METHODOLOGY

Besides providing rankings of entities, the I-distance method can create a more stable ranking methodology by modifying its official weights. The process of assigning adequate weights is referred to as the *Composite I-distance Indicator* (CIDI) methodology. In order to obtain weights which are not subjectively assigned, first, the correlation coefficients of each entity or domain with the I-distance value are calculated. Correlations are used as I-distance provides information on how valuable each domain is (Jeremic et al., 2011). The next step in the proposed methodology is calculating the new weights for each compounding domain based on the appropriate correlations. Weights are formed by dividing the values of correlations by the sum of correlations. The final sum of weights equals 1, thus forming a novel appropriate weighting system. The equation for determining weights is:

$$w_i = r_i / \sum_{j=1}^{k} r_j \tag{4}$$

where $r_i (i = 1,...,k)$ is the Pearson correlation coefficient of the i-th input indicator with the I-distance values (Dobrota et al., 2015).

3.3 POST HOC I-DISTANCE APPROACH

Besides providing a ranking list of entities, I-distance can be used for an in-depth analysis of the rank consistency. Namely, it can act as a post hoc approach. The post hoc approach is conducted in the following way. Using all *k* initially chosen indicators, the I-distance is calculated and the importance of the indicators for the ranking process is observed by calculating the Pearson's correlation coefficient between the indicator values and the obtained I-distance value. After each iteration, an indicator whose correlation coefficient with the I-distance value is the lowest is excluded from further analysis in the second iteration (Markovic et al., 2015; Savic et al., 2016). So, after each iteration, the number of indicators used to rank the entities is reduced. In the next iteration, the new I-distance rank is formed, and the procedure is repeated.

The I-distance post hoc approach is an iterative process, so the question that arises is when to stop excluding indicators from the framework. In the study by Markovic et al. (2015), further iterations were stopped when the sum of correlation coefficients started to plummet. However, another case can also appear when the sum of correlation coefficients increases throughout the iterative process. In that specific situation, the procedure stops when two indicators are left.

Using the post hoc I-distance method, it is not only possible to obtain information on the importance of indicators for the ranking process but also to get an insight on how the ranking of entities is sensitive to indicator exclusion. Ranking sensitivity of entities provides additional information on the contribution and consequences of including or excluding an individual indicator. Therefore, the post hoc I-distance method can be used to assess the composite indicator structure and suggest its simplification.

3.4 COMPARISON OF THE I-DISTANCE APPROACH WITH OTHER APPROACHES

This research focuses on the data-driven methodologies and one particular method, the I-distance method. Nevertheless, one can ask why the I-distance was chosen in this paper among many other available methods? Therefore, in the following paragraphs, we compare the I-distance method to several other data-driven and non-participatory methods used in the composite indicator literature.

Pena's Method

The Pena's method was devised in the same period as the I-distance method in the 1970s by Peña (1977). Interestingly, the first usage was as well in the field of quality of life composite indicators. The P2 distance or DP2 method functions similarly to the I-distance method: it calculates the distance an entity in relation to an object. This method is said to solve several issues, such as: aggregation of variables expressed in different measures, arbitrary weights and duplicity of information (Somarriba and Peña 2009). The formula for the Pena distance for a chosen entity *s* is:

$$DP2_s = \sum_{i=1}^{k} \left(\frac{\left| x_{is} - x_{ir} \right|}{\sigma_i^2} \left(1 - r_{i,i-1,\dots,1}^2 \right) \right) \tag{5}$$

The initial part of the Pena distance is the same as the initial part of the regular I-distance method – calculating the discriminant effect and taking the variability of the indicator into account. The main difference occurs in the weighting part. Pena distance considers partial coefficients of determination, while the I-distance method considers partial coefficients of correlation and coefficients of correlation. Both methods depend on the order of variables in the algorithm (Montero et al., 2010; Maricic et al., 2016).

Data Envelopment Analysis (DEA) and DEA-Like Approaches

DEA is an optimization method devised by Charnes et al. (1978), used to calculate the relative efficiency of decision-making units (DMUs) based on the measured values of inputs and outputs. The DEA method looks for data-driven weights which will maximize the overall score of the DMUs. Because the importance of the inputs and outputs does not depend on the analyst's or the experts' opinion, the application of the DEA method quickly increased, especially in policy-related settings (Cherchye et al., 2008). In the field of composite indicator creation, a special type of DEA method has been widely employed: the Benefit-of-the-Doubt (BoD) model. The BoD model is, in fact, an input-oriented DEA model (Melyn and Moesen, 1991). The goal function of the model is to maximize the value of the composite indicator by changing the weights assigned to individual indicators. The main issue with both DEA and BoD models is full freedom (Rogge and Van Nijverseel, 2019). Namely, if no weight constraints are imposed, all entities will achieve the maximum value of the composite indicator. Therefore, different approaches to weight restriction have been proposed: The upper and lower bounds of weight were generated via participatory methods, Intervals around the government-defined weights, symmetric interval ± 25% around CIDI weights, and others (Maricic and Jeremic, 2023). Although this also is a data-driven weighting approach, it significantly differs from the I-distance method. DEA and BoD models are optimization models, while the I-distance is a distance-based mathod.

Displaced Ideal Method (DI)

Displaced Ideal Method (DI) is a method based on the Euclidean distance proposed by (Zelany, 1974). The idea of the DI is to show the smallest distance of an entity from its ideal scenario. The Euclidean distance, as a distance metric, is not robust over a range of scales, which means that the computed results can be skewed if the units of the variables used have very different variabilities (Saranya and Manikandan, 2013). Similar as the I-distance method, the DI method posits that there is an inherent connection among the representative variables of a phenomenon being studied. On the other hand, the difference occurs when we observe from which the distance is observed. In the I-distance, the distance is observed between the worst case scenario, while in the ID method, iIt suggests that the optimal system should strive to minimize the gap between its current state and the ideal scenario (Magalhães-Timotio et al., 2022).

