Spatial Economic Analysis

ISSN: 1742-1772 (Print) 1742-1780 (Online) Journal homepage: www.tandfonline.com/journals/rsea20

Assessing multilevel spatial inequality in EU regions: a multidimensional approach

Cinzia Di Nuzzo, Venera Tomaselli & Gianpiero Torrisi

To cite this article: Cinzia Di Nuzzo, Venera Tomaselli & Gianpiero Torrisi (05 Nov 2025): Assessing multilevel spatial inequality in EU regions: a multidimensional approach, Spatial Economic Analysis, DOI: 10.1080/17421772.2025.2579250

To link to this article: https://doi.org/10.1080/17421772.2025.2579250

SPATIAL ECONOMIC ANALYSIS
https://doi.org/10.1080/17421772.2025.2579250

Regional Studies Association
Routledge
Taylor & Francis Group

OPEN ACCESS

Assessing multilevel spatial inequality in EU regions: a multidimensional approach

Cinzia Di Nuzzo, Venera Tomaselli and Gianpiero Torrisi

Department of Economics and Business, University of Catania, Catania, Italy

ABSTRACT

The study proposes a novel multilevel and multidimensional Gini inequality index for longitudinal data. The index provides scores (i.e., weights for inequality) for dimensions, occasions and levels. This methodology is used to measure inequalities in well-being in the EU countries and their NUTS (nomenclature of territorial units for statistics) 2 regions, employed as nested spatial levels. A marked centre-periphery pattern emerges. Besides, different nuances of regional inequality are detected when EU regions themselves are used as a reference point.

ARTICLE HISTORY
Received 6 May 2025
Accepted 26 September 2025

KEYWORDS
Multidimensional well-being; multidimensional inequality Gini index; inequality measure; three-way analysis; Candecomp/Parafac model

JEL
O15; O18

1. Introduction

In the aftermath of the pandemic crisis, measuring inequality in well-being is gaining additional momentum as it is crucial for both academia and policy-making. Indeed, the strand of literature mainly building upon the seminal contribution of Stiglitz et al. (2009), in terms of measurement of economic performance and social progress, Atkinson (2015), and Piketty (2014), in terms of measurement of inequality, fostered the debate both in academic and policy circles, highlighting the importance of proper measurement for effective policies aiming at its reduction. Within this research strand, the analysis of inequality at the EU level has attracted relatively less interest than the analysis at the country or regional level, see Filauro (2018). Moreover, to the best of our knowledge, the analysis of the relative contribution of EU regional inequality to overall observed inequality in the EU has been limited to income measures only and achieved through decomposition techniques (Doran & Jordan, 2013; Filauro, 2018). Nonetheless, monitoring regional inequality according to different spatial levels of analysis is crucial for the achievement of the institutional goal stated in Article 151 in TFEU (Treaty on the Functioning of the EU), which explicitly considers each Member State as a potential source of heterogeneity.

Moreover, considering only income (or gross domestic product (GDP)) the analysis offers a rather limited view. Indeed, the aforementioned Stiglitz-Sen-Fitoussi Commission urged a ‘shift of emphasis from a production-oriented measurement system to one focused on the well-being of current and future generations, towards broader measures of social progress’ (Stiglitz et al., 2009, p. 10). This issue has been recently addressed at the regional level by adopting a multidimensional framework of analysis, e.g., Greco et al. (2021) and Torrisi (2021). Once shifted towards a multidimensional setting, however, the measurement of well-being becomes more complex, and unavoidably, the measurement of regional inequality poses additional challenges. In this regard, it is worth mentioning the attempt by Greco et al. (2018) who propose a multidimensional Gini index to measure the regional inequality in the Italian context. However, their attempt is limited to cross-sectional data – i.e., it does not allow for explicitly considering repeated observations of the dimensions of well-being.[1]

CONTACT Gianpiero Torrisi gianpiero.torrisi@unict.it

This might be a substantial limitation if measuring the evolution of disparities over time (for example, to assess if a convergence process occurred) is an objective of analysis. Furthermore, their proposed inequality measure does not allow the consideration of the spatial structure of the data. Especially from the policy-making perspective, taking into account the nations to which each region belongs might offer significant insights to interpret the observed inequality in a way that is both more comprehensive and effective. Hence, the approach proposed in Greco et al. (2018), while offering a *multidimensional* view on inequality, does not allow a *multilevel* and inter-temporal analysis of it.

In the framework of multivariate inequality measurement, several multivariate Gini indices have been proposed in the literature. Notably, Koshevoy and Mosler (1997) introduced a framework for multivariate Gini indices based on Lorenz zonoids, while Arnold and Sarabia (2018) explored majorisation and the Lorenz order as a way to generalise Gini-based measures. More recently, Bernard and Mller (2020) studied dependence uncertainty bounds in relation to the multivariate Gini mean difference. Capaldo and Navarro (2025) have proposed new efficiency multivariate Gini indices for any semi-coherent system by discussing their interpretation. Additionally, Gavilan-Ruiz et al. (2024) have developed alternative inequality measures, including the nth Gini Index, providing new insights into multivariate disparity measurement. While these indices contribute various perspectives on measuring multivariate inequality, they differ from the proposed approach in this paper.

The purpose of this paper is to contribute to filling this gap in both respects. Indeed, we propose a multidimensional Gini index that; (i) explicitly considers repeated observations of dimensions of well-being (i.e., longitudinal data) and (ii) allows for the disentangling of the contribution to overall observed inequality according to a multilevel structure of data, namely, in the case at hand, the EU as a whole, its countries, and the related NUTS (nomenclature of territorial units for statistics) 2 regions.

More specifically, departing from the aforementioned attempts, rather than decomposing overall inequality, we propose a method to compute sets of weights representing; (i) the relevance of each occurrence (e.g., year) for overall inequality and (ii) the extent to which both regions and countries – individually considered – determine overall EU inequality. Similarly, (iii) the contribution to overall inequality is computed through a set of weights for each considered dimension of well-being. Therefore, our method overcomes existing limitations by introducing a longitudinal and multilevel approach to inequality measurement. The three-way multidimensional inequality Gini index (3wMDIGI) captures disparities over time and across nested spatial levels (e.g., EU, countries and regions), offering a more detailed perspective. Using the Candecomp/Parafac (CP) model (Carroll and Chang, 1970; Harshman, 1970), it optimally weights dimensions and time periods, enhancing interpretability and accuracy in assessing multidimensional inequality.

The paper is organised as follows. Section 2 recalls a brief literature review and explains the motivation of the present work. Section 3 introduces the novel 3wMDIGI. Section 4 illustrates the multilevel inequality scores. Section 5 presents an application to the EU regions. Section 6 concludes.

2. About the methods to build synthetic indicators: a novel tool to measure inequality

The proposed multidimensional multilevel inequality Gini index entails the construction of synthetic indicators. Generally speaking, despite some criticism, synthetic indicators play an important role in statistical analysis, offering a comprehensive measure of complex phenomena by combining multiple elementary indicators or variables into a single index (Greco et al., 2019). This approach allows for synthesising diverse information and capturing the multidimensional latent nature of real-world phenomena. Building synthetic indicators involves careful consideration of various factors, including the selection of relevant variables, weighting schemes and aggregation methods.

Several notable works have contributed to the development and application of synthetic indicators in scientific research for measuring complex phenomena in various domains (Greco et al., 2019; Munda & Nardo, 2009; Tomaselli et al., 2021). They are recognised as good informative tools to understand and simplify manifold information about the multidimensional real-world as a whole and avoid misleading conclusions. Quite recently, Cavicchia et al. (2022) have proposed advanced synthetic indicator construction with the ultrametric synthetic indicator model, offering a hierarchical approach to synthesising multidimensional phenomena. Following the same ideas, Cavicchia and Vichi (2021) and Cavicchia et al. (2024)

have presented model-based approaches for constructing hierarchical synthetic indicators, ensuring robustness and addressing methodological challenges.

To build synthetic indicators, aggregative-compensative and non-aggregative methods can be used with different *pros* and *cons* for each one, depending on the nature and properties of the indicators. The aggregative-compensative method is dominant for cardinal measures. It combines or mathematically aggregates a set of indicators employing methodologies to build synthetic indicators (Nardo et al., 2005; OECD, 2008; Saisana & Tarantola, 2002). In the present study, all the elementary indicators are cardinal, and then the aggregative-compensatory approach is employed to build the system of indicators. To increase measurement accuracy and precision and compensate for random errors, the researcher has to consider the selection of elementary indicators, adopting a multi-indicator approach to identify and select several indicators for each latent or non-observable variable related to underlying dimensions.

Elementary measures can present different measurement units and ranges, and the normalisation process is required to obtain pure numbers and then make indicators comparable. The first step in the normalisation procedure is the definition of the polarity of the basic indicators through the sign of the relationship between the indicator itself and the phenomenon. After normalisation, all indicators must have positive polarity; that is, an increase in the normalised indicators corresponds to an increase in the synthetic index (Maggino, 2017). Specifically, in our application, as regards the normalisation step, elementary measures can present different measurement units and ranges, and therefore require a normalisation process to obtain dimensionless quantities and ensure comparability across variables. In this study, we apply a minmax normalisation (see Equation (12)), which rescales each indicator to the interval [0, 1] within each time period, thereby preserving the relative distribution of values. The first step in this procedure is the definition of the polarity of the basic indicators, according to the sign of the relationship between the indicator and the phenomenon. Indicators with negative polarity (e.g., mortality, unemployment) are subsequently transformed as $1 - \mathbf{x}^{*}$, so that all normalised indicators have positive polarity. This choice of normalisation facilitates the integration of variables into the weighting procedure and maintains interpretability in the inequality measures.

The process of building multi-indicator systems requires a synthetic approach based on both a conceptual model and formalised rules for combining several indirect, elementary indicators, observed at different measurement scale levels on statistical units, and organised in a matrix of data typical of multivariate statistics, to measure latent dimensions of phenomena not otherwise measurable. In this work, we consider multi-indicator systems that are multidimensional data arrays starting from three-way arrays of a matrix, like N objects $\times$ J variables $\times$ H occasions, but it can be possible to extend the process of building multi-indicator systems to a multi-way case, see Simonacci and Gallo (2024) and Bro (1997).

The importance of constructing synthetic indicators lies in their ability to provide insights into complex phenomena by taking into account a common, latent structure. Synthetic indicators are used to extend the measurement of inequality, especially in a multidimensional context. Therefore, the proposed method for measuring inequality is based on the computation of synthetic indicators of regional socio-economic performance. This approach to the measurement of regional inequality aligns with the tenet that synthetic indicators represent a way for addressing the complexity of real-world phenomena.

The main novelty of this paper is to build a multidimensional inequality Gini index by integrating weighting procedures by the Candecomp/Parafac (CP) model (Carroll & Chang, 1970; Harshman, 1970), which enables the determination of the best combination of indicators to maximise the effectiveness of the synthetic index used to measure multidimensional regional inequality. This process can assist in balancing the various dimensions of the phenomenon to improve and reflect its complexity. Moreover, introducing a multidimensional analysis approach on a three-way dataset ensures that the index is more effective in measuring variations of regional inequality over both time and dimensions.

