UNIVERSITAT JAUME I

PROGRAMA DE DOCTORAT DE DESENVOLUPAMENT LOCAL I COOPERACIÓ INTERNACIONAL
ESCOLA DE DOCTORAT UNIVERSITAT JAUME I

COMPOSITE INDICATOR AS A DISTANCE

A Thesis submitted by **Eduardo Jiménez Fernández** for the degree of Doctor of Universitat Jaume I

Supervised by:
Luisa Alamá Sabater
Firmado por MARIA LUISA ALAMA SABATER -
and
Jose M. Calabuig Rodríguez
JOSE MANUEL| CALABUIG| RODRIGUEZ Firmado digitalmente por JOSE MANUEL| CALABUIG| RODRIGUEZ Fecha: 2022.10.17 17:43:21 +02'00'

Eduardo Jiménez Fernández
JIMENEZ FERNANDEZ EDUARDO - 40990333D Firmado digitalmente por JIMENEZ FERNANDEZ EDUARDO - 40990333D Fecha: 2022.10.17 15:52:54 +01'00'

Castelló de la Plana, Octubre 2022

"The author of this thesis did not receive any type of funding during the entire period of preparation of this work".

Publications

1. Jiménez-Fernández, E. & Ruiz-Martos, M.J. (2020). Review of some statistical methods for constructing composite indicators. *Estudios de Economía aplicada-Studies of applied economy, 38*(1), 1-15.
https://doi.org/10.25115/eea.v38i1.3002

2. Jiménez-Fernández, E., Sánchez, A. & Ortega-Pérez, M. (2022). Dealing with weighting scheme in composite indicators: an unsupervised distance-machine learning proposal for quantitative data. *Socio-Economic Planning Sciences, 101339*
https://doi.org/10.1016/j.seps.2022.101339

3. Sánchez, A. & Jiménez Fernández. E. (2022) EU Cohesion Policy: socio-economic vulnerability of the regions and the COVID-19 shock. *Applied Research in Quality of Life.* To be published.

"*This thesis has been accepted by the co-authors of the publications listed above that have waved the right to present them as a part of another PhD thesis*"

Agradecimientos

La realización de una tesis doctoral siempre es un reto que un estudiante afronta con ilusión, dedicación y mucho esmero. Un paréntesis de la vida que es enriquecedor y que deja una señal inolvidable en su forma de pensar y en el modo de entender la ciencia. Fruto de todo ese esfuerzo, se obtiene un trabajo que pretende aportar algo de luz a unas preguntas planteadas, tratando de resolverlas con menor o mayor fortuna. Sin embargo, no podemos dejar pasar, que dicho fruto no es solo el resultado de una tarea individual, más concretamente, es el resultado de las sinergias de varios investigadores/as, del apoyo de los compañeros/as, del soporte de las instituciones y por último, de la familia, capaz de soportar los inconvenientes, delirios y divagaciones de un estudiante. Por consiguiente, es justo y necesario, recordar y enumerar a todas esas personas o instituciones que han jugado un papel relevante en el desarrollo de esta tesis. En primer lugar, agradecer a mis tutores/as Lluisa Alamá y Jose M. Calabuig, por todo el apoyo ofrecido, consejos y ayuda incansable en la elaboración de este trabajo. A mis coautores/as, Angeles Sánchez, María J. Ruiz Martos, y Mario Ortega Pérez, por sus enormes aportaciones e inestimable báculo. A todos/as mis compañeros/as de la Universitat Jaume I, a los cuales estoy profundamente agradecido por brindarme unos años maravillosos de compañerismo y amistad. A la propia Institución, que gracias a sus políticas, instalaciones y actividades facilita el recorrido de los objetivos de sus profesionales y estudiantes. A mis padres, los cuales han intervenido de forma irreemplazable en toda mi formación académica. A Kiko, que me ha acompañado con sus profundos sueños en mis largas horas de estudio. Y por último, a mis dos Marías, mi hija en la cual me veo reflejado en sus avatares con el estudio y a mi esposa, la cual siempre ha sabido devolverme con una sonrisa y amor, los momentos más delicados de este arduo proyecto.

Eduardo Jiménez Fernández

Contents

Abstract

There is increasing interest in the construction of composite indicators to evaluate socio economic concepts. In general, the mathematical approaches on which the most commonly used techniques are based do not allow for reliable benchmarking. Moreover, the determination of the weighting scheme in composite indicators remains one of the most problematic issues. In this thesis, different methodologies are analyzed to extract their strengths and weaknesses. From this analysis it emerges that few of these tools allow comparison between observations. Using the vector space formed by all observations, a new method to construct composite indicators is proposed: a distance or metric that considers the concept of proximity between units. To do so, we take Pena Trapero's P2 Distance method as a starting point. The proposed methodology eliminates the linear dependence of the model and looks for functional relationships that allow us to build the most efficient model. This approach reduces the subjectivity of the researcher by assigning the weighting scheme with unsupervised machine learning techniques. Monte Carlo simulations confirm that the proposed methodology is robust. As an application of this new approach, a vulnerability index is constructed for the European Union (EU) regions. The cohesion policy is analyzed for the period 2021-2027 focuses on five objectives to make the EU smarter, greener, more connected, more social and closer to citizens. A macroeconomic index is proposed as the predominant criterion for allocating Structural Funds among regions. It is hypothesized that it is possible to take into account new complementary criteria that better reflect the quality of life of citizens. To this end, a composite index of socio-economic vulnerability is constructed for the 233 regions studied. The results show that, following our multidimensional approach for the allocation of Structural Funds, there are notable differences in the maps of priority regions. Moreover, the COVID-19 pandemic represents a welfare threat. Multilevel models are estimated from which it follows that country characteristics interact with those of regions to alter vulnerability patterns. More specifically, increased public spending on education and improved political stability would reduce regional vulnerability or build resilience, while increased poverty would be associated with increased vulnerability. Likewise, the most vulnerable regions would be the most exposed to the negative socio economic effects of COVID-19.