Partial Least Squares-Path Modeling

Structural equation modelling (SEM) is a statistical multivariate analysis which lies on the principles of factor analysis and regression analysis (Kline, 2005). Therefore, the analysis allows for grouping of individual indicators and exploration the relationship between the newly formed latent variables. Both these features are quite valuable in the process of composite indicator creation as they allow for considering the role (formative and reflective) of the manifest variables (MVs) (Lauro et al., 2018). Two main approaches within the SEM literature are the covariance-based (CB-SEM) and partial least squares (PLS-SEM). The first is seen as the parametic SEM, while the second is observed as the non-parametric SEM. Within the composite indicator literature, the PSL-SEM approach is more common due to the fact that the goal of a composite indicator is to estimate the latent variables, and PLS-SEM does just that (Trinchera et al., 2008). Using SEM algorithms to create composite indicators considers taking a model based approach which creates a multidimensional latent variable measurable directly and related to its single indicators or MVs by a reflective or formative relationship or by both (Lauro et al., 2018). Although both SEM and I-distance approaches are data-driven, the first one is model based, while the other is distance-based.

The presented lietarure review indicates that within the field of composite indicator creation and evaluation there is a plethora of statistical analysis and methods of operational research which can be employed. Neither approach is flawless (Greco et al. 2018). Therefore, it is suggested that the composite indicator creator considers several approaches in quest of determing the final composite indicator methodology. The I-distance method applied in this paper is just one of the possible solutions.

4. RESULTS

This section presents the results of scrutinizing the GEI using the I-distance method and the related CIDI and Post hoc approaches. The results are organized into three subsections for better paper flow and presentation.

4.1 APPLICATION OF THE TWO-FOLD I-DISTANCE APPROACH TO THE GEI

The first direction in our research implied calculating the Total I-distance values for the GEI. Within the aim of the study to scrutinize the GEI framework, we applied the twofold I-distance approach to all EU member states and compared their rankings to the official GEI rankings. The first step in the analysis is applying the I-distance method on the sub-domain values to create new I-distance domain values. The second step sees the application of the I-distance method on the previously obtained six I-distance domain values.

As presented, the analysis used sub-domain data for one reason: data availability. The data scores available were scores of the sub-domains, domains and overall GEI. The data for indicators is available, but separately for males and females. Although the data on the indicator level is available, as there is no clear and straightforward information on how the final scores are calculated, we decided to focus on the sub-domain data. Therefore, we assumed that on the level of aggregation from indicators to sub-domains, there should be no change in weights and that equal weighting is appropriate.

The rankings within each domain after applying the I-distance method are presented in Table 3.

Tab. 3: The rankings of countries by *Work, Money, Knowledge, Time, Power* and *Health* domains after applying the I-distance method

Member state	Work rank	Money rank	Knowledge rank	Time rank	Power rank	Health rank
Sweden	**1**	11	**1**	**1**	2	2
The Netherlands	3	10	5	3	5	3
Denmark	2	6	4	2	7	13
Belgium	14	2	2	13	6	12
Luxembourg	10	**1**	3	8	12	5
Ireland	12	7	7	7	10	**1**
Finland	5	4	11	6	4	9
Spain	18	21	6	10	3	7
France	16	12	9	11	**1**	14
Austria	8	8	12	16	14	4
Malta	9	14	8	14	19	6
Slovenia	13	3	22	9	16	16
Slovakia	21	5	10	25	24	20
Germany	15	17	25	12	8	8

Estonia	7	23	19	5	22	22
Latvia	6	26	27	4	15	27
Italy	27	18	13	17	11	10
Czech Republic	20	9	14	21	23	18
Cyprus	17	13	17	19	26	11
Lithuania	4	25	18	22	18	21
Bulgaria	19	27	15	26	9	23
Hungary	23	16	20	18	27	17
Portugal	11	19	21	23	13	24
Poland	24	15	16	20	20	25
Greece	26	22	23	27	25	15
Croatia	22	20	26	24	17	19
Romania	25	24	24	15	21	26

Source: Authors' own work

The I-distance domain results provide interesting results. Namely, two Scandinavian member states (Sweden and Denmark) are in the top 5 for three domains, which leads to the conclusion that these countries are committed to a multidimensional approach to reducing gender inequality. The domains in which the results of the "Scandinavian duo" are not the leading ones are *Money, Power* and *Health* domains. In the case of the *Money* domain, Luxembourg tops the list, followed by Belgium. Looking at the results of *Power,* France and Sweden lead the way, closely followed by Spain and Finland. On the other hand, the results of the *Health* domain pointed out Ireland as a country where both men and women have the same treatment and access to basic medical care.

Finally, the results of Romania and Croatia should be pointed out. The results of these countries are in the bottom 5 for three domains. These findings do not mean these countries are not trying to create a gender-equal society, just that there are discrepancies between them and other member states.

The final step of the two-fold I-distance approach saw the appliance of the I-distance on previously obtained domains. Table 4 shows the results of the two-fold approach: the Total I-distance value, Total I-distance ranks, and official GEI ranks.