The motivation behind this study stems from the need to address key limitations in existing approaches to the measurement of multidimensional inequality, particularly in contexts where data exhibit a three-way structure encompassing spatial, temporal and dimensional features. Previous studies and proposals offer valuable insights into regional inequalities but are limited to standard units-by-variables data. As such, they fail to capture the temporal dynamics of disparities, which are crucial for understanding how inequality evolves. Additionally, these approaches overlook the spatially nested structure of data, which is increasingly relevant considering the growing demand for regional statistics.

This paper fills these gaps by proposing a novel multidimensional Gini index tailored for longitudinal data, explicitly incorporating temporal dynamics (Section 3) and spatial multilevel structures (Section 4). This index extends traditional measures by accounting for inequalities not only within specific intervals of time but also across the entire temporal space. This dual focus provides information on the impact of long-term inequalities on annual disparities and the changes in overall inequality over time. Moreover, this approach considers the hierarchical structure of data, such as regions nested within countries and countries nested within the EU, and enables the detection of the inequality within- and between-cluster. Specifically, the main methodological innovation lies in the use of the CP model to determine weights (of variables and times) for building the synthetic index. This extends the use of principal component analysis (PCA) to three-way data, incorporating temporal dynamics by redefining variable weights to reflect the influence of time. This approach enhances the interpretability of the inequality measures and ensures that the index captures the complexity of multidimensional disparities.

Moreover, as for the aforementioned spatial structure of inequality, the methodology allows us to take into account the potential multilevel framework of data. It is worth noting that, while we apply this method to EU regions considering three levels of spatial inequality (namely EU, countries, and NUTS 2 regions), the method in itself is flexible enough to be applied to a variety of data with a multilevel (nested) structure and considering a different number of levels. This might be of particular interest in light of the increased momentum for timely, granular and harmonised sub-national statistics to facilitate informed decision-making (see, for example, the effort made by Eurostat,[2] the OECD,[3] ISTAT,[4] and the ONS[5]) in which the multilevel structure of data is more and more relevant.

The analysis offers interesting insights in different respects. To begin with, the extent of overall EU regional multidimensional inequality is substantially in line with the one based on income (measures) only, with GDP inequality being generally higher than the multidimensional counterpart. Besides, once considered jointly with other dimensions of well-being, GDP plays only a relatively minor role in shaping overall inequality. Indeed, the relative contribution to inequality (i.e., the score) of GDP in our multidimensional analysis is lower than other measures of well-being, such as the score of life expectancy, Human Resources in Science and Technology (HRST) and education.

Put differently, we find evidence that; (i) the extent of inequality captured by GDP is in line with the one obtained by adopting a multidimensional approach and (ii) mainly due to the substantial overlap between GDP and other considered measures of well-being, GDP's relative contribution to overall inequality is significantly reduced in a multidimensional setting. Our reading of this empirical evidence is that, despite the copious debate about its ability to measure well-being, GDP remains a solid variable to *measure* well-being and, consequently, inequality. However, only a multidimensional approach can unveil the domains in which GDP inequality translates into higher inequality, and, in turn, it is potentially able to better inform policies aiming to reduce it.

Moreover, once considering the multilevel structure of data, it emerges that the main driver of inequality is within-countries. This result is in line with analysis based on income only (Filauro, 2018). In this regard, our analysis unveils a rather marked centre-periphery spatial pattern, with relatively more developed central regions contributing more to inequality. This pattern is enhanced if one considers the EU regions themselves as a reference point, rather than their countries. As for the convergence process, our multidimensional approach somehow records an increase in disparities during the considered sample. Despite the short timeframe, jointly read with the results on GDP, this evidence contrasts with the trend shown by our GDP measures of inequality.

3. Three-way multidimensional inequality Gini index (3wMDIGI)

One of the key issues for building a synthetic indicator is the weighting chosen for the dimensions under analysis. Specifically, an objective approach in the weighting process of a synthetic indicator is an important aspect of ensuring the reliability and validity of complex assessments.

In this framework, the use of PCA in constructing the weights of these indices has garnered significant attention within various domains, with proponents advocating for its utility in aggregating diverse variables into comprehensive measures, see Jolliffe (2002) and McGillivray (2007). The use of PCA in determining weights for a multidimensional inequality Gini index offers significant advantages within the context of assessing inequality across multiple dimensions. Some studies have explored the application of PCA in weight determination for the multidimensional Gini inequality index. For instance, the seminal work by Bourguignon and Chakravarty (2003) discusses the use of PCA in identifying variable combinations that contribute most to overall inequality in a multidimensional context. PCA enables the identification of optimal linear combinations of dimensions to maximise the overall variance of the data, thereby identifying the most influential dimensions contributing to overall inequality.

In the three-way context, Bocci et al. (2021a) introduced a three-way approach using the STATIS method to define indexes for analysing the competitiveness of European regions, providing a valid alternative to existing indices such as the regional competitiveness index (RCI); moreover, Bocci et al. (2021b) conducted a regression tree analysis to identify key drivers of regional competitiveness, offering insights into the main determinants of the EUROSTAT RCI and classifying European regions into homogeneous groups.

Within a three-way context, constructing a multidimensional Gini index requires thoughtful consideration of weight assignment to various dimensions. The methodology used here aims to build a multi-dimensional inequality Gini index (MDIGI), which takes into account the information of the variables considered for each object (for example, an EU region) and for each occasion (for example, a specific year). Therefore, it is important to introduce a measure that describes the inequalities for each time considered and, simultaneously, a measure that summarises the inequalities over the entire time interval. Here, the weights are chosen by using an extension of PCA to three-way data. In two-way data, PCA has often been adopted to obtain synthetic indicators, where the loadings represent the weights of the synthetic measure.

Frequently, in practical applications, we confront data arrays with three dimensions. Adaptations of PCA tailored for such three-way data are recognised as three-way PCA (3wPCA). This variant has garnered significant attention due to its relevance and applicability in handling multi-dimensional data.

Formally, a three-way data can be denoted as:

$$\boldsymbol{X} = \{x_{ijh}, \text{ for } i = 1, \ldots, N, \ j = 1, \ldots, J, \ h = 1, \ldots, H\}. \tag{1}$$

For example, x_{ijh} could stand for the brightness of the (i, j) element recorded at time h.

In our approach, to build an MDIGI for three-way data (3wMDIGI), a weighting process is carried out by the Candecomp/Parafac (CP) model, which is an extension of PCA for three-way data (see Carroll and Chang, 1970; Harshman, 1970), is employed to ascertain weights attributed to different measured dimensions, considering their relative contribution to overall inequality. Specifically, in this study, we want to extend the multidimensional inequality Gini index as defined by Decancq and Lugo (2012) for three-way data.

With this aim, we recall the reduced version of rank 1 of the CP model that is used to build the inequality index. The model of rank 1 is described by,

$$\boldsymbol{X}_h = \boldsymbol{A} c_h \boldsymbol{B}^\top + \boldsymbol{E}_h, \qquad \text{for } h = 1, \ldots, H, \tag{2}$$

where:

- $\boldsymbol{A}$ is a N-vector that represents the scores of objects;
- $\boldsymbol{B}$ is a J-vector that represents the scores of the variables;
- c_h represents the score of the h-th occasion;
- $\boldsymbol{E}_h$ denotes the error matrix of size $N \times J$ at occasion h.

Note that c_h has the subscript h, which is not possessed by $\mathbf{A}$ and $\mathbf{B}$: they are invariant across different h, while c_h serves to explain the differences in $\mathbf{X}_h$ across h. An explanation of the CP model of rank greater than one can be found in Adachi (2020), Harshman (1970) and Carroll and Chang (1970). The CP model is a powerful tool for decomposing multi-way arrays. Unlike traditional matrix decomposition techniques that deal with two-dimensional data, CP handles data in three (or more) dimensions. Two of the main properties of this model are: (i) uniqueness: under mild conditions, the CP decomposition is unique up to scaling and permutation of the factors, which is a significant advantage over other tensor decomposition models; (ii) interpretability: the factor matrices **A**, **B** and **C** provide direct insights into the underlying structure of the data.

Let $\boldsymbol{C} = (c_1, \ldots, c_H)^\top$ be the H-vector contain the score on each occasion from Equation (2). To provide the weights of the 3wMDIGI, we consider the following products,

$$\bar{\boldsymbol{W}} = \boldsymbol{BC}^\top, \tag{3}$$

where $\bar{\mathbf{W}}$ is an $J \times H$ matrix of unnormalised weights on J variables over H occasions. Since the weights have to sum to one, we normalise the previous quantity as follows,

$$w_{jh} = \frac{\bar{w}_{jh}}{\sum_{J}^{m=1} \bar{w}_{mh}}, \qquad \text{for } j = 1, \ldots, J, \;\; h = 1, \ldots, H. \tag{4}$$

Therefore, one can build a $J \times H$ matrix of the weights of the multidimensional index as,

$$\boldsymbol{W} = (w_{jh}), \qquad \text{for } j = 1, \ldots, J, \;\; h = 1, \ldots, H. \tag{5}$$

So, following the same idea of Decancq and Lugo (2012), the scores of the 3wMDIGI are computed by,

$$s_{ih} = \left(\sum_{J}^{j=1} w_{jh} (x_{ijh})^\beta \right)^{1/\beta}, \qquad \text{for } h = 1, \ldots, H \text{ and } i = 1, \ldots, N \tag{6}$$

where β is a parameter and x_{ijh} is the original data at i-th object, j-th variable and h-th occasion; thus s_{ih} can be interpreted as the scores of the multidimensional inequality Gini index of the i-th object over the h-th occasion, for $h = 1, \ldots, H$ and $i = 1, \ldots, N$. To produce a ranking of the N objects, we sort $\mathbf{s}_i$ in non-increasing order, for all $h = 1, \ldots, H$. Let r_{ih} be the rank of the object i at the occasion h based on s_{ih}. Let $\mathbf{R} = (r_{ih})$ be the $N \times H$ matrix of the ranks for $i = 1, \ldots, N$ and $h = 1, \ldots, H$.[6] Let $\boldsymbol{M} = (\mu_{jh})$ be the means matrix, where $\mu_{jh} = \frac{\sum_{i=1}^{N} x_{ijh}}{N}$, for $j = 1, \ldots, J$ and $h = 1, \ldots, H$.

Therefore, the 3wMDIGI is defined by,

$$I_h(\boldsymbol{X}) = 1 - \frac{\sum_{N}^{i=1} \left[\left(\frac{r_{ih}}{N} \right)^\delta - \left(\frac{r_{ih} - 1}{N} \right)^\delta \right] \left(\sum_{J}^{j=1} w_{jh} (x_{ijh})^\beta \right)^{1/\beta}}{\left(\sum_{J}^{j=1} w_{jh} \mu_{jh}^\beta \right)^{1/\beta}}, \quad \text{for } h = 1, \ldots, H, \tag{7}$$

where δ and β are two positive parameters.

According to the reference literature, admissible parameter intervals satisfying the unfair rearrangement (UR) principle[7] are $\beta < 1$ and $1 \leq \delta \leq 2$. More in detail, $\beta = 1 - 1/\sigma$ where σ represents the elasticity of the substitution. This parameter captures the degree of (allowed) substitutability between the dimensions of well-being related to the elasticity of substitution between dimensions. Hence, the case in which $\beta = 1$ refers to dimensions as perfect substitutes in a fully compensatory framework. By contrast, if β tends to $-\infty$, then the relationship between dimensions tends towards a perfect complementarity.