Resumen

Cada vez hay más interés en la construcción de indicadores compuestos para evaluar conceptos socio-económicos. En general, los enfoques matemáticos en los que se basan las técnicas más utilizadas no permiten realizar una evaluación comparativa de forma fiable entre las observaciones analizadas. Además, la determinación del esquema de ponderación en los indicadores compuestos sigue siendo una de las cuestiones más problemáticas. En esta tesis se analizan diferentes metodologías para extraer sus fortalezas y debilidades. De este análisis se desprende que pocas de estas herramientas permiten la comparación entre observaciones. Utilizando el espacio vectorial formado por todas las observaciones, se propone un nuevo método para construir indicadores compuestos: una distancia o métrica que considera el concepto de proximidad entre unidades. Para ello, se toma como punto de partida el método de la Distancia P2 de Pena Trapero. La metodología propuesta elimina la dependencia lineal del anterior modelo y busca relaciones funcionales que permitan construir un modelo más eficiente. Este enfoque reduce la subjetividad del investigador al asignar el esquema de ponderación con técnicas de aprendizaje automático no supervisado. Las simulaciones de Monte Carlo confirman que la metodología propuesta es robusta. Finalmente, como aplicación a este nuevo enfoque se construye un indice de vulnerabilidad en las regiones de Unión Europea (UE). Las políticas de cohesión para el periodo 2021-2027 se centran en cinco objetivos para hacer que la UE sea más inteligente, más ecológica, más conectada, más social y más cercana a los ciudadanos. Partiendo de la hipótesis de que es posible tener en cuenta nuevos criterios complementarios que reflejen mejor la calidad de vida de los ciudadanos, se construye un índice compuesto de vulnerabilidad socio-económica para las 233 regiones estudiadas como criterio predominante para asignar los Fondos Estructurales. Los resultados muestran que, siguiendo nuestro enfoque multidimensional existen notables diferencias en los mapas de regiones prioritarias. Adicionalmente, la pandemia de COVID-19 ha representado una amenaza para el bienestar. Se estiman modelos multinivel de los cuales se desprenden que las características de los países interactúan con las de las regiones para alterar los patrones de vulnerabilidad. Más concretamente, el aumento del gasto público en educación y la mejora de la estabilidad política en el nivel país reducirían la vulnerabilidad regional o fomentarían la capacidad de resiliencia, mientras que el aumento de la pobreza se asociaría a una mayor vulnerabilidad. Asimismo, las regiones más vulnerables serían las más expuestas a los efectos socio-económicos negativos del COVID-19.

Chapter 1

Introduction

The use of composite indicators is becoming increasingly widespread in the social sciences. Most of the phenomena studied within this framework are multidimensional, e.g., development, poverty, welfare (Maggino, 2017), therefore, synthesizing these into a single indicator has favored the emergence of methodologies that offer different approaches to address this complexity (Greco et al. 2019). However, the construction of a composite indicator goes beyond the purely mathematical operation involved in reducing the dimensionality of the data (Mazziotta and Pareto, 2018). Such a construction should follow a respectful methodological approach to ensure that the overall picture captures fundamentally what is intended (OECD, 2008). The methodological process starts with the precise definition of the conceptual framework (a defined measurement process, Maggino, 2017, p.87), which conditions the selection of single indicators that (attempt to) measure the different dimensions of the concept and the aggregation method - differential weighting allowed - of the resulting indicator system, and ends with the robustness analysis of the composite indicator. This measurement process inevitably involves some subjective choices whose consequences must be clearly stated by the researcher (Maggino, 2017, p.89). There are different aggregation approaches to construct composite indicators. We can distinguish between compensatory and non-compensatory methods. This refers to the possibility that low values of a single indicator may or may not be offset by high values of another indicator. The appropriateness of the (degree of) compensability of the aggregation technique depends on the conceptual framework . Examples of compensatory methods are linear and geometric aggregation (e.g., Saisana and Tarantola, 2002; Bandura, 2008, 2011; Greco et al. 2019). Examples of non-compensatory techniques are the Electre and Promethee methods. The drawback of non-compensatory approaches is their computational complexity, which minimizes their popularity (Greco et al., 2019).