Tab. 4: The results of the I-distance method, Total I-distance value, Total I-distance ranks and official GEI ranks of EU member states

EU Member state	Total I-distance value	Total I-distance rank	Official rank
Sweden	45.718	1	1
The Netherlands	22.157	2	3
Denmark	21.722	3	2
Belgium	19.420	4	8
Luxembourg	19.393	5	9
Ireland	19.375	6	7
Finland	17.396	7	4
Spain	14.755	8	6
France	14.300	9	5
Austria	11.298	10	10
Malta	10.300	11	13
Slovenia	8.570	12	12
Slovakia	7.387	13	24
Denmark	6.872	14	11
Spain	6.266	15	17
Latvia	6.236	16	16
Italy	6.000	17	14
Czech Republic	5.249	18	23
Cyprus	5.122	19	22
Lithuania	2.962	20	20
Bulgaria	2.697	21	18
Hungary	2.684	22	25
Portugal	2.619	23	15
Poland	2.294	24	21
Greece	1.519	25	27
Croatia	1.452	26	19
Romania	0.540	27	26

Source: Authors' own work

Consequently, Sweden and the Netherlands top the list. The obtained value of Sweden, 45.718, indicates that Sweden is furthest away in the multidimensional space from the fictive entity with minimal values of all six domains. On the other side of the ranking, Romania is quite close to the fictive entity, with a Total I-distance value of just 0.540.

After applying the two-fold I-distance method, out of 27 countries, 11 have improved their rank; five did not change their rank, and 11 dropped their rank.

The I-distance results show countries that lead the way in terms of creating a more gender-equal society, the ones that have visibly improved their attitude towards gender equality, the ones that were expected to be more gender equal, and the ones that still have a way to go when implementing this concept.

Scandinavian countries have a rich history of gender equality that originates from the last decades of the 19th century. More than a hundred years later, these countries are still leading in terms of women's rights in education, voting, and political representation. One of the reasons for this lies in the famous Nordic welfare system that stands out for its universalism and its devotion to creating a society that gives women equal rights in all spheres of life (Borchorst and Siim, 2008). In several central European countries, the gender equality legislation drastically changed after they joined the EU. The EU's legal system and the potent guidance of the EU were just what these countries needed to trace their path towards raising gender equality. This can be especially noted in the case of Hungary and the Czech Republic (Velluti, 2014).

After having applied the two-fold approach, the list of countries in the top 10 did not change. Countries which improved their rank are Luxembourg and Belgium, while Denmark, France, Spain, and Finland dropped ranks. This means some aspects of gender equality need more attention and that there is place for further legislation improvements.

Younger EU member states have quite a distinctive work and family policy history compared to the EU's. In these countries, there is a strong gender – tradition-based ideology that denotes men as breadwinners and women as housewives (Hofacker et al., 2013). These discrepancies mean there are essential political, economic, and social changes related to gender equality ahead of them (Witkowska, 2013).

4.2 ASSIGNING WEIGHTS BY APPLYING THE I-DISTANCE METHOD

The second direction of our research was to scrutinize the GEI weighting scheme on the domain level. To get an in-depth analysis of the GEI, besides applying the two-fold I-distance approach, we used the CIDI methodology. Newly obtained domain weights by CIDI are calculated by dividing the correlations of domains to the Total I-distance value and the sum of domains' correlations to the Total I-distance (Dmitrovic et al., 2016). To perform the analysis, we calculated the correlations between Total I-distance values and each of the I-distance domain values

The comparison of the GEI weights and the weights proposed by CIDI is presented in Table 5. The largest differences are with the domains *Power* and *Time. Power* is weighted at 19% according to the official GEI ranking and 15.9% according to the CIDI methodology. As most member states have established minimum quotas for female representation in *Political* and *Economic* sphere, this domain does not need such high importance (Mateos de Cabo et al., 2011). Similarly, *Knowledge* dropped importance by our method to 20.3% from 22%. On the other hand, *Time* is assigned 15% weight according to the AHP experts' opinions, while our method gives it a higher significance, 17.3%. This domain deserves more attention, as women are facing constraints on their leisure time both within and outside home (Aitchison, 2013).

Tab. 5: Differences in CIDI weights and original GEI domain weights.

Domain	Correlation Coefficient	CIDI weights	Official GEI weights	Change from official GEI weight
Knowledge	0.897	20.3%	22%	-1.7%
Work	0.788	17.8%	19%	-1.2%
Time	0.768	17.3%	15%	2.3%
Money	0.725	16.4%	15%	1.4%
Power	0.705	15.9%	19%	-3.1%
Health	0.546	12.3%	10%	2.3%

Source: Authors' own work

The results of the CIDI methodology indicate that changes to the domain could be implemented. The smallest suggested change is for the domain *Work* and its weight should be decreased for 1.2% points, followed by the domain *Money,* whose weight should be increased by 1.4% points. The greatest change suggested is for the domain *Power,* whose importance could be reduced from 19% to 15.9%. The detected differences indicate that even though the GEI weights at the domain level are expert-driven, they are close to data-driven weights and that for only one domain, substantial changes are advised.

4.3 APPLICATION OF THE POST HOC I-DISTANCE APPROACH

Following the idea of Markovic and associates (2015), we applied the post hoc I-distance approach. Accordingly, the third phase of our study refers to 13 I-distance iterations for 14 GEI sub-domains. Namely, the iterative process was stopped when the highest average coefficient of correlation was obtained, in this case, 13 iterations. In the last iteration, we stopped excluding indicators because there were only two sub-domains left – Financial resources (B1) and Attainment and segregation (C1). Further iteration would have shown Financial resources (B1) as the most important sub-domain. Table 6 shows the ranks after the first two iterations, the ranks in the last iteration, the change in rank after the first and the last iteration, the median and the interquartile range (IQR) for all EU member states after the application of the post hoc I-distance approach.