Furthermore, the specific case $\beta = 0$ relates to the unitary-elasticity case $\sigma = 1$ for which the ratio-scale invariance is satisfied only by a proper Cobb–Douglas well-being function (Decancq & Lugo, 2012, p. 728). As for δ, it captures the welfare weights, i.e., the judgments on the relative importance of the different units involved in the evaluation. The higher the δ, the more importance is given to the bottom of the distribution. Therefore, the limit case δ tending to $-\infty$, identifies a Rawlsian approach where only the worst-off performance counts towards the evaluation exercise. Generally speaking, for $\delta \in [0, 1]$, the best-off units are given more weight. By setting $\delta = 2$, the standard Gini social evaluation function is obtained.[8]

Hence, it is worth noting here how both β and δ capture different judgments and beliefs underlying the evaluation exercise. Put differently, while their overall effect on the measure(s) of inequality depends on the actual data under investigation, the choice of β and δ *per se* inherently establishes the degree of complementarity between and within dimensions, respectively.

The 3wMDIGI defined in Equation (7) gives a measure of the disparities among statistical objects for each occasion. This is useful to compare how to change the disparities of variables over a certain time interval. However, it can be important to define a 3wMDIGI that simultaneously provides a unique measure of the disparities among objects, variables and occasions. Because the simple mean of the Gini indicators found for each year/occasion $h = 1, \ldots, H$ can be misleading, as there may be certain specific years that have a greater impact on the inequalities among objects compared to other years/occasions. Therefore, we define a *GLOBAL 3wMDIGI* as,

$$I(\mathbf{X}) = 1 - \frac{\sum_{N}^{i=1}\left\{\left[\left(\frac{\tau_i}{N}\right)^{\delta} - \left(\frac{\tau_i - 1}{N}\right)^{\delta}\right]\left(\sum_{J}^{j=1} \omega_j(\tilde{x}_{ij})^{\beta}\right)^{\frac{1}{\beta}}\right\}}{\left(\sum_{J}^{j=1} \omega_j[\mu(\tilde{x}_j)]^{\beta}\right)^{\frac{1}{\beta}}}, \tag{8}$$

where the elements have all the same role as Equation (7) except for the ω_j that, according to Equation (2), $\omega_j = b_j$ represents the weight of indicator such that $\omega_j > 0$ and normalised to have $\sum_{j=1}^{J} \omega_j = 1$; and $\tilde{x}_{ij}$ that is defined as $\widetilde{\mathbf{X}} = \mathbf{X}_i\mathbf{C}$ is a matrix of size $N \times J$, where $\mathbf{X}_i$ is the slice of three-way data of size $J \times H$. Note that $\widetilde{\mathbf{X}}$ represents the data of N objects by J variables but scaled for the loadings for each occasion. Therefore, as in Equation (7), τ_i is a shorthand for $\tau_i(\phi)$ the ranks of unit i on the basis of $\phi_i = \left(\sum_{j=1}^{J} \omega_j(\tilde{x}_{ij})^{\beta}\right)^{\frac{1}{\beta}}$.

Moreover, we underline that the choice between the annual 3wMDIGI (Equation (7)) and the Global 3wMDIGI (Equation (8)) depends on the analytical or policy question at hand. The 3wMDIGI is preferable when the objective is to monitor and compare disparities over specific occasions/years, for example, to identify years in which inequality sharply increases or decreases, or to evaluate the short-term impact of policy interventions. By contrast, the Global 3wMDIGI aggregates information across occasions/years and yields a single synthetic measure of disparities among objects, variables and time. This is particularly useful for summarising multi-year inequality patterns into one index, facilitating cross-sectional benchmarking, and supporting policy discussions that require a long-term, structural view of disparities rather than annual fluctuations. In practice, the two measures are complementary: the 3wMDIGI provides a time-resolved perspective, whereas the Global 3wMDIGI condenses the temporal dimension into a unified inequality assessment.

Finally, as aforementioned, it is worth noting that in Equations (7) and (8) the two hyperparameters β and δ must be selected a priori. In this regard, although Decancq and Lugo (2012) provides an admissible range in which these two hyperparameters should vary, that is $\beta < 1$ and $1 \leq \delta \leq 2$; their selection turns out to be quite challenging and the choice depends on the objective of the analysis.

Since, in this work, the aim is to construct an indicator that highlights disparities among the objects, it is important to choose β and δ such that they yield the highest possible value of the indicator to be constructed. Therefore, to select the two hyperparameters, a sensitivity analysis must be carried out and β and δ are subsequently chosen to maximise the inequality indicator in Equations (7) and (8). Hence, a cross-analysis is performed between a set of values selected for $0 < \beta < 1$ and a set of values selected for $1 \leq \delta \leq 2$ to highlight the inequalities or to maximise the outcome.

3.1. CP vs PCA

Here, we want to note that the motivation behind the development of 3wMDIGI stems from the need to extend traditional multidimensional inequality measures to data structures involving a third dimension, such as time. Existing approaches typically measure inequality within a single timeframe, neglecting the broader temporal dynamics. The proposed index addresses this limitation by enabling the evaluation of inequality across units and variables within each year while simultaneously considering disparities over an extended period.

This dual focus provides two critical insights; (i) how inequality across the entire timeframe impacts individual years and (ii) how overall inequality evolves over time, offering a foundation for assessing the effectiveness of policies or interventions aimed at reducing disparities. To achieve this, we have proposed a new way to weigh the units and variables, specifically extending the use of PCA to three-way data. While PCA has been widely used to construct indices for two-way data, it has not, to the best of our knowledge, been applied in contexts that incorporate temporal dynamics as an additional dimension.

Our contribution lies in redefining variable weights, as presented in Equations (3)–(6), to integrate the influence of time, and introducing data weighting in Equation (7) to account for temporal variations. These innovations enhance the interpretability and utility of inequality measures in complex data structures. This approach is particularly relevant for datasets with nested structures (see Section 4), such as those found in socio-economic and demographic studies. By capturing temporal disparities and their interplay with variable and unit-level features, 3wMDIGI provides a more comprehensive framework for inequality analysis, filling a critical gap in the existing literature.

Moreover, we note that the three-way structure of the data offers benefits that cannot be replicated by collapsing dimensions or performing subset analyses. This structure captures the interrelations between dimensions (e.g., regions, variables and time) that would otherwise be lost. For instance, collapsing the temporal dimension into a single average obscures year-to-year variability and the direction of disparities over time, while analysing subsets of data (e.g., individual years or variables) isolates specific components without providing a total view of their interactions. Maintaining the three-way structure enhances interpretability by supporting the decomposition of inequality into contributions across dimensions and time.

From a methodological point of view, the CP model avoids over-representation of certain dimensions that might occur in collapsed data, particularly when one dimension (e.g., time) has a higher cardinality than others. By treating the three dimensions independently, our approach ensures a balanced representation of all dimensions in the computation of weights. Indeed, the CP model is explicitly designed to analyse three-way data arrays, enabling one to account for interactions between units, variables and time in a unified framework. Collapsing the last two dimensions into one, as with PCA on a two-dimensional array (of size $N \times JH$), results in the loss of critical information about the temporal and inter-variable relationships.

This simplification neglects the temporal dynamics and the interplay between variables within each temporal instance, which are useful for capturing the complexity of multidimensional inequality over time. For example, a CP model allows for assigning weights that vary across variables and occasions, offering a richer understanding of how each contributes to overall inequality. In contrast, collapsing the data for PCA would lead to uniform weights within the combined dimension, failing to capture the contributions of specific years and variables. In other words, if we consider the collapsed data of size $N \times JH$ and the PCA on this new dataset, the weights of the Gini indicator lie in a JH-vector $\mathbf{w}$, when we consider the sum in Equation (6) we lose the dimension on occasion.

Therefore, collapsing dimensions can introduce misleading results: applying PCA to collapsed data (e.g., regions × variables) cannot identify the distinct influence of time or temporal variation in variable contributions, limiting the depth of insights. For example, aggregating GDP over the years may allow years with extreme disparities to dominate, masking the contributions of other years. Subsetting the data may produce wrong results by focusing narrowly on one aspect of the data without considering its broader context. Retaining the three-way structure decreases these risks by ensuring that results reflect true patterns rather than artifacts of aggregation or selection.

Thus, while collapsing dimensions or using subset analyses is technically possible, the three-way structure of the data provides unparalleled advantages: it preserves multidimensional relationships, avoids misleading, enables advanced methodologies and ensures practical relevance, making it essential for the proposed approach.

4. Multilevel inequality scores

As argued, our approach allows for the measurement of inequality explicitly considering the multilevel structure of data. More specifically, it provides inequality scores for nested – i.e., organised by levels – statistical objects. Figure 1 illustrates the multilevel structure of data. Multilevel systems, such as the EU's nested structure of countries and regions, require a framework that preserves the relationships to assess inequality at each level.

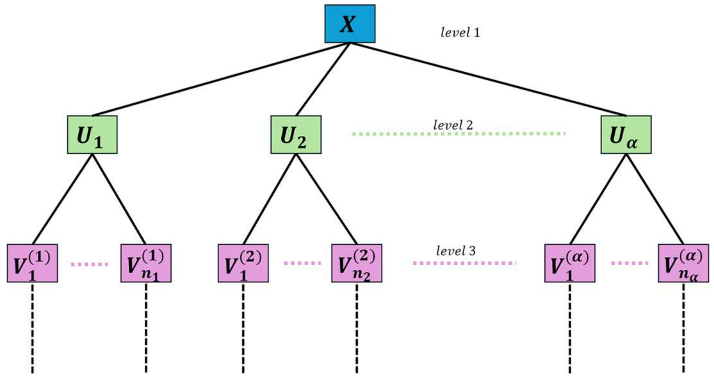

Figure 1. Graphical representation of multilevel inequality scores. Source: authors' elaboration.

For illustrative purposes, Figure 1 presents a simplified case with only three levels, although the methodological framework is flexible enough to accommodate more granular structures, as suggested by the final dotted vertical lines.

Level 1 represents the highest level, including all the objects. In level 2, objects are grouped into α disjoint sets U_k. Elements of each $\{\mathbf{U}_1, \ldots, \mathbf{U}_\alpha\}$ are, in turn, organised into n_k disjoint sets $V_{n_k}^{(k)}$ (for $k = 1, \ldots, \alpha$).

Within this multilevel framework, the quantities,

$$\eta_{ih} = \frac{\left(\sum_{J}^{j=1} w_{jh}(x_{ijh})^{\beta}\right)^{1/\beta}}{\left(\sum_{J}^{j=1} w_{jh}\mu_{jh}^{\beta}\right)^{1/\beta}}, \qquad i = 1, \ldots, N, \quad h = 1, \ldots, H, \tag{9}$$

and

$$\gamma_i = \frac{\left(\sum_{J}^{j=1} \omega_j(\tilde{x}_{ij})^{\beta}\right)^{1/\beta}}{\left(\sum_{J}^{j=1} \omega_j\mu(\tilde{x}_j)^{\beta}\right)^{1/\beta}}, \qquad i = 1, \ldots, N, \tag{10}$$

can be interpreted as the multilevel inequality scores of the inequality Gini index among objects and occasions in Equation (9); and among objects in Equation (10).

For the sake of conciseness, from this point onward, we elucidate the utility of these scores for the *Global 3wMDIGI* (Equation (8)). However, it is worth noting that the discussion applies to the 3wMDIGI defined in Equation (7), the only difference being the repetition of what we will state for each occasion $h = 1, \ldots, H$.