There are very different methodologies whose purpose is to construct a good indicator that optimally summarizes the problem to be addressed, its applicability being subject to the type of phenomenon under study. Some popular aggregation methods for constructing composite indicators, e.g., of human well-being, are characterized by computing weights based on statistical methods (Decanq and Lugo, 2013, p. 19 in Greco et al., 2019). For instance, Data Envelopment Analysis (DEA), P2 Distance (DP2), Mazziotta-Pareto Index (MPI) or Principal Component Analysis (PCA) have been widely used (e.g. Saisana and Tarantola, 2002; Bandura, 2008; Somarriba and Peña, 2009; Greyling and Tregenna, 2016; Yang et al. 2017; Sanchez and Ruiz-Martos, 2018; for a more comprehensive review see Greco et al. 2019). The approaches used in each case to compute the weights are inherently different, resulting in very different measures and making each appropriate for a specific measurement exercise. The selection of the methodology for the construction of a composite index requires a refinement of the conceptual framework to define the ultimate objective of the measurement exercise. In other words, it is not enough to state the objective multidimensional concept, e.g., well-being. It is necessary to establish how exactly we intend to measure it. For instance, if the aim of the research is to produce a ranking of observations (countries, regions, etc.) with respect to, for example, welfare, then PCA and DP2 should be applied, with preference for the latter (see Mazziota and Pareto, 2019 and below). If the objective of the research is, however, to determine which dimension(s) (or individual indicator(s)) is/are more efficient in maximizing the welfare of each observation (e.g., in which dimensions of welfare is each country more efficient in addressing public policies), then DEA and MPI type methodologies should be applied.

The first chapter of this thesis reviews the literature and analyzes four aggregation methodologies of very different nature. This chapter corresponds to the work published in the journal Studies of Applied Economic (Jiménez-Fernández & Ruiz-Martos, 2020). The P2 distance, principal components, data envelopment analysis and the Mazziotta-Pareto index are analyzed. P_2 Distance, defined by Pena Trapero (1977), is an iterative method to obtain a metric (distance of each observation to a reference vector) by aggregating several individual indicators as a weighted sum. Principal component analysis (PCA) is a mathematical tool whereby an orthogonal transformation of the reference system of the set of observations transforms the set of indicators of possibly correlated variables into a set of linearly uncorrelated variables. On the other hand, Data Envelopment Analysis (DEA) is a methodology originally related to Management Science used to analyse the technical-efficiency of public-sector decision-making units (DMU's). This methodology optimizes for each individual observation that one may compute a discrete piecewise frontier through the set of Pareto-efficient Decision Making Units (DMU)(Charnes et. al. 1979). Lastly, the Mazziotta-Pareto index (MPI) is a composite index, (OECD, 2008[2]) whose purpose is to group a set of individual indicators that are not fully substitutable[3]. Starting from a normalization, the indicators are aggregated to a non-linear function to which a penalty is introduced for units with unbalanced values of the indicators (De Muro et al., 2011[4]). This chapter analyzes the weaknesses and strengths of these four indicators and provides a first approximation of which methodologies are most appropriate depending on the study to be carried out. The PhD student reviews together with co-author María Jose Ruiz Martos. The experiments and the conclusions derived from them are also contributed by the student where R software was used for computations, graphical developments and algorithms. In addition, he also participates in the drafting, proofreading and submission of the article.

As a result of the above analysis, Chapter 2 proposes a new methodology. This chapter corresponds to the work published in the journal Socio-Economic Planning Sciences, Jiménez-Fernández et al. (2022). Starting from the P2 Distance, the weaknesses of this method are improved, incorporating machine learning for weight computation and correcting some technical aspects that did not satisfy the classical methodology. Using the vector space (each vector is formed by the coordinates resulting from each simple indicator), a distance or metric is constructed that considers the concept of proximity between units as a tool to synthesize the indicators. This approach enables comparisons between the units being studied, which are always quantitative. The proposed methodology eliminates the linear dependence on the model and seeks functional relationships that enable constructing the most efficient model. This approach reduces researcher subjectivity by assigning the weighting scheme with unsupervised machine learning techniques. Monte Carlo simulations confirm that the proposed methodology is robust. The PhD student is the architect of the algorithm and the relevant mathematical developments to address the issues and challenges presented by the proposed problem. Again, R software is used for computation, graphics development and algorithms. In addition, the student also participates in writing, drafting, proofreading of the article.

Finally, as an example of the application of the above methodology, Chapter 3 hypothesizes that it is possible to take into account new complementary criteria that better reflect the quality of life of citizens. To that end, a composite index of socio-economic vulnerability for the 233 NUTS2 european regions is constructed. The results show that following our multidimensional approach for allocating the Structural Funds, there are remarkable differences in the maps of priority regions. In addition, the COVID-19 pandemic represents a threat to well-being. Are all regions equally exposed to COVID-10 in terms of their socio-economic vulnerability? To address this issue, we estimate multilevel models which indicate that country characteristics interact with regions? characteristics to alter patterns of vulnerability. More specifically, increases in government expenditures in education and an improvement in political stability would reduce the regional vulnerability or foster the capacity for resilience, whereas increases in poverty would be associated with greater vulnerability. Likewise, more vulnerable regions would be the most exposed to the negative socio-economic effects of COVID-19. However, it is remarkable that several regions of Sweden and Finland would be among the group of regions whose socio-economic vulnerability would be the most negatively affected. The paper corresponding to this chapter is accepted and pending publication in the journal Applied Research in Quality of life. The PhD student develops the methodological part related to the construction of the composite indicator and participates in the multilevel analysis to obtain conclusions. In addition, he also participates in the drafting, proofreading of the article.

The thesis ends with conclusions, where the findings or discoveries of the study are addressed, highlighting new information, advantages and disadvantages, future lines of research, and the bibliography provided for each chapter.

Chapter 2

Review composite indicator methodologies

Abstract

The methodology for the construction process of composite indicators is reviewed in a step-by-step approach ranging from the ex-ante definition of the latent variable that is intended to be measure, through the construction process of the composite indicators. We focus particularly on four aggregations methods in order analayze weighting and aggregation approach, P2 Distance, Principal Component Analysis, Data Envelopment Analysis and Mazziotta-Pareto Index. An empirical comparison among them is provided and the composite indices divergences are shown.