Tab. 6: I-distance iterations, change in rank between the first and the last iteration, median, and interquartile range

Country	1st iteration	2nd iteration	…	13th iteration	Change in rank	Median	IQR
Sweden	1	1	…	6	-5	1	1
The Netherlands	2	2	…	2	0	2	0
Denmark	3	3	…	3	0	3	2
Ireland	4	5	…	5	-1	5	0
Finland	5	4	…	4	1	4	2
Luxembourg	6	6	…	1	5	4	3
Belgium	7	7	…	7	0	7	2
Spain	8	8	…	11	-3	8	2
Austria	9	10	…	8	1	9	2
France	10	9	…	10	0	9	2
Malta	11	12	…	12	-1	12	2
Slovenia	12	11	…	15	-3	13	1
Germany	13	13	…	9	4	11	1
Italy	14	15	…	14	0	15	2
Cyprus	15	18	…	13	2	17	3
Czech Republic	16	17	…	18	-2	20	3
Estonia	17	14	…	16	1	14	2
Slovakia	18	16	…	26	-8	26	5
Portugal	19	20	…	21	-2	18	3
Hungary	20	22	…	22	-2	22	4
Lithuania	21	21	…	17	4	17	5
Bulgaria	22	24	…	27	-5	25	4
Latvia	23	19	…	24	-1	23	7
Croatia	24	25	…	23	1	22	2
Poland	25	23	…	19	6	24	2

Greece	26	26	…	20	6	21	6
Romania	27	27	…	25	2	27	0

Source: Authors' own work

Indicator *Change in rank* shows an interesting result: three countries held the same rank throughout the iterations, and three countries moved up for just one place. The highest oscillations, as a result of sub-domains reduction, occurred at the bottom of the ranks, to Poland and Greece. Namely, these countries improved their ranks by six places respectively. On the other hand, the largest decrease in ranks occurred in the case of Slovakia, which dropped eight ranking places. These three countries proved to be the most sensitive to excluding indicators. A detailed frequency analysis of the change in rank is presented in Table 7. As can be seen, the majority of Member states, as much as five of them, have not changed their rank after significant indicator removal. This can be an indication that the subdomain list can be reduced.

Tab. 7: Frequency analysis of the change in rank of Member states

Change in rank	Frequency	Member states
-8	1	Slovakia
-5	2	Sweden, Bulgaria
-3	2	Slovenia, Spain
-2	3	Hungary, Portugal, Czech Republic,
-1	3	Latvia, Malta, Ireland
0	5	The Netherlands, Denmark, Belgium, France, Italy
1	4	Croatia, Estonia, Austria, Finland
2	2	Romania, Cyprus
4	2	Lithuania, Germany
5	1	Luxembourg
6	2	Greece, Poland

Source: Authors' own work

It is valuable to observe the rank median and compare it with the rank in the last iteration. For example, Sweden has a median 1, while its 13th rank is six. This indicates that Sweden dropped rank almost at the end of the analysis. On the other hand, Luxembourg has a median 4, while its 13th rank is 1, which indicates that Luxembourg improved rank almost at the end of the analysis. For the majority of member states, the median and the 13th rank are the same or close, indicating mostly stable ranks throughout the analysis.

IQR can also be used to detect the member states whose ranks oscillated. For example, the IQRs of Latvia and Greece are seven and six, respectively, indicating sharp rank changes. On the other hand, the IQR of Romania is 0, showing that the ranks of Romania have been stable.

As the final part of the descriptive analysis of the measured changes in rank, we provide a scatterplot between the rank in iteration 1 and iteration 13. The countries below the regression line are the countries whose rank improved, while the countries above are those whose rank decreased.

Insert Figure 1 here

Fig. 1: Scatterplot between rank in iteration 1 and 13

Besides solely analyzing country's ranks throughout the iterations, we observed the ranks of indicators by their correlation coefficient with the Total I-distance value and the order by which they were ruled out from further observation. Table 8 presents the rank of indicators, their coefficient of correlation and the average coefficient of correlation after each iteration. As one can see, after the first iteration, the change of the average coefficient of correlation was for +0.024. With every next iteration, the quality of our model improved, as the average correlation grew from 0.640 in the first iteration to 0.891 in the last. Another interesting fact is that the importance of the sub-domains changed. In the first iterations, the sub-domain *Social activities* (D2) was top-ranked, while at the end of the analysis, the sub-domain *Financial resources* (B1) came to be the most important. The order of domains by which they were ruled out of the framework is *Power-Health-Time-Work-Knowledge-Money.*

Tab. 8: Review of excluding indicators, their coefficient of correlation and the average coefficient of correlation after each iteration

1st iteration	r	2nd iteration	r	…	12th iteration	r	13th iteration	r
D2	0.877	D2	0.884	…	B1	0.940	B1	0.941
A2	0.826	C1	0.822	…	C1	0.851	C1	0.842
C1	0.815	A2	0.803	…	A2	0.833		
B1	0.780	F2	0.760	…				
F2	0.761	B1	0.747	…				
E3	0.736	E3	0.742	…				
E1	0.716	E1	0.733	…				
F1	0.632	D1	0.674	…				
D1	0.618	E2	0.622	…				
E2	0.614	F1	0.577					
C2	0.545	C2	0.510					
B2	0.362	A1	0.399					
A1	0.348	B2	0.358					
F3	0.335							
Average r	0.640		0.664	…		0.875		0.891

In order to get a glimpse of the oscillations in I-distance ranks, Figure 2 gives a graphic overview of how the rank of the top 5 countries (according to the Total I-distance) changed during the analysis. Firstly, Sweden and the Netherlands swapped places after nine iterations. The rank of Sweden continued to decline in the next iterations, while the Netherlands remained in the top ranks. The rank of Denmark was stable until iteration 7, when it started to decrease, up to rank 6 (iteration 10). After that, it improved its rank and finished third best. The remaining two countries swapped after the first two iterations, after which Finland sharply decreased to rank 8. During the next iterations, the rank of Finland visibly changed, leading to a final improvement to rank 4. Similar movement could be detected for Ireland.