The inequality scores in Equation (10) are of paramount importance when ranking objects. They provide a measure of the effect that the i-th object has on the disparities among all other objects. Furthermore, they provide measures of disparity relative to objects belonging to a specific set. For example, (the set of) European regions belonging to a given country, that is, EU regions within a country, see Section 5.

More formally, let $\mathbf{X}$ be the three-way data and let $\mathcal{U} = \{\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_\alpha\}$ be the common partition for all occasions; i.e., given $\mathbf{X}_h$ for $h = 1, \ldots, H$, $\mathbf{X}_h = \mathbf{U}_1 \cup \cdots \cup \mathbf{U}_\alpha$ such that $\mathbf{U}_l \cap \mathbf{U}_r \neq \emptyset$ (for all $l \neq r$ $l, r = 1, \ldots, \alpha$) for all $h = 1, \ldots, H$.

Let us suppose that a partition exists for each $\mathbf{U}_l$ (for $l = 1, \ldots, \alpha$); i.e., given a certain $\mathbf{U}_l$, $\mathbf{U}_l = \mathbf{V}_1 \cup \cdots \cup \mathbf{V}_{n_l}$ with $\mathbf{V}_{l_1} \cap \mathbf{V}_{l_2} \neq \emptyset$ for each $l_1 \neq l_2$ $(l_1, l_2 = 1, \ldots, n_l)$, where $n_1 + n_2 + \cdots + n_\alpha = N$. In practice, given a certain occasion, data can be structured as in Figure 1, where we can consider $\mathbf{X}$ as a nested hierarchical structure organised by levels: level 1 corresponds to the initial data $\mathbf{X}$, level 2 corresponds to the sets belonging to the partition $\mathcal{U}$, and level 3 corresponds to the union of the set of the partitions $\mathcal{V}_p$ for $p = n_1, \ldots, n_\alpha$.

Therefore, referring only to a Global 3wMDIGI (Equation (8)), but similar considerations hold for MDIGI (Equation (7)), it is possible to define the following quantities,

$$\vartheta_{VU} = \frac{\boldsymbol{\gamma}_{|\mathbf{V}^*}}{\boldsymbol{\gamma}_{|\mathbf{U}^*}}, \qquad \vartheta_{UX} = \frac{\boldsymbol{\gamma}_{|\mathbf{U}^*}}{\boldsymbol{\gamma}} \tag{11}$$

which can be interpreted as the multilevel inequality scores of the objects within and between sets of objects and where $\mathbf{V}^* \subseteq \mathbf{U}^* \subseteq \mathbf{X}$. Specifically, ϑ_{VU} represents the contribution to disparities among the objects within level 3 over the objects within level 2; and so on. It is worth stressing here that this interpretation stems from the assumption that the Gini for a higher level can be somehow interpreted as a mean of the Gini for the lower level.

In this multilevel context, it is important to note that structures with more than three nested levels can also be considered. Each level represents a partition of the initial set, allowing us to identify groups of objects that share common characteristics in terms of inequality loadings. Thus, every level in the hierarchy provides an additional layer of granularity in examining the disparities among objects. Furthermore, the sets of objects comprising each partition can be interpreted as clusters of objects. For instance, when considering territorial regions or spatial objects, these clusters can represent groups of countries that exhibit similar characteristics in terms of inequality loadings for the initial variables under consideration.

Specifically, if we extend our framework beyond three levels, we might have a fourth level, Level 4, where objects in each $\mathbf{V}_{l_k}$ are further partitioned into other disjoint sets. This additional partitioning allows for a more refined analysis of inequality within increasingly specific subgroups of the original dataset.

This adaptability is particularly useful in the analysis of the territorial or spatial data, where clusters of regions or countries can be examined in terms of their contributions to overall inequality, providing insights into the multilevel nature of disparities.

Finally, to clarify the main features of 3wMDIGI, Table 1 lists the differences between our proposal and the previous works on multidimensional Gini index (e.g., Capaldo & Navarro, 2025; Decancq & Lugo, 2012; Gavilan-Ruiz et al., 2024; Greco et al., 2019; Koshevoy & Mosler, 1997).

5. Disparities among EU regions via the multidimensional inequality Gini index

For the sake of illustration, we apply our methodology to the measurement of inequality among EU NUTS 2 regions. We use data on multidimensional well-being from EUROSTAT, ranging from 2013 to 2019 ($H = 7$). In addition to economic indicators such as GDP per capita and proxies for human capital represented by HRST (human resources in science and technology, persons with tertiary education (ISCED) and/or employed in science and technology, percentage of population in the labour force) and education (participation rate in education and training, population from 25 to 64 years) the considered $J = 10$ variables include demographic indicators such as population, fertility rate, mortality rate (deaths per thousand people), infant mortality rate and life expectancy.

Furthermore, aligning with other notable previous attempts to measure multidimensional well-being, such as the EU Regional Competitiveness Index (Annoni et al., 2017, p. 30) and the OECD's Better Life Index (Durand, 2015, p. 10), we include three variables related to the job market. Namely, employment (per thousand persons), unemployment rate (per thousand persons), and rate of NEET (young people neither in employment nor in education and training, from 15 to 29 years). It is worth noting that the economic rationale for including both employment and unemployment measures lies in their ability to capture the different nuances of their effect on well-being in a more comprehensive way, according also to a theoretical standpoint.

Indeed, while employment somehow reflects the extent of opportunities available in each region and has a marked direct effect on disposable income, unemployment, in addition to its implications in terms of (lack

Table 1. Main differences and features between previous works and our proposal for multidimensional Gini index.

Feature	Previous works	Proposed 3wMDIGI
Time consideration	Mostly cross-sectional	Longitudinal data
Weighting process	PCA or decomposition methods	CP model for three-way data
Spatial structure	Limited to national or regional levels	Multilevel: EU, countries, NUTS2 regions. Scope: Specific within- and between-region disparities

Source: authors' research.

of) opportunities, shows persistent adverse effects on an individuals subjective perception of social status, integration (Pohlan, 2024) and, eventually, social exclusion (Haataja, 1999).

Due to data availability, we focused on $N = 234$ European regions in the following countries: Belgium, the Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, the Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia and Sweden.

Tables 2 and 3 report, respectively, the summary statistics of the data and the correlation between variables before the transformation reported in Equation (12). The analysis of Tables 2 and 3 provides some relevant preliminary information for our study.

Indeed, in terms of disparities, Table 2 shows a quite significant degree of variation across EU regions. For example, the Population variable (representing a general proxy for the attractiveness of territories) demonstrates substantial variation, with values ranging from 83,620 to 12,300,000. This is further captured by the high standard deviation (1,626,865). It shows a between-areas standard deviation of 1,629,614, and internal variation (within) of 27,945.94, indicating, as expected in this case, that most variability is due to differences between areas rather than within the same areas over time.

Similarly, the variable Deaths varies from 3.11 to 20.35 (deaths per thousand inhabitants), indicating considerable disparity among different regions. It exhibits a between-areas standard deviation of 2.073 and a much lower within-areas standard deviation of 0.33, indicating that most variability is found between different regions. The related variable Infant Mortality rate ranges from 7 to 11.3, with a relatively small standard deviation of 1.57 and a similar pattern of between and within-areas standard deviation, respectively of 1.42 and 0.690. On a more economic note, it is worth noting how the rate of NEET (persons) varies significantly from 4.1 to 40.3, highlighting notable regional differences unfolding mainly between regions (standard deviation of 7.1 as compared to the within datum of about 1.930).

Table 2. Summary statistics of the raw data.

Variable		Mean	Std. dev.	Min	Max	Observations
Population	overall	1,854,427	1,626,865	83,620	12,300,000	$M = 1652$
	between		1,629,614	84,452.43	12,100,000	$N = 234$
	within		27,847.08	1,596,372	2,038,748	$H = 7$
Fertility	overall	1.548	0.2765812	0.94	3.82	$M = 1638$
	between		0.2686716	1002857	3.604286	$N = 234$
	within		0.0676558	1.266129	1.816129	$H = 7$
Deaths	overall	10.47431	2.095648	3.1148	20.3515	$M = 1638$
	between		2.073179	3.324321	19.66157	$N = 234$
	within		0.3307926	9.08423	11.63342	$H = 7$
Infant Mortality	overall	3.577411	1.574791	0.7	11.3	$M = 1638$
	between		1.418385	1.714286	9.042857	$N = 234$
	within		0.6895831	−4797314	8.84884	$H = 7$
Life Expectancy	overall	80.68327	2.575234	73	85.8	$M = 1638$
	between		2.557859	73.47143	85.11429	$N = 234$
	within		0.3364003	79.42613	82.02613	$H = 7$
NEET	overall	14.46422	7.344448	4.1	40.3	$M = 1638$
	between		7.099462	5.114286	38.77143	$N = 234$
	within		1.929585	5.035653	23.16422	$H = 7$
HRST	overall	41.0348	9.523968	15.1	70.7	$M = 1638$
	between		9.344681	16.214429	66.91429	$N = 234$
	within		1.924308	33.44909	49.24908	$H = 7$
Employment	overall	804.3739	692.5581	23.5	5488.4	$M = 1638$
	between		692.8663	25.92857	5338.743	$N = 234$
	within		36.50476	532.7309	1079.731	$H = 7$
Unemployment	overall	80.02149	118.5681	3.8	1460.6	$M = 1638$
	between		115.5962	4.771429	1149.171	$N = 234$
	within		27.2926	−230.5499	391.4501	$H = 7$
GDP	overall	27,153.54	14,440.75	3600	100,400	$M = 1638$
	between		14,283.64	4400	95,814.29	$N = 234$
	within		2293.63	−1932.173	52,167.83	$H = 7$
Education	overall	10.29713	7.11582	0.6	36	$M = 1638$
	between		7.0417	0.8	32.14286	$N = 234$
	within		1.082037	5.011417	15.46856	$H = 7$

Note: Please note that for the ease of interpretation, departing from the usual notation used with panel data and consistently with the notation used in our 3wMDIGI we use M, N and, H to index the total number of observations, the number of regions included in the panel, and the number of periods, respectively.

Source: authors' elaboration.

Table 3. Correlation table.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Population	1.000										
(2) Fertility	−0.045	1.000									
(3) Death	−0.120	−0.257	1.000								
(4) Infant Mortality	0.830	0.162	−0.084	1.000							
(5) Life Expectancy	0.155	−0.027	−0.530	−0.116	1.000						
(6) NEET	0.002	−0.070	0.062	0.068	−0.128	1.000					
(7) HRST	0.153	0.169	−0.350	0.054	0.374	−0.636	1.000				
(8) Employment	0.977	−0.022	−0.114	0.819	0.143	−0.136	0.233	1.000			
(9) Unemployment	0.755	−0.145	−0.196	0.565	0.210	0.273	−0.021	0.648	1.000		
(10) GDP	0.112	0.180	−0.432	0.007	0.556	−0.537	0.718	0.192	−0.048	1.000	
(11) Education	0.026	0.338	−0.375	−0.052	0.465	−0.445	0.553	0.056	0.002	0.590	1.000

Source: authors' elaboration.

Besides these, Fertility has a narrower range, with minimum and maximum values of 0.94 and 3.82, respectively, indicating less disparity compared to the population. Fertility has a between-areas standard deviation of about 0.270 and a within-areas standard deviation of 0.0677, suggesting again that variability between-areas is greater than the temporal variability within the same areas. Life Expectancy also shows less disparity, with a range between 73 years and 85.8 years and a relatively low standard deviation of 2.58, indicating more consistency in life expectancy across EU regions.