2.1 Introduction

In social sciences, the use of indicators is ever spreading. Indicators, single and composite, aim to measure some concept or latent variable. Most socioeconomic phenomena are multidimensional, which renders a single indicator unable to capture the inherent complexity in, for example, development, poverty, well-being (Maggino, 2017; Greco et al. 2019), and favours a multi-indicator approach. Composite indicators, which synthesize the information conveyed by an usually wide range of indicators, constitute a popular alternative. Constructing a composite indicator, however, goes beyond the purely mathematical operation involved in reducing data dimensionality (Mazziotta and Pareto, 2018). The construction of composite indicators should follow a respectful methodological approach to ensure that the *big picture* fundamentally captures what it is meant to OECD, 2008). The methodological process to construct a composite indicator starts with the precise definition of the conceptual framework (a *defined process* of measurement, Maggino, 2017, p.87), which conditions the selection of single indicators that (attempt to) measure the various dimensions of the concept and the appropriate aggregation method -differential weighting allowed- of the resulting system of indicators, and finishes with the robustness analysis of the composite indicator. This measurement process inevitably involves some subjective choices whose consequences should be clearly stated by the researcher (Maggino, 2017, p.89).

This paper reviews the methodological steps[1] that ought to be followed in the construction

[1] For a more thorough methodological discussion see, for instance, OCDE 2008 and Maggino, 2017.

process of a composite indicator, and some of the most popular methodologies used to construct composite indicators of human well-being. In particular, we discuss four aggregation methods that elicit data-driven weights (Decanq and Lugo, 2013, p. 19 in Greco et al., 2019): Data Envelopment Analysis (DEA), P_2 Distance (DP2), Mazziotta-Pareto Index (MPI) and Principal Components Analysis (PCA)[2]. We focus on these aggregation methodologies because, first, they are widely used (for instance: Saisana and Tarantola, 2002; Bandura, 2008; Somarriba and Pena, 2009; Greyling and Tregenna, 2016; Yang et al. 2017; Sanchez and Ruiz-Martos, 2018; for a more thorough survey see Greco et al. 2019). Secondly, despite all being data-driven techniques, their approaches to the computation of weights are intrinsically different, which results in severely dissimilar measures and makes each one of them appropiate for a specific measurement exercise. We review the desired properties of an aggregation method and discuss the properties verified by the four methodologies. Finally, we compare these methods with respect to their weighting schemes; and the consequences of eliminating observations (countries, regions, etc) and adding noise (introducing an indicator which is a lineal combination of the other indicators).

Our main conclusion is that the selection among these aggregation methods requires a refinement of the conceptual framework so as to define the ultimate purpose of the measurement process. That is, it does not suffice to state that, e.g., well-being is the multdimensional latent variable of interest. It is also necessary to establish how exactly we aim to measure it. One may be interested in comparing observations regarding well-being or, alternatively, in selecting a particular dimension for each observation to maximize its well-being. In particular, if the research goal is to produce a ranking of observations (countries, regions, etc) in terms of the latent variable (e.g. well-being) then PCA and DP2 should be applied, with a preference for the latter one (see Mazziota and Pareto, 2019 and discussion below). If the research goal is, however, to determine which dimension/s (or individual indicator/s) is/are more efficient to maximize the latent variable for each observation (e.g. in which dimensions of well-being each country is more efficient so as to address public policies), then DEA type methodologies and MPI should be applied.

Next section defines the four methodologies we compare. We discuss in section 3 the methodological steps required to construct a composite indicator and why human well-being requires a very specific model of measurement. Section 4 reviews the desired properties of an aggregation method. Section 5 compares the methodologies and Conclusions follow.

2.2 Some aggregation methods

The measurement of multidimensional socio-economic phenomena has been developed in last decades. In this paper, we focus on several methodologies in order to analyze the weakness and robustness of each of them. Data Envelopment Anlalysis (DEA), Distance P_2 (DP2), Principal Components Analysis (PCA) and Mazziotta-Pareto Index (MPI) are analysed. Notice that there is not universal method for composite indices construction, therefore, depending on which phenomena is measured and on how it is measured, a methodology is more suitable than other. To carry out this analysis, we check a set of properties for each of the above methods. Throughout the section, sub-index i will correspond to an observation (region, country, etc.) and sub-index j to a simple indicator. First, we will provide a brief description of each of them.

[2]Greco et al. (2019) do not discuss the DP2 methodology.

2.2.1 Data Envelopment Anlalysis (DEA)

Data Envelopment Analysis (DEA) is a methodology originally related to Management Science used to analyse the technical-efficiency of public-sector decision-making units (DMUs). This methodology optimizes for each individual observation that one may compute a discrete piece-wise frontier through the set of Pareto-efficient Decision Making Units (DMU) (Charnes, A. et. al. 1979). One of the virtues of this methodology with respect to other parametric approaches is that it does not require specific assumptions on the distribution of the error terms. In addition, a crucial feature of DEA is that the weights assigned to well-being domains are endogenously generated at the observation level.

DEA gathers a set of methodologies for evaluating performance. We focus on the so-called DEA-BoD approach, where BoD is the abbreviation of Benefit-of-the-Doubt (BoD) principle (Cherchye et al. 2007). The basic idea of the DEA-BoD approach is to impose on each observation under evaluation the optimal set of weightings such that the observation achieves the best relative position with respect to all other observations. As a example, Mariano et al. (2015) provide a literature review on the research using DEA to measure and analyse human development. Another approaches combine DEA and Multi-Criteria-Decision-Making (MCDM) techniques to improve DEA-BoD while retaining a structure scheme of weightings for well-being domains across observations (Despotis, 2005; Peiro & Picazo 2018).