Insert Figure 2 here

Fig. 2: Overview of the oscillations in the ranks of the top five countries according to the total I-distance value

Knowing that the bottom-ranked countries are more prone to rank oscillations, Figure 3 gives a graphic overview of how the rank of the bottom 5 countries (according to the Total I-distance) changed during the analysis. Romania had a stable rank up until the 10th iteration, when it improved its rank to 25th place on which it remained. The volatility of the remaining four observed countries is greater. Greece and Latvia had visible rank increase and decrease depending on the domains in the composite index framework. Poland and Croatia also have oscillations, but they are not that drastic.

Insert Figure 3 here

Fig. 3: Overview of the oscillations in the ranks of the bottom five countries according to the total I-distance value

4.4 FUTURE DIRECTIONS OF THE STUDY AND STUDY LIMITATIONS

The three performed analyses based on the I-distance are just one direction of the GEI revision. Namely, other statistical methods could also be applied to inspect whether the GEI is statistically sound and how to enhance it.

One of the possible directions is applying the CIDI methodology to sub-domains. Namely, CIDI can be performed in order to get an evaluation of weights assigned to sub-domains and their importance for the ranking process. Further, sensitivity and uncertainty analysis can be carried out in order to get a complete evaluation of the indicators that make the GEI framework. Also, the results of the post hoc analysis can be applied to reduce the number of framework indicators, which might result in lower calculation costs and even in annual index publications.

Besides further statistical analysis, the number of countries covered is to be expanded. In addition, other countries which are not member states may build a gender equality index modelled on the GEI. Their results can be later compared with the EU average or on a regional level. Serbia, an EU candidate country, is the first to use this opportunity as it has the know-how and expertise to do so (EIGE, 2014). Other candidate countries, like Montenegro, Bosnia and

Herzegovina, and others, should also try to follow Serbia on the road for measuring gender equality and trying to create a gender-equal society.

Besides exploring the list of sub-domains and indicators of the GEI as well as the weighting schemes assigned, it would be of interest to inspect the initall structure of the GEI. For that, advanced multivariate analysis such as Hierarchical Disjoint Non-Negative Factor Analysis (Cavicchia et al., 2021), Second-Order Disjoint Exploratory Factor Analysis (Cavicchia and Sarnacchiaro, 2021), or variants of the Benefit-of-the-Doubt method (for example Verbunt and Rogge, 2018).

A limitation of the presented two-fold approach pulls along another direction of future studies. Namely, the I-distance can provide information on the relative performance of the observed countries. This statistical method calculates the distance to a referent entity rather than to a previously defined standard. This comparative analysis draws attention to the fact that there are countries that are doing better than others in terms of achieving gender equality, without elaborating the detailed reasons for such differences. The identification of the determinants of the discrepancies among the observed countries requires the assistance of sociologists, lawyers, and other specialists in the topic. Such multidisciplinary analysis could give answers to policy makers and point them out which legislative segment should be tackled in the future.

Finally, a limitation due to the data availability should be observed. As mentioned in Section 4.1, in this study, the focus was on the sub-domains and domains without deeper analysis of indicator scores and weights. The assumption made in this study was that the equal weighting from indicator scores to sub-domain scores was reasonable. Nevertheless, such a claim can be challenged, as no proof exists that all indicators within one sub-domain are equally important. Therefore, this study focused on the two levels of the GEI, and not on all three. If the indicator scores were publicly available, three-fold I-distance could have been applied, shedding light on the importance of individual indicators.

5. CONCLUDING REMARKS

Out of eight United Nations Millennium Development Goals (MDGs), which were established at the United Nations Millennium Summit in 2000, one is related to gender equality, precisely to “Promote gender equality and empower women” (UN, 2000). Despite the steady progress in the field of gender equality in access to education, labour market and politically influential positions, there is still place for further improvements (Eurostat, 2022). To continuously monitor the progress of achieving gender equality, relevant, accurate and timely gender data is of high importance. The collected data is used to calculate explicit measures of gender (in)equality. The academic community and the international institutions have acknowledged the importance of such metrics, and consequently, several gender equality indices have been created. The idea behind these indices is to draw attention of public and, more importantly, policymakers to the issue of gender-related policies and research (UN, 2012). The index which stands out is the one devised by the European Institute for Gender Equality, the Gender Equality Index (GEI).

In the previous years, the I-distance method was used with success for assessing composite indices of different purposes (Maricic and Kostic-Stankovic, 2016; Markovic et al., 2015; Maricic et al., 2019). What differentiates the I-distance method and related methods from other statistical methods is its objectiveness. Namely, the method applied in this study does not place any weighting factor on its indicators (Jeremic et al., 2014), meaning subjectively assigned weights cannot influence the final ranking. Within the aim of this research to attempt to improve the measuring system of GEI we proposed the I-distance method to be applied, together with Composite I-distance Indicator (CIDI) methodology, and post hoc analysis.