The correlation table (Table 3) offers interesting insights into the potential patterns of multidimensional regional inequality. Indeed, Population, while having a positive correlation with Infant Mortality rate (0.830) and, to a lesser extent, with Life Expectancy (0.155), has a rather weak negative correlation with both Fertility (−0.045) and the number of Deaths (−0.120). Furthermore, the high correlation between Population and Employment (0.977) suggests that more populous areas tend to have higher employment levels, and vice versa. This confirms, as aforementioned, the general tenet of the population as a proxy for regional attractiveness. Life Expectancy is positively correlated with HRST (0.374) and GDP (0.556), but negatively correlated with Infant Mortality (−0.116) and NEET (−0.128). NEET, in turn, is negatively correlated with HRST (−0.636) and GDP (−0.537). HRST has significant positive correlations with GDP (0.718) and Education (0.557), showing, to some extent, synergies applied to economic productivity. GDP, in turn, shows significant positive correlations with Education (0.590).

Definitely, the analysis of descriptive statistics and correlations reveals substantial disparities and variability across different regions, highlighting significant inequality in socio-economic indicators:

- Population and Employment: Regions with larger populations tend to have higher levels of employment, indicating a concentration of economic activity and resources in more populous areas. The wide range and high standard deviation in population and employment figures suggest significant spatial inequality in economic opportunities and resources.
- Health Indicators (Fertility, Deaths, Infant Mortality, Life Expectancy): Health indicators such as fertility, deaths, infant mortality and life expectancy show considerable variability. High infant mortality rates and death rates in certain regions signal severe spatial disparities in healthcare access and quality. The negative correlation between infant mortality and life expectancy suggests that areas with higher infant mortality rates suffer from lower overall life expectancy, reflecting broader issues in health and socio-economic conditions.
- Economic Indicators (GDP, HRST, NEET): Economic indicators exhibit significant disparities, with GDP varying widely between regions. Regions with higher GDP also tend to have higher educational attainment and better human resources in science and technology. The negative correlation between NEET and both GDP and HRST suggests that regions with higher economic development and educational resources potentially provide better stimulus to participate in economic activities and personal development. Overall, GDP confirms a substantial correlation with other indicators.
- Education: Education levels also show substantial variability, with some regions exhibiting significantly higher levels of educational attainment. The positive correlation between education and GDP highlights the critical role of education in economic development and reducing inequality.

To summarise, the analysis of summary statistics underscores the existence of significant multidimensional spatial inequalities. These disparities indicate that socio-economic resources and opportunities are unevenly distributed, leading to uneven levels of well-being and development. Hence, this evidence provides further rationale for constructing a multidimensional inequality Gini index that highlights disparities among the various variables considered across European regions and for different years.

5.1. Three-way approach to multidimensional inequality Gini index

Building upon the argument presented in the previous section, in this section, we present our 3wMDIGI as applied to EU regions. Preliminary, to make data comparable, given $\bar{h} \in \{1, \ldots, H\}$, we adopt the following transformation:

$$X^*_{ij\bar{h}} = \frac{X_{ij\bar{h}} - \min x_{j\bar{h}}}{\max x_{j\bar{h}} - \min x_{j\bar{h}}}, \tag{12}$$

where $x_{j\bar{h}}$ represents the jth column of X at a given occasion $\bar{h} \in \{1, \ldots, H\}$. Afterward, following the transformation in Equation (12), for variables with negative polarisation such as mortality and unemployment, we reordered the data as $1 - x^*_{i\bar{j}\bar{h}}$, for those specific $\bar{j} \in \{1, \ldots, J\}$, where $x^*_{i\bar{j}\bar{h}}$ represents the ith row of X^* at variable $\bar{j}$ and occasion $\bar{h}$.[9] Therefore, each variable assumes values in [0, 1].

Based on the transformed dataset $\mathbf{X}^*$, we compute the 3wMDIGI defined in Equation (7), where, consistently with Decancq and Lugo (2012), the parameters δ and β are set equal to 2 and 0.005, respectively. More in detail, while the value of δ is set to obtain the standard Gini social evaluation function, the value of β has been set to maximise the computed 3wMDIGI over a rather comprehensive grid of parameter pairs, as outlined in the sensitivity analysis in Section 5.1.1. More generally, as aforementioned, for a given δ, the overall value of the Gini index depends on the interplay between the weighting scheme and the parameter β. As already pointed out, the choice of the pair (β, δ) reflects the underlying judgments on both substitutability between the dimensions and on the relative importance of lower values of the selected dimensions (bottom sensitivity). However, this illustrative evaluation exercise adopts a data-driven approach to both weighting and parameter selection. Hence, we propose to set β to let the multidimensional inequality emerge to the highest extent.[10] Having said that, it is worth noting that the fact that such a β is close to zero reflects a very limited extent of substitutability between considered dimensions of well-being and, furthermore, it is consistent with the aforementioned normative space given by bottom sensitivity (δ) and substitutability (β) detected in Decancq and Lugo (2012).

Results are shown in Figure 2 and reported in the first row of Table 4. Figure 2 reports the 3wMDIGI for each year in our sample. Once more, it is worth stressing that in the computation of the Gini index for each year, the whole pattern of regional inequality during the whole sample is taken into account. According to

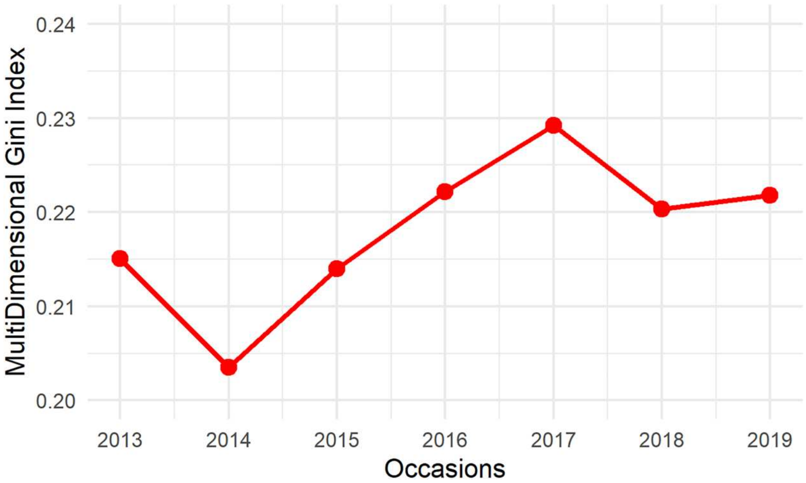

Figure 2. 3wMDIGI according to Equation (7) by varying $h \in \{2013, \ldots, 2019\}$. Source: authors' elaboration.

Table 4. Multidimensional inequality Gini index.

	2013	2014	2015	2016	2017	2018	2019
3wMDIGI	0.2150593	0.2035066	0.2139918	0.2221749	0.2291844	0.2203557	0.2217859
3wMDIGI-GPD	0.2296349	0.2246350	0.2317876	0.2369283	0.2345096	0.2417920	0.2432570
Gini-GPD	0.291096	0.290316	0.290712	0.291047	0.286800	0.282940	0.280563

Note: A comparison among indices.
Source: authors' elaboration.

this novel multidimensional metric of inequality, a convergence pattern is observed in the EU during the sample considered. More in detail, although in 2018 a slight increase in disparities with respect to the previous year is observed, overall a clear decreasing trend is registered during the whole sample.

Figure 3a and 3b reports the loadings of each dimension and each year in the overall inequality index, respectively. Hence, Figure 3 shows that (a) the mainstream measure of GDP plays a relatively minor role concerning, for example, the prominent role of life expectancy in shaping the multidimensional Gini index. This evidence somehow confirms the case for the simultaneous consideration of measures other than GDP. Furthermore, the aforementioned result concerning life expectancy confirms its role as a proxy for general living conditions.

As for the single time-occurrences, Figure 3b shows that, in the case at hand, more recent values have a higher relative weight in determining the multidimensional Gini index. While this signals the strength of the path-dependence in EU regional disparities, we acknowledge that the considered time sample is

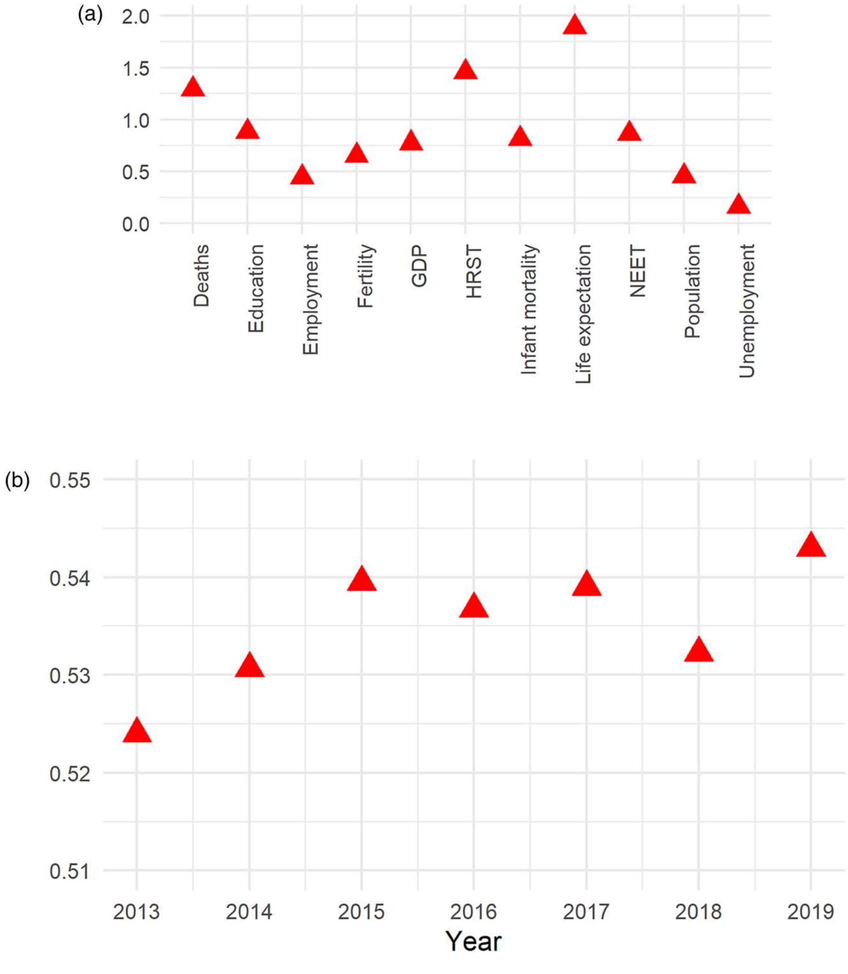

Figure 3. (a) Loadings of variables B; (b) Loadings of occasions $C = (c_1, \ldots, c_H)^\top$; according to the CP model (Equation (2)). Source: authors' elaboration.

relatively short to properly appreciate the time pattern. However, it is worth noting how our proposed methodology is innovative, also to the extent it allows us to consider this aspect of regional disparities.

Moreover, to compare and contrast our multidimensional Gini index with the (mainstream) Gini index computed on GDP, we include in Table 4 two additional Gini indices. Namely, 3wMDIGI-GDP and Gini-GDP. The former is our proposed multidimensional Gini index computed excluding the GDP variable; the latter is, indeed, the Gini index computed on GDP only.