We concentrate on the additive model DEA-BOD, which has been used to compute well-being and quality of life composite indicators, for instance, González et al. (2010), Reig-Martínez (2012), Mizobuchi (2014). We suppose that X is a $n \times m$ dimension matrix, where each column represents a single indicator. Let y_i be the composite index associated to the i-observation. For each $i \in \{1,...,n\}$, the Additive model DEA consists on maximising n DMU problems as follows:

$$Maximise_{\alpha_{ij}} \quad y_i = \sum_{j=1}^{m} \alpha_{ij} x_{ij}$$

$$\textit{Subject to:}$$

$$\sum_{j=1}^{m} \alpha_{ij} x_{\ell j} \leq 1 \; \textit{for all} \; \ell \in \{1,..,n\}$$

$$\alpha_{ij} \geq 0 \quad j \in \{1,...m\}.$$

The DEA computations maximize the relative efficiency score of each DMU, where the constraint condition is that the set of weights so obtained for each DMU must also be feasible for all the other DMUs included in the optimization. Then, the main difference with respect to others parametric approaches is that the analysis is based on individual observations and not on population estimations. Such a strategy provides a single aggregate measure for each observation (DMU) through the input factors (single indicators), i.e., produces its respective composite indicator. A priori, the optimization procedure does not require a suitable specification of weights for each single indicator. As stated above, it is not necessary to impose the functional relationship between the composite index and the set of single indicators that define it, (Charnes et al. 1997).

2.2.2 P2 Distance

P2 Distance, defined by Pena Trapero (1977), is an iterative method to obtain a metric (the composite indicator) by aggregating various single indicators as a weighted sum. Let X be a $n \times m$-dimension matrix, in which n is the number of observations and m is the number of single indicators. We define the reference vector $X_* = (x_{*1}, ..., x_{*m})$ as a fictitious vector that his coordinates are composed by the results of a theoretical observation with the best-worst possible scenario for all the single indicators. For each observation we define $d_{ij} = |x_{ij} - x_{i*}|$ as the distance from the i-observation $\imath \in \{1, ..., n\}$ to the j-coordinate of the reference vector $\jmath \in \{1, ..., m\}$. For instance, if the composite indicator is measuring the regional development, the composite indicator measures the distance between each region and a fictitious reference. Thus, the reference vector summarizes the results of a fictitious region with the worst possible scenario for all the indicators, for more clarity see Sánchez & Ruiz (2018).
The Frechet Distance (DF) corresponding to the i-observation is defined as follows

$$DF_i = \sum_{j=1}^{m} \frac{d_{ij}}{\sigma_j} = \sum_{j=1}^{m} \frac{|x_{ij} - x_{i*}|}{\sigma_j} \tag{2.1}$$

where σ_j is the standard deviation of the j-single indicator $j \in \{1, ..., m\}$. Regarding the weights, this method allows several options. First, the researcher can assign to each element of the sum, i.e. to $|x_{ij} - x_{i*}|/\sigma_j$, a weight according to the relative importance of each indicator (e.g., experts opinion, political agreement). Secondly, we can assume the DF distance so that we are implicitly assuming that all indicators have the same weight. These options are to be considered in those cases where statistical relationships among single indicators do not represent the actual influence among them (Saisana and Tarantola 2002). However, when there is no information about nor agreement on the importance of the indicators, Distance P_2 provides an iterative method to assign weights to the model based on the linear correlations between indicators. This strength is also its weakness, as it depends solely on the linear relationships that may exist. The Distance P_2 is defined as follow:

$$DP2_i = \sum_{j=1}^{m} \frac{d_{ij}}{\sigma_j} = \sum_{j=1}^{m} \frac{|x_{ij} - x_{i*}|}{\sigma_j}(1 - R^2_{j,...,1}) \tag{2.2}$$

where for each $j \in \{1, ..., m\}$, $R^2_{j,...1}$ represents the coefficient of determination in the multiple linear regression of x_j over $x_{j-1}, ..., x_1$ assuming $R_1^2 = 0$. The weight $(1 - R^2_{j,...,1})$ deletes the information contained in the preceding indicators and, thus, avoids the duplication of information. This property will be called as "completeness". This method is sensitive to the order in which the indicators are introduced. To avoid subjectivity of choice, the pairwise correlation coefficients between each indicator and DF are assessed, and then the indicators are sorted from highest to lowest according to the absolute values of the pairwise correlation coefficients. This property is called "neutrality" by Zarzosa Espina (1996). The indicators are introduced in the model following the previous order and the weights calculated following this criterion. The process continues iteratively until the difference between two averages adjacent DP2s is zero. As example of this methodology see Sánchez & Ruiz (2018).