First, using the two-fold I-distance approach for assessing GEI allowed us to identify the leaders, who follows them, but who is very inefficient in that. The results revealed that Scandinavian member states top the list, that some of the founding members (France and Germany) scored below expectation, whereas South European countries are showing low efficiency in applying the gender equality concept. The obtained ranks comply with the research conducted by Earles (2014). The country ranks by the two-fold approach and the official GEI differ in the middle of the rankings, while the ranks of the top and the bottom countries have not changed. These results led to a significant level of correlation between the two ranking methodologies measured by Spearman's correlation coefficient (r_s=0.880, $p<0.01$). Both GEI and I-distance ranks point out one: there are considerable differences inside the EU regarding gender equality that are connected with history, tradition, culture, and the welfare of member states (Witkowska, 2013).

In the official framework, the ranks were obtained as weighted geometric mean of domain values, whilst weights were obtained through the analytical hierarchy process (AHP). Although AHP is widely preferred for solving multi-criteria decision-making problems, one of its drawbacks, which might easily be challenged, are the subjectively assigned evaluation measures (Kilincci and Asli Onal, 2011). By proposing the CIDI methodology for measuring member states' performance and gender equality, we attempted to improve GEI's ranking methodology by assigning objective weights to domains. The aim of this analysis was to test the domain weights obtained by AHP. The obtained result differs from the official ones. The biggest difference can be seen in weights assigned to domains *Power* and *Time.* Besides these, the CIDI methodology confirms the latter AHP weights.

The third analysis performed was the post hoc analysis, whose results imply that five member states had the same GEI and 13th iteration rank and that two of the top five countries did not change throughout the analysis. The correlation coefficient between official GEI ranks and the last iteration is significant (r_s=0.833, p<0.01), meaning that the country's ranks are stable. One should notice that the correlation coefficient between the first (r_s=0.894, p<0.01) and the last iteration (r_s=0.833, p<0.01) with the GEI differs for just 0.063. This leads to the conclusion that the number of indicators which comprise the GEI could be refined without significant changes to the member states ranks.

All of the above-mentioned findings provide an in-depth analysis of the GEI domains and weights assigned to them. We hope our study could act as a confirmation of the GEI's methodology and its results, but also as a guidance for possible future slight enhancements of this composite indicator. This research can also be an impetus for innovative research approaches in the field of gender equality evaluation on the EU level, which might eventually impact EU policy and legislation.

Declarations

Ethical Approval
Not applicable

Competing interests
Not applicable

Authors'contributions
The authors equally contributed to the analyses and the manuscript.

Funding
Not applicable

Availability of data and materials
Available on request from the corresponding author

REFERENCES

Amin, E. and Sabermahani, A. (2017). Gender inequality index appropriateness for measuring inequality. *Journal of Evidence-Informed Social Work. 14*: 8-18.

Aassve, A., Fuochi G. and Mencarini, L. (2014). Desperate Housework: Relative Resources, Time Availability, Economic Dependency, and Gender Ideology Across Europe, *Journal of Family Issues. 35*: 1000-1022.

Aitchison, C. C. (2013) *Gender and Leisure: Social and Cultural Perspectives*. London: Routledge. ISBN: 978-0-415-26155-5.

Amici, M. and Stefani, M. L. (2013) *A gender equality index for the Italian regions* (No. 190). Bank of Italy, Economic Research and International Relations Area. ISSN: 1972-6643.

Appiah, E. N. and McMahon, W. W. (2002). The Social Outcomes of Education and Feedbacks on Growth in Africa. *Journal of Development Studies.* 38: 27–68.

Baba, K., Shibata, R., & Sibuya, M. (2004). Partial correlation and conditional correlation as measures of conditional independence. *Australian & New Zealand Journal of Statistics, 46*(4), 657-664.

Becker, W., Saisana, M., Paruolo, P. and Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological indicators. 80*, 12-22.

Bergmann, N., Scheele, A. and Sorger, C. (2019). Variations of the same? A sectoral analysis of the gender pay gap in Germany and Austria. Gender. *Work & Organization. 26:* 668-687.

Bericat, E. (2012). The European Gender Equality Index: Conceptual and Analytical Issues. *Social Indicators Research.* 108: 1-28.

Birindelli, G., Iannuzzi, A. P. and Savioli, M. (2019). The impact of women leaders on environmental performance: Evidence on gender diversity in banks. *Corporate Social Responsibility and Environmental Management. 26:* 1485-1499

Booysen, F. (2002). An overview and evaluation of composite indices of development. *Social indicators research.* 59: 115-151.
Borchorst, A. and Siim, B. (2008). Woman-friendly policies and state feminism: Theorizing Scandinavian gender equality. *Feminist theory. 9*(2), 207-224.
Cavicchia, C., Sarnacchiaro, P. and Vichi, M. (2021). A composite indicator for the waste management in the EU via Hierarchical Disjoint Non-Negative Factor Analysis. *Socio-Economic Planning Sciences. 73:* 100832.
Cavicchia, C. and Sarnacchiaro, P. (2021). The effects of a new e-tivity on students'performance and satisfaction in an online course. *Statistica Applicata, 33*: 163-175.
Cela, E., Dankelman, I. E. M. and Stern, J. (2014). *Bookshelf: Powerful Synergies: Gender Equality, Economic Development and Environmental Sustainability*. New York: UNDP.
Celis, K. and Lovenduski, J. (2018). Power struggles: gender equality in political representation. *European Journal of Politics and Gender. 1* (1-2): 149-166.
Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., Saltelli, A., ... and Tarantola, S. (2008). Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index. *Journal of the Operational Research Society. 59*(2): 239-251.
Crompton, R. (2006). *Employment and the family: the reconfiguration of work and family life in contemporary societies*. Cambridge: Cambridge University Press.
Dilli, S., Carmichael, S. G. and Rijpma, A. (2019). Introducing the historical gender equality index. *Feminist Economics. 25*(1): 31-57.
Dmitrovic, V., Dobrota, M. and Knezevic, S. (2017). A statistical approach to evaluating bank productivity. *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies. 20* (75)*:* 47-56.
Dobrota, M., Bulajic, M., Bormnann, L. and Jeremic, V. (2015). A New Approach to QS University Ranking Using Composite I-distance Indicator: Uncertainty and Sensitivity Analyses. *Journal of the American Society for Information Science and Technology*. *67* (1): 200-211.
Earles, K. (2014). Gender Equality Identity in Europe: The Role of the EU', in D. B. MacDonald and M. M. DeCoste (ed.) *Europe in Its Own Eyes, Europe in the Eyes of the Other,* pp. 103-24. Wilfrid Laurier University Press. ISBN: 978-1554588404.
EIGE. (2022). *Gender Equality Index 2022: The COVID-19 pandemic and care. European Institute for Gender Equality*. Italy: EIGE. Available at:

https://eige.europa.eu/publications/gender-equality-index-2022-covid-19-pandemic-and-care. Last access: 15/09/2022

EIGE. (2017). Gender Equality Index 2017: Methodological Report. Available at: https://eige.europa.eu/publications-resources/publications/gender-equality-index-2017-methodological-report. Last access: 26/10/2023

EIGE. (2014). EIGE's Gender Equality Index invigorates Serbia's path to Gender Equality. Available at: http://eige.europa.eu/content/news-article/eiges-gender-equality-index-invigorates-serbias-path-to-gender-equality. Last access: 23/11/2022

Elias, J. (2013). Davos Woman to the Rescue of Global Capitalism: Postfeminist Politics and Competitiveness Promotion at the World Economic Forum. *International Political Sociology. 7* (2): 152-169.

EUR-Lex. (2012). Charter of Fundamental Rights of the European Union. 2012/C 326/02

Eurostat. (2022). SDG 5 - Gender equality - Achieve gender equality and empower all women and girls. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=SDG_5_-_Gender_equality. Last access: 10/02/2022

Girón, A., and Kazemikhasragh, A. (2021). Gender equality and economic growth in Asia and Africa: empirical analysis of developing and least developed countries. *Journal of the knowledge economy*. 13: 1433-1443.

Greco, S., Ishizaka, A., Matarazzo, B. and Torrisi, G. (2018). Stochastic multi-attribute acceptability analysis (SMAA): An application to the ranking of Italian regions. *Regional Studies, 52*(4), 585–600.

Greco, S., Ishizaka, A., Tasiou, M. and Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social indicators research. 141*, 61-94.

Heise, L., Greene, M. E., Opper, N., Stavropoulou, M., Harper, C., Nascimento, M., ... & Gupta, G. R. (2019). Gender inequality and restrictive gender norms: framing the challenges to health. *The Lancet. 393* (10189): 2440-2454.

Hendriks, S. (2019). The role of financial inclusion in driving women's economic empowerment. *Development in Practice. 29* (8)*:* 1029-1038.

Hofacker, D., Stoilova, R. and Riebling, J. R. (2013). The gendered division of paid and unpaid work in different institutional regimes: Comparing West Germany, East Germany and Bulgaria. *European Sociological Review. 29* (2)*:* 192-209.

Ivanovic, B. (2007) *Classification Theory*. Belgrade: Institute for Industrial Economic.

Jeremic, V., Bulajic, M., Martic, M. and Radojicic, Z. (2011). A fresh approach to evaluating the academic ranking of world universities. *Scientometrics. 87*: 587-596.

Jeremic, V., Kostic-Stankovic, M., Markovic, A. and Martic, M. (2014). Towards a Framework for Evaluating Scientific Efficiency of World-Class Universities. *International Journal of Social, Management, Economics and Business Engineering. 8* (2): 590–595.

Kabeer, N. and Natali, L. (2013). Gender equality and economic growth: Is there a win-win?. *IDS Working Papers*. 2013(417): 1-58.

Klasen, S. (1999). *Does Gender Inequality Reduce Growth and Development? Evidence from Cross - country Regressions*, Policy Research Report on Gender and Development. Washington DC: World Bank

Klasen, S. (2006). UNDP's gender-related measures: Some conceptual problems and possible solutions. *Journal of Human Development. 7* (2): 243–274.

Kline, R. (2005). Principles and Practice of Strucutral Equation Modeling 2nd Ed.

Kilincci, O. and Asli Onal, S. (2011). Fuzzy AHP approach for supplier selection in a washing machine company. *Expert Systems with Applications. 38* (8): 9656-9664.

Lauro, N. C., Grassia, M. G. and Cataldo, R. (2018). Model based composite indicators: New developments in partial least squares-path modeling for the building of different types of composite indicators. *Social Indicators Research. 135*: 421-455.

Leeves, G. D. and Herbert, R. (2014). Gender differences in social capital investment: Theory and evidence. *Economic Modelling. 37*: 377-385.

Magalhães-Timotio, J. G., Barbosa, F. V. and Ferreira, B. P. (2022). Constructing a composite financial inclusion index for Brazil. *Revista Gestão & Tecnologia. 22*(1): 168-192.

Maricic, M., & Jeremic, V. (2023). Imposing unsupervised constraints to the Benefit-of-the-Doubt (BoD) model. *METRON*, Online first: 1-38.

Maricic, M., Egea, J. A. and Jeremic, V. (2019). A hybrid enhanced Scatter Search—Composite I-Distance Indicator (eSS-CIDI) optimization approach for determining weights within composite indicators. *Social Indicators Research*. *144:* 497-537.