The rationale is as follows. 3wMDIGI-GDP aims to capture the inequality due to variables other than GDP, i.e., the multidimensional inequality measured regardless of the strictly economic aspect represented by GDP. Gini-GDP, as it is well-known, captures the degree of inequality as measured by the single metric represented by GDP. Therefore, by simultaneously considering the three indices, one can gather information on the potential overlaps or complementarity between them. Indeed, the closer the alternative measures are, the more similar the information on the extent of inequality they can convey. In the case at hand, the three different indices share the same spatial pattern and, therefore, the information they convey in terms of the measurement of the evolution of inequality can be considered substantially overlapping.

More precisely, the index computed on GDP-only shows values slightly higher than those computed with the alternative multidimensional measures. However, the general picture is substantially unchanged (Figure 4). Hence, if the aim of the analysis were (limited to) 'measuring' inequality, then GDP encapsulates enough information on overall levels of well-being.

However, only a multidimensional approach can disentangle the extent and the eventual patterns in which differences in GDP translate into differences in dimensions of well-being. This peculiar aspect is missing in the approach based on GDP only and rather, it is reflected in the different weights for each dimension, like the ones reported in Figure 3. Therefore, our analysis somewhat confirms that a multidimensional approach to inequality offers a more effective standpoint to both design and implement policies aiming at 'reducing' inequality by addressing the way it affects the lives of people.

Consistent with the methodological framework reported in Section 4, in addition to the Gini indices presented above, we break the contribution to inequality of EU regions according to Equations (9)–(11). The graphical representation of our application is represented in Figure 5, hence, there are three levels of a set of territorial objects in the EU: level 1 is the EU; level 2 is the partition of the EU into countries; and level 3 is the collection of the set of regions of each EU country.

Results of multilevel inequality scores (Equation (10)) and the ratio scores (Equation (11)) are shown in Figures 6–8. More specifically, Figure 6 shows the overall regional contribution to inequality (Equation (10)) as measured by the 3wMDIGI for each of the NUTS 2 regions in our sample. Figure 6 reports a quite spatially even pattern of weights in which a clear spatial drive of inequality cannot be detected. Figures 7 and 8 show the contribution to inequality of each European NUTS 2 region under two different, yet related, respects.

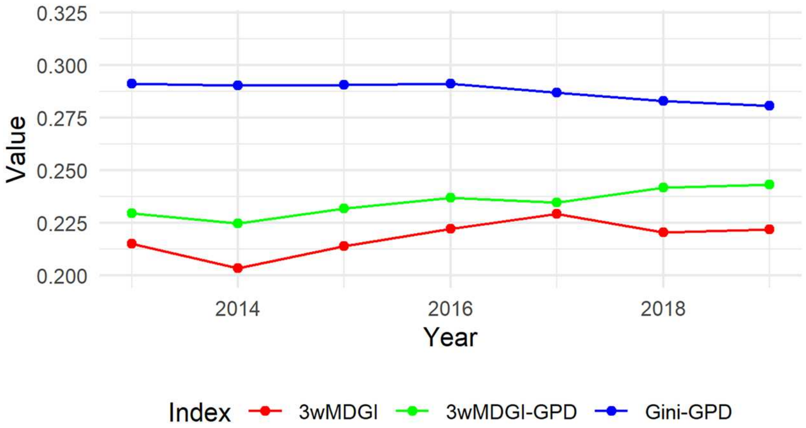

Figure 4. Multidimensional inequality Gini index plots. Comparison among Gini indices in Table 4 by varying $h \in \{2013, \ldots, 2019\}$. Red: 3wMDIGI. Green: 3wMDIGI-GDP. Blue: Gini-GDP. Source: authors' elaboration.

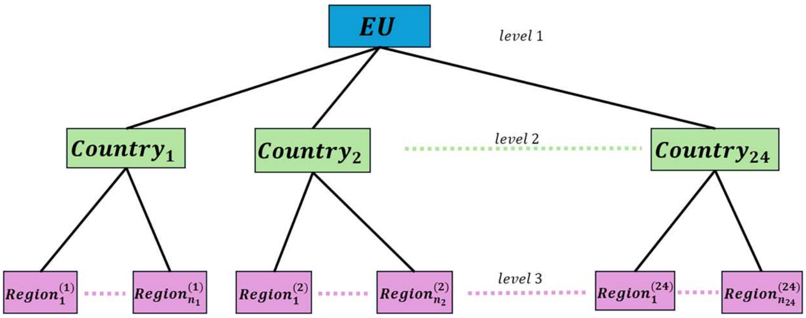

Figure 5. Graphical representation of EU-data structure to multilevel inequality scores. Source: authors' elaboration.

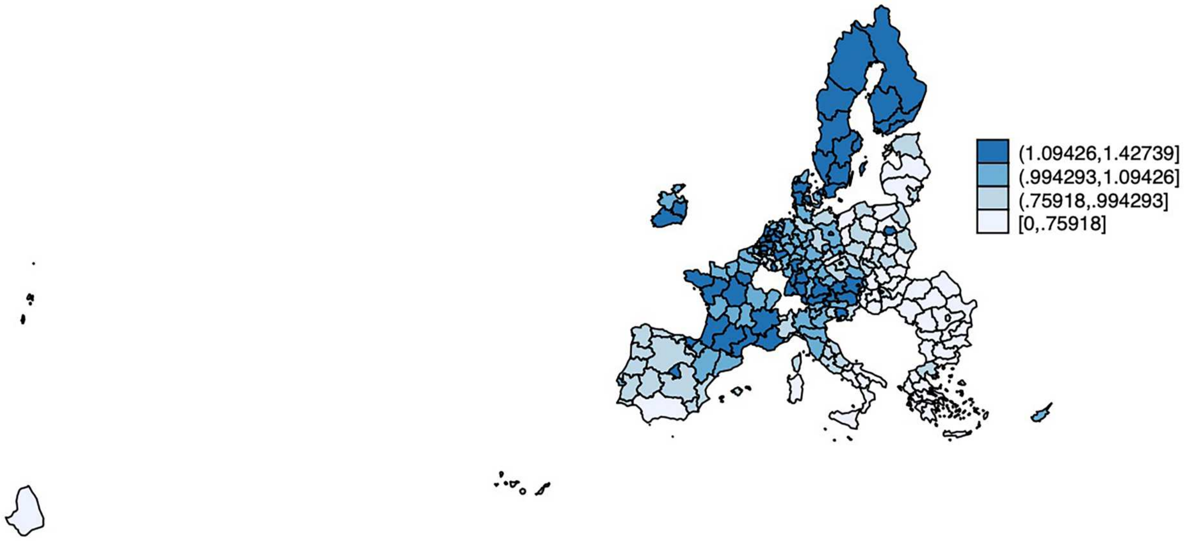

Figure 6. Global Gini inequality scores γ_i (10) for $i = 1, \ldots, N$. Source: authors' elaboration.

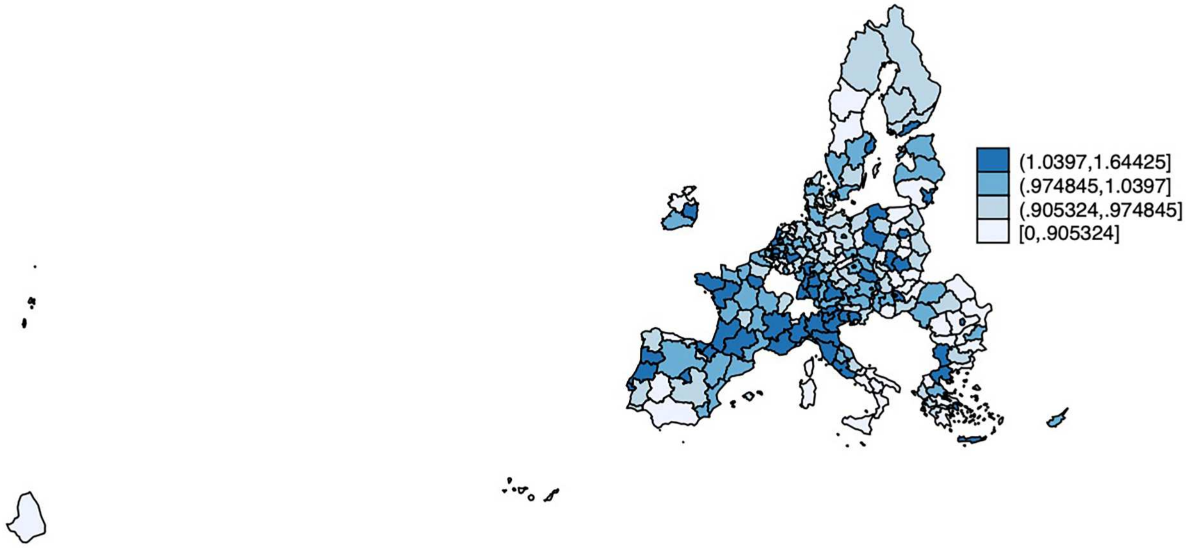

Figure 7. Within-countries multilevel inequality scores. Source: authors' elaboration.

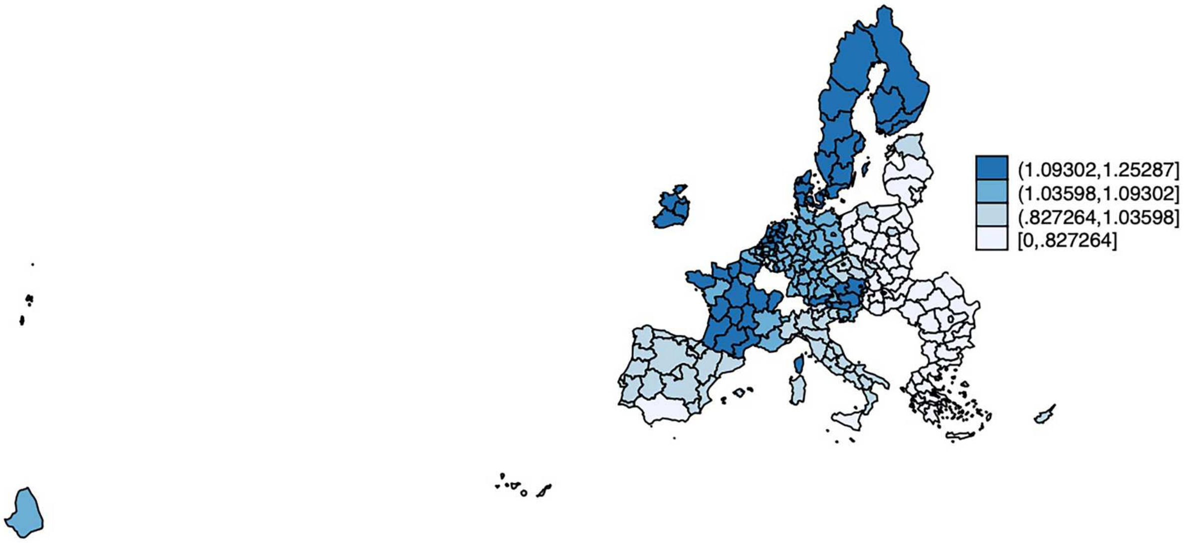

Figure 8. Between-regions multilevel inequality scores. Source: authors' elaboration.