2.2.3 Principal Components Analysis

Principal component analysis (PCA) is a mathematical tool whereby an orthogonal transformation of the reference system of the set of observations transforms the set of indicators of possibly correlated variables into a set of linearly uncorrelated variables. Its origin can be imputed to Pearson (1901) or even Cauchy (1829). Let X_j be the j-single indicator $j \in \{1,...,m\}$ and let $p \leq m$ be the number of principal components $Y_1,...,Y_p$ that they are obtained as linear combinations of the original data $X_1,...,X_m$.

$$\begin{aligned} Y_1 &= a_{11}X_1 + ... + a_{1m}X_m \\ \vdots &= \quad \vdots \\ Y_p &= a_{p1}X_1 + ... + a_{pm}X_m \end{aligned}$$

where the factor loadings $A_\ell = (a_{\ell 1},...,a_{\ell m})$ satisfy that

$$\sum_{j=1}^{m} a_{\ell j} = 1, \; \textit{for all } \ell \in \{1,...,p\} \tag{2.3}$$

and $Y_1,...,Y_p$ is a orthogonal set of vectors (i.e., uncorrelated). The goal of this method is to maximize the variance

$$\begin{aligned} &Var(Y_1) = Var(A_1X) = A_1'\Sigma A_1 \\ &\textit{Subject to:} \\ &A_1'A_1 = 1 \end{aligned}$$

where Σ is the covariance matrix of the data set $X_1,...,X_m$. Let $\lambda_1 \geq ... \geq \lambda_m$ denote the set of eigenvalues of the covariance matrix Σ. Using Lagrange Multiplier approach and Roché-Frobenius theorem $Var(Y_1) = \lambda_1$. Therefore, the maximum eigenvalue provides the maximum variance, and the corresponding eigenvector whose coordinates are the factor loadings. The second component Y_2 is computed solving the previous optimization problem but also by imposing $Y_1'Y_2 = 0$. The process continues iteratively until all the components are computed. Karamizadeh et al. (2009) provide a recent description of this methodology. A detailed review of the literature and applicability of this approach is shown in Greco et al. (2019, p.71).

2.2.4 Mazziotta-Pareto Index

Let x_{ij} be the ith-observation corresponding on the jth-indicator. Min-max method is used as normalization method through the following formulation:

- If the polarity of the j-indicator is positive:

$$y_{ij} = \frac{x_{ij} - min(X_j)}{max(X_j) - min(X_j)} \tag{2.4}$$

- If the polarity of the j-indicator is negative:

$$y_{ij} = \frac{max(X_j) - x_{ij}}{max(X_j) - min(X_j)} \tag{2.5}$$

The Mazziotta-Pareto Index (MPI) is a non-linear function composite index (Maggino, F. 2017b). After Min-Max normalization, we define $z_{ij} = 100 + 10(y_{ij} - M(Y_j))/\sigma(Y_j)$, where the $M(Y_j))$ is the mean of the j-Min-Max normalized indicator and $\sigma(Y_j))$ is the standard deviation of the j-Min-Max normalized indicator. The MPI index is defined as follows:

$$MPI^{\pm}_{z_i} = M_{z_i} \pm \sigma_{z_i} cv_{z_i} = \frac{1}{m}\sum_{j=1}^{m} z_{ij} \pm \sigma_{z_i} cv_{z_i} \quad i \in \{1,...,n\} \tag{2.6}$$

where M_{z_i}, σ_{z_i} and $cv_{z_i} = \sigma^2_{z_i}/M_{z_i}$ denote the mean, standard deviation and the coefficient variation of each i-th observation. The sign $\pm$ is related with the phenomenon to be measured. Penalty direction is positive in case of increasing composite indicator (for instance, well-being index)). On the other hand, penalty direction is negative in case of decreasing composite index (for example, poverty index).

2.3 Conceptual Framework

Measuring in social sciences requires a robust conceptual definition of the target of measurement (latent variable), a consistent collection of observations and a subsequent analysis of the relationship between observations and defined concepts (Maggino, 2017, p. 87). This relationship between variables and indicators (Maggino, 2017, p. 94) determines the model of measurement and conditions the construction process of the composite indicator, in particular the appropriate aggregation method (Maggino, 2017, p.97).

The model of measurement may be reflective or formative (Maggino, 2017; Mazziota and Pareto, 2018). In a reflective model, indicators are functions of the latent variable, i.e., the latent variable is the independent variable (changes in the latent variable trigger changes in the indicators). In a formative model, the latent variable depends on the indicators, i.e., the indicators are the independent variables (changes in the indicators imply changes in the latent variable). In the following, let R represent the multidimensional latent variable we aim to measure and X_j, individual indicator j.

2.3.1 Reflective model

In a reflective model, indicators are manifestations of the latent variable R. Hence, causality is from the concept to the indicators. The latent variable R represents the common cause shared by all indicators X_j that reflect the concept, with each indicator corresponding to a linear function of the underlying variable R plus a measurement error:

$$X_j = \lambda_j R + \epsilon_j$$

where λ_j is the coefficient or loading that captures the effect of R on X_j and ϵ_j is the measurement error of that indicator.

Measurement errors are assumed to be independent and unrelated to the latent variable. In reflective models, individual indicators are: interchangeable (removing one of the indicators does not affect essentially the latent variable); intercorrelated (two uncorrelated indicators cannot be caused by the same latent variable); and, moreover, positively correlated if they share equal polarities (i.e., they are equally related to the latent variable), conversely, negatively correlated (Maggino, 2017; Mazziota and Pareto, 2018).

The key for a reflective model being that single indicators must be highly correlated (Maggino, 2017, p.121). The right approach to this conceptual framework is to reduce dimensionality by a factor or scaling model, i.e., factor analysis (its main goal is to test a reflective approach and allows to synthesize indicators belonging to the same dimension or latent variable) and PCA (its main goal is to summarize the whole variance of the data by fewer variables, called components, than the original indicators) (Maggino, 2017, Mazziota and Pareto, 2019). For example, measuring intelligence through a questionnaire, the more intelligence the more correct answers in all dimensions (Simonetto, 2012 in Mazziota and Pareto, 2019, p. 454).