Maricic, M. and Kostic-Stankovic, M. (2016). Towards an impartial Responsible Competitiveness Index: a twofold multivariate I-distance approach. *Quality & Quantity. 50* (1)*:* 103-120.

Maricic, M., Bulajic, M., Dobrota, M., & Jeremic, V. (2016). Redesigning the global food security index: A multivariate composite I-distance indicator approach. *International Journal of Food and Agricultural Economics (IJFAEC), 4*(1), 69-86.

Markovic, M., Zdravkovic, S., Mitrovic, M. and Radojicic, A. (2016). An iterative multivariate post hoc I-distance approach in evaluating OECD Better Life Index. *Social Indicators Research*. *126*: 1-19.

Mateos de Cabo, R., Gimeno, R. and Nieto, M. J. (2011). Gender Diversity on European Banks' Board of Directors. *Journal of Business Ethics. 109*: 145-162.

Melyn, W., & Moesen, W. (1991). *Towards a synthetic indicator of macroeconomic performance: Unequal weighting when limited information is available*. Leuven University, Working paper ID 26175691.

Montero, J. M., Chasco, C. and Larraz, B. (2010). Building an environmental quality index for a big city: A spatial interpolation approach combined with a distance indicator. *Journal of Geographical Systems, 12*(4), 435–459.

Munda, G. (2008). Social multi-criteria evaluation for a sustainable economy (Vol. 17). Berlin: Springer.

OECD. (2005) *Handbook On Constructing Composite Indicators: Methodology And User Guide*. STD/DOC(2005)3

Papadimitriou, E., Norlen, H. and Del Sorbo, M. (2020). JRC Statistical Audit of the 2020 Gender Equality Index. European Commission, Joint Research Centre.

Permanyer. I. (2010). The Measurement of Multidimensional Gender Inequality: Continuing the Debate. *Social Indicators Research. 95*: 181-198.

Permanyer. I. (2013). Are UNDP Indices Appropriate to Capture Gender Inequalities in Europe?. *Social Indicators Research. 100*: 927-950.

Peña, J. B. (1977). Problemas de la medición del bienestar y conceptos afines (Una aplicación al caso español). (Madrid: Instituto Nacional de Estadística (INE)).

Plantenga, J., Remery, C., Figueiredo, H. and Smith, M. (2009). Towards a European Union Gender Equality Index. *Journal of European Social Policy. 19* (1): 19-33.

Radojicic, M., Savic, G. and Jeremic, V. (2018). Measuring the efficiency of banks: the bootstrapped i-distance gar dea approach. *Technological & Economic Development of Economy*, *24* (4)*: 1581-1605*.

Reichelt, M., Makovi, K. and Sargsyan, A. (2021). The impact of COVID-19 on gender inequality in the labor market and gender-role attitudes. *European Societies*. *23* (1)*:* S228-S245.

Rogge, N. and Van Nijverseel, I. (2019). Quality of life in the European Union: A multidimensional analysis. *Social Indicators Research. 141*(2): 765-789.

Saisana, M., Saltelli, A. and Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of Royal Statistical Society. 168* (2): 307-323.

Saranya, C. and Manikandan, G. (2013). A study on normalization techniques for privacy preserving data mining. *International Journal of Engineering and Technology (IJET). 5*(3): 2701-2704.

Savic, D., Jeremic, V. and Petrovic, N. (2016). Rebuilding the pillars of sustainable society index: A multivariate post hoc I-distance approach. *Problemy ekorozwoju– problems of sustainable development. 12* (1)*:* 125-134.
Somarriba, N. and Peña, B. (2009). Synthetic Indicators of quality of life in Europe. *Social Indicators Research, 94*(1), 115–133.
Tarantola, S. and Saltelli, A. (2007). Composite indicators: the art of mixing apples and oranges. Presented on Composite Indicators – Boon or Bane. Statistisches Bundesamt, Germany. JRC43341.
Trinchera, L., Russolillo, G. and Lauro, C. N. (2008). Using categorical variables in pls path modeling to build system of composite indicators. *Statistica Applicata, 20*(2): 309–330.
UN. (2012). Indicators of Gender Equality, *Working paper 21,* Economic Commission for Europe, Conference of European Statisticians, Switzerland, Geneva, 5th March 2012.
UN. (2014a). *World survey on the role of women in development 2014 - gender equality and sustainable development*. New York: United Nations. ISBN 978-92-1-130330-8.
UN. (2000). 55/2. United Nations Millennium Declaration. Available at: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N00/559/51/PDF/N0055951.pdf?OpenElement. Last access: 15/02/2023.
UNDP. (2023). Gender Development Index (GDI). Available at: https://hdr.undp.org/gender-development-index#/indicies/GDI. Last access: 20/01/2023.
Velluti, S. (2014). Gender regimes and gender equality measures in Central Eastern European Countries post-accession: The case of Hungary and Poland. *Journal of International and Comparative Social Policy. 30* (1)*:* 79-91.
Verbunt, P. and Rogge, N. (2018). Geometric composite indicators with compromise Benefit-of-the-Doubt weights. *European Journal of Operational Research. 264* (1)*:* 388-401.
WEF. (2022). The Global Gender Gap Report 2022. Available at: https://www.weforum.org/reports/global-gender-gap-report-2022/. Last access: 10/12/2022
Witkowska, D. (2013). Gender Disparities in the Labor Market in the EU. *International Advances in Economic Research. 19*: 331-354.
Zelany, M. (1974). A concept of compromise solutions and the method of the displaced ideal. *Computers & Operations Research. 1*(3–4): 479–496