First, the regional contribution to inequality is considered for the country to which each region belongs. Second, its contribution to inequality is considered relative to (i.e., in proportion to) the contribution to European regional inequality. In other words, Figure 7 plots the ratio between $\vartheta_{Region-Country}$ and $\vartheta_{Country-EU}$; where, as mentioned, the former measures the weight of the region in determining the overall observed inequality as compared to the remaining regions within the same country, and the latter measures the weight of the country as a whole in determining the overall regional inequality in Europe.

Following this interpretation, the scores reported in Figure 7 are labelled 'within-countries multilevel inequality scores' to highlight the nested nature of regional data *within* countries. Figure 6 shows the extent to which each region contributes to regional inequality in the EU, taking into account possible country-specific multilevel patterns of inequality. Hence, across Figures 6–8, darker colours signal a higher inequality.

Overall, a marked multilevel pattern of contribution to inequality emerges. Indeed, the picture is characterised by a clear centre-periphery multilevel pattern. While central regions show a relatively higher level of contribution to inequality, peripheral regions show a substantially lower level of contribution to inequality.

Given the multidimensional nature of the measure of inequality at hand, our exercise shows that multilevel disparities involve dimensions going well beyond the mainstream measures based on GDP and the like. Furthermore, within the more comprehensive framework characterised by multidimensional inequality, the pattern of European multilevel disparities is mainly due to central regions (i.e., higher relative weight symbol) and only to a substantially minor extent to peripheral regions.

As mentioned, Figure 8 performs a similar exercise with the difference being that the relative contribution to inequality of each region, in this case, is compared to the contribution to inequality that the region itself has for the EU. Hence, in the formula, we consider the ratio between $\vartheta_{Region-Country}$ and $\vartheta_{Region-EU}$; hence, this score aims to take into account the contribution to overall regional inequality *between* EU regions, i.e., considering EU regions themselves – not their countries – as a reference point. Also, according to this perspective, the regional contribution to inequality shows the aforementioned spatial pattern. Indeed, the centre-periphery spatial pattern is confirmed. However, it is worth noting how Sweden's regions have a more marked contribution to inequality in this case.

Overall, our finding aligns with previous analyses based on income alone, demonstrating the robustness of our multidimensional approach to the extent that the analysis uncovers a centre-periphery spatial pattern where peripheral regions exhibit a lower level of regional development with respect to core regions with substantially higher performance.

Nonetheless, as a novel contribution to the extant literature, it also shows a slightly different pattern when each region is directly compared to other EU regions. This evidence stresses the case for considering multidimensional inequality according to different spatial perspectives. Furthermore, from a policy point of view, it highlights the opportunity for a more granular approach to inequality to enhance policy efficiency and effectiveness. For example, the result indicating that within-country differences are the primary drivers of inequality highlights the importance of regional disparities within individual countries. This insight is particularly valuable for policymakers aiming to address inequality according to a multilevel governance framework.

5.1.1. Sensitivity analysis

To address concerns about the robustness of our results to the normative assumptions embedded in the choice of the parameters (β, δ), we performed a systematic sensitivity analysis. As discussed in Section 3, the admissible ranges are $\beta < 1$ and $1 \leq \delta \leq 2$ (Decancq & Lugo, 2012). Therefore, we computed the Global 3wMDIGI over a grid of parameter pairs:

$$\delta \in \{\, 1, 1.25, 1.5, 1.75, 2\} \,, \; \beta \in \{\, 0.001, 0.005, 0.1, 0.25, 0.5, 0.75, 0.999\}.$$

Expectedly, two patterns emerge (see Table 5): (i) for any fixed δ, increasing β lowers the Global 3wMDIGI, reflecting the higher degree of substitutability between well-being dimensions; (ii) for any fixed β, increasing δ slightly raises the Global 3wMDIGI, as more weight is given to the worst-performing regions.

The maximum is attained at ($\delta = 2$, $\beta = 0.005$), which was therefore chosen in our application as the optimal case. This data-driven choice allows the index to highlight disparities to the greatest extent, consistent with our empirical aim of revealing inequality patterns. In this case, the annual values of the 3wMDIGI range between 0.203 (2014) and 0.229 (2017), confirming a gradual convergence trend with only a minor reversal in 2018. Moreover, loadings analysis shows that GDP plays only a modest role compared to life expectancy and other social indicators, reinforcing the case for a multidimensional perspective. Nevertheless, to verify robustness, we re-estimated the 3wMDIGI under two alternative settings: ($\delta = 2$, $\beta = 0.25$) and ($\delta = 2$, $\beta = 0.75$). The index declines in both scenarios, which is consistent with our expectations. Crucially, however, the overall pattern persists:

- For ($\delta = 2$, $\beta = 0.25$), the 3wMDIGI values are:

$$0.1804, 0.1697, 0.1599, 0.1638, 0.1576, 0.1532, 0.1604,$$

  showing the same decreasing trajectory as the optimal case, albeit at slightly lower levels (around 10–15% smaller). The spatial maps of inequality scores (Figure 9) reproduce the centre–periphery divide observed in the optimal case (Figure 6), with core regions systematically contributing more to EU inequality. The relative role of GDP remains limited.

- For ($\delta = 2$, $\beta = 0.75$), the annual values are:

$$0.0703, 0.0548, 0.0461, 0.0305, 0.0465, 0.0326, 0.0426,$$

  which are substantially lower in magnitude due to the high degree of substitutability implied by large β. Yet, despite this rescaling, the main findings are unchanged: the downward trend persists, the centre–periphery structure is preserved (Figure 10), and GDP remains only a secondary contributor relative to health and education dimensions.

Comparing across specifications, one can conclude that: (i) the choice of (δ, β) affects the inequality (with higher β compressing disparities); (ii) however, the *ranking of regions, the spatial patterns, and the dimensional contributions* remain virtually unaffected; (iii) in all cases, a centre–periphery disparity

Table 5. Results of the global 3wMDIGI for different values of δ and β parameters.

	$\beta = 0.001$	$\beta = 0.005$	$\beta = 0.1$	$\beta = 0.25$	$\beta = 0.5$	$\beta = 0.75$	$\beta = 0.999$
$\delta = 1.00$	0.1989	0.1988	0.1743	0.1408	0.1078	0.0496	0.0044
$\delta = 1.25$	0.2004	0.2004	0.1782	0.1431	0.1045	0.0493	0.0041
$\delta = 1.50$	0.2019	0.2020	0.1814	0.1451	0.1010	0.0487	0.0035
$\delta = 1.75$	0.2034	0.2035	0.1834	0.1471	0.0977	0.0475	0.0022
$\delta = 2.00$	0.2044	**0.2045**	0.1827	0.1495	0.0954	0.0452	0.0002

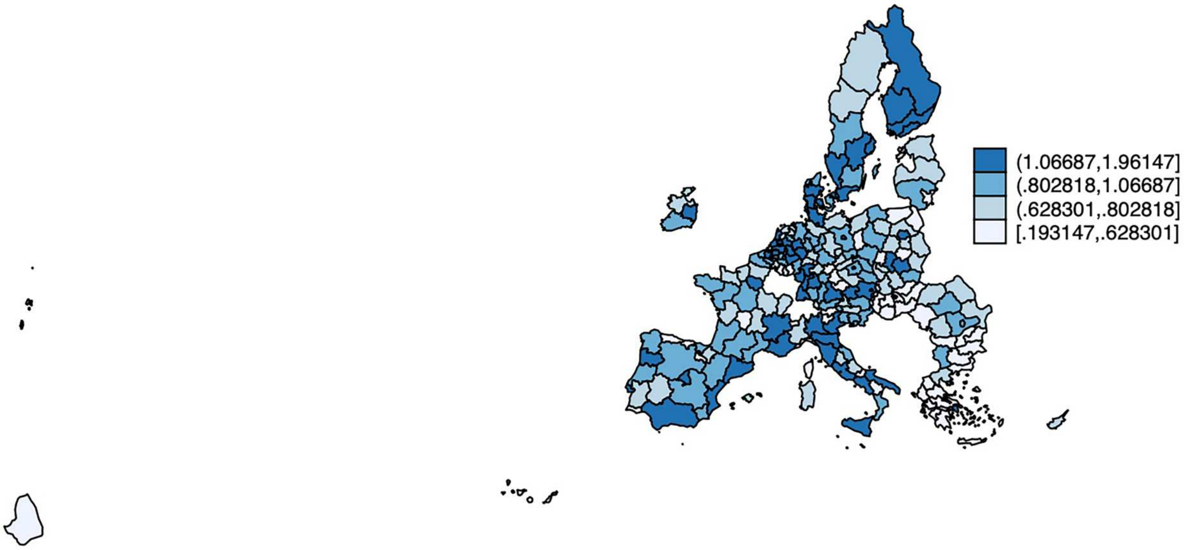

Figure 9. Global Gini inequality scores γ_i for $i = 1, \ldots, N$ and for $\beta = 0.25$ and $\delta = 2$. Source: authors' elaboration.

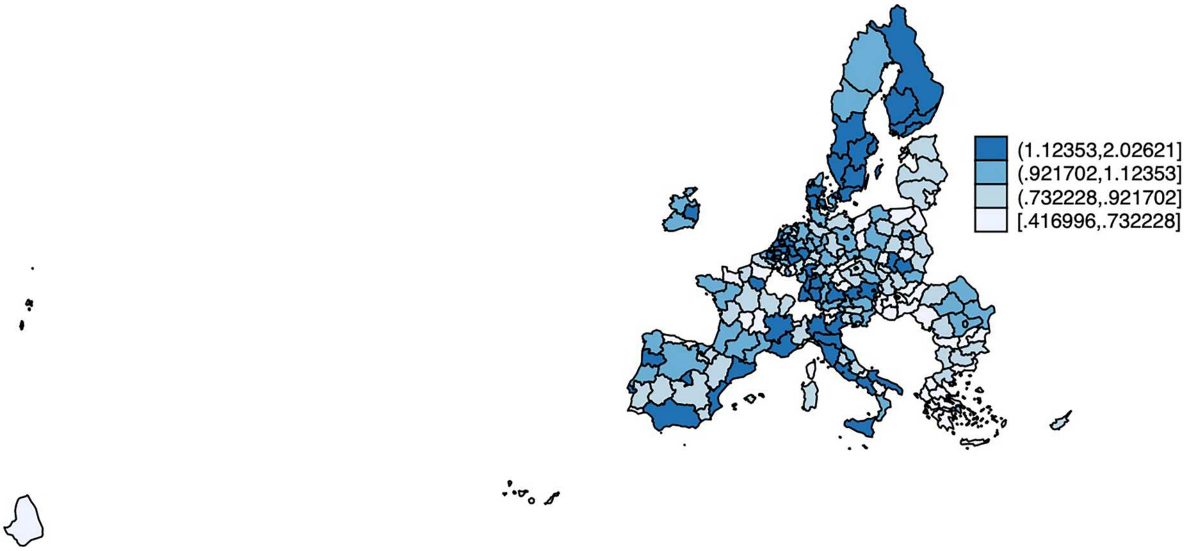

Figure 10. Global Gini inequality scores γ_i for $i = 1, \ldots, N$ and for $\beta = 0.75$ and $\delta = 2$. Source: authors' elaboration.

emerges, with central EU regions exerting stronger contributions to overall inequality; and (iv) GDP contributes less than health and education indicators regardless of parameter choice.