2.3.2 Formative model

In a formative model, indicators cause the latent variable and, thus, a change in the latent variable does not necessarily imply changes in all its measures (Mazziotta and Pareto, 2018). The concept is defined by the indicators.

$$R = \sum_j \lambda_j X_j + \zeta$$

where λ_j captures the effect of indicator $X_j$on the latent variable R, and ζ is the error term.

Indicators are not interchangeable (omitting one of the indicators is omitting a part of the latent variable), and correlations between indicators are not explained by the measurement model (high correlations are possible but not generally expected and will cause a multicollinearity problem). In fact, correlated indicators may be redundant and make the conceptual component measured by both to get more weight in the composite indicator (Maggino, 2017). Moreover, correlations and polarities are independent. Since such a model does not assume that indicators are correlated, the correlation structure of the data cannot be used to determine the latent variable. Instead, the latent variable is estimated by taking a weighted average of the indicators that conform the concept (Shwartz et al. 2015 in Mazziota and Pareto, 2019, p. 4). DEA-BoD, Distance P_2 and MPI are examples of formative models.

The measurement of human well-being requires a formative model (Diamantopolus et al. 2008). Well-being depends on health, income, occupation, services, safety, environment, etc. So the improvement of any one of these indicators would imply an improvement in well-being, even if the other indicators remain invariant. Subsequently, an improvement of well-being does not necessarily imply and improvement in all its indicators. When the composite indicator follows a formative model, the following issues are critical (Diamantopolus and Winklhofer, 2001 in Maggino, 2017, p. 120):

- Content specification refers to the content domain that the composite indicator aims to capture and is inextricably linked to the specification of the indicators.
- Indicator specification: indicators must cover the entire latent variable domain. Neither is good an excessive number of indicators nor the exclusion of a neccessary indicator.
- Indicator collinearity. Excessive collinearity among indicators is problematic as it makes difficult to disentangle particular influences of individual indicators in the latent variable. Multicollinear indicators may be redundant and one of them may be eliminated.

- External validity. The composite indicator should be related to other measures, this is accomplished by individual indicators being related to other indicators external to the composite indicator.

2.3.3 Selection of indicators

The selection of indicators must address the complexity of the targeted phenomena (multidimensionality, nature -objective versus subjective, quantitative versus qualitative-, distinct level of observations -micro and macro-, dynamics -internal and external conditions, trends and relationships between phenomena), allow for relativity and comparability (e.g., the same concept may be measured by different indicators in different areas; careful interpretation of the results) and avoid overreductionism (the system of indicators may be simplified either by reducing the number of indicators following the conceptual model or by synthesizing indicators into a composite indicator) (Maggino, 2017).

Indicators have diverse measurement units and ranges, for which, before aggregation, they should be made comparable. This is achieved by normalization, i.e., by transforming indicators into pure, dimensionless, numbers. Several normalization techniques are available: ranking, standardization, re-scaling or indicization. Each of these techniques has pros and cons mainly regarding the interval level of information, sensitivity to outliers and implicit weighting (Mazziota and Pareto, 2017, p.170). The selection of a particular one should be justified attending to the conceptual framework, indicators variability among observations and the ultimate purpose of the measurement exercise.

Moreover, polarity of indicators must be defined (the sign of the relationship between the indicator and the latent variable, positive or negative). Indicators with negative polarity must be "inverted" by a linear or non-linear transformation (Mazziota and Pareto, 2017). The aim of normalizing and dealing with polarity is that an increase in the normalized indicator implies an increase in the composite index (Salman, 2003; Mazziota and Pareto, 2017, p.166).

Furthermore, as socio-economic data are mainly ordinal and discrete, it should be clearly stated how ordinal indicators are dealt with, i.e., how metric analysis is carried out of non-metric data because this may be not be consistent with the true nature of the targeted phenomena (Maggino, 2017, p.127).

2.3.4 Weighting criteria

Regarding the weights, the methodologies discussed here are all data-driven techniques. However, there can be sound reasons to state the importance of indicators in measuring the dimension or latent variable (Maggino, 2017; Greco et al. 2019). These reasons ought to be clearly declared in the construction process. Moreover, as we discussed above, the DP2 distance method could be used while imposing predetermined weights.

2.3.5 Synthesizing indicators: aggregative-compensative approach

The aggregative-compensative approach assumes that only one latent variable is being measured. Critical issue is the correlation among the indicators to be aggregated. The interpretation of this correlation depends, as we have discussed above, on the model of measurement.

A reflective model requires very high correlation among indicators as they all are manifestations of the same latent variable, which may be multidimensional. Hence, indicators referring to the same dimension may be aggregated.

A formative model usually encompasses indicators that measure independent dimensions of the multidimensional latent variable. Correlation among indicators suggests they overlap and may induce discarding one of them, which should be done attempting to preserve comparability among observations and over time (Maggino, 2017, p.122).

The main criticism of this aggregative approach charges against its main strength, unidimensionality, and claims that conveying into an unidimensional measure a multidimensional, complex and dynamic concept, such as well-being, rises critical conceptual, methodological and technical issues (Maggino, 2017). For instance, aggregating may result in two distinctive observations being assigned the same score. Hence, it is fundamental to identify the befitting aggregation technique. This step should consider the issues of comparability and measurement homogeneity, which both refer to normalization and polarity as discussed above, and compensability.