Overall, while the 3wMDIGI is sensitive in levels to (β, δ), the main empirical findings – namely, the centre-periphery pattern and the relatively limited contribution of GDP – prove robust across a broad range of admissible parameter values. This robustness strengthens the substantive validity of the proposed measure.

6. Concluding remarks

This study proposes a novel contribution to the measurement of inequality by developing the three-way multidimensional inequality Gini index (3wMDIGI), a synthetic indicator that captures the joint effects of multidimensionality, temporality and nested spatial structures. The index is based on a methodological innovation: the application of the Candecomp/Parafac (CP) model to estimate weights reflecting the relevance of each variable, time period and territorial level. This enhances the interpretability and sensitivity of inequality measures in complex data environments.

Applying the index to the European Union provides several substantive insights. Firstly, overall levels of multidimensional inequality in EU regions are comparable to those obtained from GDP-based measures alone. However, GDP's contribution within the broader multidimensional framework is notably lower than that of other well-being indicators, such as life expectancy, education and human resources in science and technology. This confirms that, while GDP remains informative, it is insufficient to capture the full complexity of inequality.

Secondly, the multilevel decomposition shows that the biggest share of inequality is due to inequalities within countries. This suggests that national-level heterogeneities, rather than differences between countries, are the primary drivers of inequality across the EU. In addition, the breakdown of regions shows a continuing and significant divide between the centre and the periphery. Central regions have a disproportionate share of the overall inequality, while peripheral areas have a much smaller share. This pattern is amplified when regional contributions are evaluated independently of national aggregates, underscoring the spatial granularity of the phenomenon.

The findings also point to a divergence in inequality trends during the observation period, with these trends showing different patterns over time. The 3wMDIGI signals are showing increasing disparities, which is in contrast to the more stable or even convergent patterns suggested by GDP-based indices, despite the relatively short time span. This highlights the importance of a multidimensional and temporally sensitive approach that can reveal dynamics which traditional methods might miss.

Importantly, a sensitivity analysis confirms the robustness of these findings to alternative parameter choices. While the absolute value of the 3wMDIGI depends on (δ, β) – with higher β compressing disparities – the key empirical insights remain stable. Specifically, the ranking of regions, the centre-periphery divide and the relative contributions of different dimensions are virtually unaffected by parameter variation; in all cases, central EU regions exert stronger contributions to inequality than peripheral areas; and GDP consistently plays a secondary role compared to health- and education-related indicators. This robustness strengthens the substantive validity of the proposed measure and confirms that the observed patterns are not artefacts of specific parameter assumptions.

Overall, the 3wMDIGI represents a significant conceptual and technical advancement in inequality measurement. Its flexibility makes it suitable for a range of applications beyond the EU, particularly in policy design contexts that require an understanding of inequality across multiple dimensions and territorial layers. These results suggest that policymakers should consider implementing more targeted interventions that focus on addressing disparities within nations, with a particular emphasis on understanding and addressing the spatial distribution of inequality across different regions.

Acknowledgements

This study was funded by the European Union – NextGenerationEU, in the framework of the GRINS – Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 – CUP E63C22002120006). The views and opinions expressed are solely those of the author and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was funded by the European Union – NextGenerationEU, in the framework of the GRINS -Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 – CUP E63C22002120006); EU Commission.

Data availability

The data that support the findings of the study reported in Section 5 are available from the corresponding author upon reasonable request and they were originally downloaded from the Eurostat database https://ec.europa.eu/eurostat/data/database.

Code availability

The custom code used to process the data, implement the methodology and generate the results presented in this paper is openly available in a GitHub repository at the following persistent link: https://github.com/CinziaDiNuzzo/3wMDIGI.

Notes

1. We acknowledge that repeated measurement over time could offer insights on the evolution of disparities over time. That said, it is worth noting that our proposed methodology; (i) does not require multiple computations, as (ii) it explicitly takes into account the evolution over time while computing the inequality measure(s) for each year in the sample.
2. https://www.oecd.org/regional/regional-policy/redefining-urban-9789264174108-en.htm
3. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=City_statistics_%E2%80%93_introduction#Policy_background
4. https://www.istat.it/it/archivio/220004
5. https://www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome/articles/subnationalstatisticsandanalysiscurrentandupcomingwork/october2023#progress-towards-our-ambitions
6. r_{ih} is a shorthand for $r(s_{ih})$, the rank of object i at occasion h on the basis of $\mathbf{s}_h = (s_{ih})$ in Equation (6), for $i = 1, \ldots, N$ and $h = 1, \ldots, H$.
7. For technical details on the normative space of compliance with the UR principle, the reader is referred to Decancq and Lugo (2012), especially pp. 731 and 736.
8. For a detailed treatment of these methodological aspects related to model parameters, the reader is referred to Decancq and Lugo (2012).
9. It is worth noting that this data transformation before applying the PCA, as mentioned, aims to reflect variables' polarisation in the subsequent PCA exercise in order to maintain their economic interpretation, as it cannot be ruled out that PCA would entail counterintuitive weights (for instance, weights with counterintuitive signs) on the variables (Banerjee, 2010; Morrison, 1976).
10. See Section 5.1.1 for details.

ORCID

Cinzia Di Nuzzo http://orcid.org/0000-0003-2062-8930
Venera Tomaselli http://orcid.org/0000-0002-2287-7343
Gianpiero Torrisi http://orcid.org/0000-0002-6389-5572

References

Adachi, K. (2020). *Matrix-based introduction to multivariate data analysis*. Springer Nature Singapore Pte Ltd.

Annoni, P., Dijkstra, L., & Gargano, N. (2017). *The EU regional competitiveness: Index 2016*. Publications Office of the European Union.

Arnold, B. C., & Sarabia, J. M. (2018). *Majorization and the Lorenz order with applications in applied mathematics and economics*. Springer.

Atkinson, A. B. (2015). *Inequality: What can be done?* Harvard University Press.

Banerjee, A. K. (2010). A multidimensional Gini index. *Mathematical Social Sciences*, *60*(2), 87–93.

Bernard, C., & Mller, A. (2020). Dependence uncertainty bounds for the energy score and the multivariate Gini mean difference. *Dependence Modeling*, *8*(1), 239–253.

Bocci, L., D'Urso, P., & Vicari, D. (2021a). A three-way approach for defining competitiveness indexes of the European regions (nuts-2). *Social Indicators Research*, *155*(3), 995–1043.

Bocci, L., D'Urso, P., & Vitale, V. (2021b). Clustering of the Italian regions based on their equitable and sustainable well-being indicators: A three-way approach. *Social Indicators Research*, *155*(3), 995–1043.

Bourguignon, F., & Chakravarty, S. R. (2003). The measurement of multidimensional poverty. *The Journal of Economic Inequality*, *1*(1), 25–49.

Bro, R. (1997). PARAFAC. Tutorial and applications. *Chemometrics and Intelligent Laboratory Systems*, *38*(2), 149–171.

Capaldo, M., & Navarro, J. (2025). New multivariate Ginis indices. *To Appear in Journal of Multivariate Analysis*, *206*, 105394.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "eckart-young" decomposition. *Psychometrika*, *35*(3), 283–319.

Cavicchia, C., Sarnacchiaro, P., Vichi, M., & Zaccaria, G. (2022, January 27–28). An ultrametric model to build a composite indicators system. In *Book of Short Papers. 10th International Conference IES 2022: Innovation and Society 5.0:*

Statistical and Economic Methodologies for Quality Assessment (p. 208211). Department of Economics, University of Campania.

Cavicchia, C., Sarnacchiaro, P., Vichi, M., & Zaccaria, G. (2024). A model-based ultrametric composite indicator for studying waste management in Italian municipalities. *Computational Statistics*, *39*(1), 21–50.

Cavicchia, C., & Vichi, M. (2021). Statistical model-based composite indicators for tracking coherent policy conclusions. *Social Indicators Research*, *156*(2), 449–479.

Decancq, K., & Lugo, M. A. (2012). Inequality of wellbeing: A multidimensional approach. *Economica*, *79*(316), 721–746.

Doran, J., & Jordan, D. (2013). Decomposing European nuts2 regional inequality from 1980 to 2009: National and European policy implications. *Journal of Economic Studies*, *40*(1), 22–38.

Durand, M. (2015). The OECD better life initiative: How's life? And the measurement of wellbeing. *Review of Income and Wealth*, *61*(1), 4–17.

Filauro, S. (2018). *The EU-wide income distribution: Inequality levels and decompositions*. Publications Office of the European Union.

Gavilan-Ruiz, J. M., Ruiz-Gndara, A., Ortega-Irizo, F. J., & Gonzalez-Abril, L. (2024). Some notes on the Gini index and new inequality measures: The nth Gini index. *Stats*, *7*(4), 1354–1365.

Greco, S., Ishizaka, A., Matarazzo, B., & 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., & Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, *141*(1), 61–94.

Greco, S., Tasiou, M., & Torrisi, G. (2021). A balanced development? The novel σ–μ efficiency of Italian regions. In S. Colombo (Ed.), *Spatial economics volume II: Applications,* (pp. 35–60). Springer.

Haataja, A. (1999). *Unemployment, employment and social exclusion* (No. 195). LIS Working Paper Series.

Harshman, R. (1970). Foundations of the parafac procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA Working Papers in Phonetics*, *16*(1), 1–84.

Jolliffe, I. (2002). *Principal component analysis*. Wiley Online Library.

Koshevoy, G. A., & Mosler, K. (1997). Multivariate Gini indices. *Journal of Multivariate Analysis*, *60*(2), 252–276.

Maggino, F. (2017). *Complexity in society: From indicators construction to their synthesis*. Springer.

McGillivray, M. (2007). Human well-being: Issues, concepts and measures. In M. McGillivray (Ed.), *Human well-being, studies in development economics and policy* (pp. 1–22). Palgrave Macmillan.

Morrison, D. F. (1976). Multivariate statistical methods-2.

Munda, G., & Nardo, M. (2009). Non-compensatory/non-linear composite indicators for ranking countries: A defensible setting. *Applied Economics*, *41*(12), 1513–1523.

Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005). Tools for composite indicators building. *Technical reports, European Commission. Institute for the Protection and Security of the Citizen . EUR 21682 EN. 2005. JRC31473.* https://publications.jrc.ec.europa.eu/repository/handle/JRC31473

OECD. (2008). *Handbook on constructing composite indicators. Methodology and user guide*. OECD Publications.

Piketty, T. (2014). *Capital in the twenty-first century*. Unfinished business: gender and inequality.

Pohlan, L. (2024). Unemployments long shadow: The persistent impact on social exclusion. *Journal for Labour Market Research*, *58*(1), 12.

Saisana, M., & Tarantola, S. (2002). State-of-the-art report on current methodologies and practices for composite indicator development. In *EUR 20408 EN.*

Simonacci, V., & Gallo, M. (2024). On four-way CP model estimation efficiency. *Computational Statistics*, *39*(1), 343–362.

Stiglitz, J. E., Sen, A., & Fitoussi, J. P. (2009). *The measurement of economic performance and social progress revisited, volume 33*. OFCE France.

Tomaselli, V., Fordellone, M., & Vichi, M. (2021). Building well-being composite indicator for micro-territorial areas through pls-sem and k-means approach. *Social Indicators Research*, *153*(2), 407–429.

Torrisi, G. (2021). A multidimensional analysis of the link between decentralization and development. In Ignacio Lago (Ed.), *Handbook on decentralization, devolution and the state - Chapter 16* (pp. 287–302). Edward Elgar Publishing.