A compensatory technique allows for low values in some indicators to be compensated by high values in other indicators. Compensation among indicators determines the interpretation of the weights. The weights may be viewed as trade-offs between indicators or as importance coefficients. The compensatory approach implies that weights should be interpreted as trade-offs, not as importance coefficients (OCDE, 2008; Maggino, 2017; Greco et al. 2019). Both the linear (the composite indicator is the sum of the weighted indicators) and geometric (the composite indicator is the product of indicators, each of them raised to the power of its weight)aggregation schemes are compensatory. In linear aggregation, compensability is constant; in geometric aggregation, compensability is lower for those indicators with worse values. All of these compensability issues should be taken into account.

A multidimensional phenomena such as well-being, where each dimension may be represented by several indicators, may require to build a composite indicator for each dimension and, then, obtain the overall index by aggregating the partial composite indicators. In which case, a compensatory approach could be followed within dimensions and a non compensatory or partially-compensatory approach[3] among dimensions (Mazziotta and Pareto, 2017). However, non compensatory approaches, such as multi-criteria, may elude compensability but are computationally costly with a high number of observations (Munda and Nardo, 2007 in OCDE, 2008 p. 33).

[3] MPI is an example that summarizes partially non compensatory indicators

2.4 Properties and weight scheme

2.4.1 Properties of the aggregation method

In table 3.1 we summarize the mathematical properties of the aggregation methods described in section 2. Following the properties pointed out by Pena Trapero (1977, 2009) and Zarzosa Espina (1996), as well as the desired properties for Indicators Construction by Maggino (2017), we present below how the methods presented in this study -Distance P_2, PCA, DEA and MPI- perform regarding these mathematical properties.

Table 2.1: Summary of the mathematical properties

Existence	
For any observation (country, region, etc.), the composite indicator derived from each method is well defined.	1.1 **Distance** P_2 True. 1.2 **PCA** True. 1.3 **DEA** True. 1.4 **MPI** False,(M_{z_i} must be different to 0 for all $i \in \{1,...,n\}$).
Monotony	
An increase(decrease) in one single indicator with positive(negative) polarity while keeping the other indicators constant produces an increase in the composite indicator.	2.1 **Distance** P_2 True. 2.2 **PCA** True (if all single indicators must have the same polarity). 2.3 **DEA** True. 2.4 **MPI** True.
Symmetry	
The composite indicator does not depend on the order of the single indicators.	3.1 **Distance** P_2 False (The methodology impose an order). 3.2 **PCA** True. 3.3 **DEA** True. 3.4 **MPI** True.
Invariance	
The aggregation method is invariant by origin and scale changes.	4.1 **Distance** P_2 True. 4.2 **PCA** False. (Depends on the normalization that has been chosen). 4.3 **DEA** False. (Depends on the normalization that has been chosen) 4.4 **MPI** True.

Completeness	
The weighs of the single indicators are introduced according to their relevance avoiding duplication of information.	5.1 **Distance** P_2 True. 5.2 **PCA** False. The weighting scheme is not related to the relevance of each indicator . 5.3 **DEA** False. The information provided by each indicator is particular to each observation. Therefore, we cannot know what information is provided by each indicator as a whole. 5.4 **MPI** False. The information provided by each indicator is particular to each observation. Therefore, we cannot know what information is provided by each indicator as a whole.
Objectivity	
The ranking or the weights are not arbitrarily determined.	6.1 **Distance** P_2 True. 6.2 **PCA** True. 6.3 **DEA** True. 6.4 **MPI** True.

2.4.2 Weight scheme

In this section, we would like to highlight some important aspects of the analysed methods. To do so, we use the Human Development database, where ten single indicators was chosen from the Human Development Report Office (HDRO) for 2017, in order to construct a composite indicator. Table 4.1 depicts the definition of the selected single indicators and additional information. All indicators have been chosen with the same polarity due to PCA cannot be used with indicators of different ones polarities. Following OECD in its Regional Well-Being Dataset, we normalize using Equation 2.7 taking *max* as the best scenario for each single indicator X_j, $j \in \{1,...,10\}$. This normalization has been used for DEA, MPI, and PCA, such that the data set will express the distance each observation has from the best scenario in each indicator. For the Distance P_2 calculation, the reference vector has as coordinates the minimum of each single indicator. The greater the distance to the minimum, the greater the calculated distance P_2

$$\widehat{x_{ij}} = \frac{x_{ij} - min(X_j)}{max(X_j) - min(X_j)} \tag{2.7}$$

Table 2.2: Single indicators at country level

Title	Definition	Source	Polarity
Population **POP**	Urban population (%)	UNDESA (2018a). World Urbanization Prospects: The 2018 Revision. New York. https://esa.un.org/unpd/wup/. Accessed 17 May 2018.	Positive
Education **EDUC**	Expected years of schooling (years)	UNESCO Institute for Statistics (2018), ICF Macro Demographic and Health Surveys, UNICEF Multiple Indicator Cluster Surveys and OECD (2017a).	Positive
Environmental sustainability **ENVSUS**	Renewable energy consumption (% of total final energy consumption)	World Bank (2018a). World Development Indicators database. Washington, DC. http://data.worldbank.org. Accessed 6 July 2018.	Positive
Gender **GENDER**	Estimated gross national income per capita, female (2011 PPP $)	HDRO calculations based on ILO (2018a), UNDESA (2017a), World Bank (2018b) and IMF (2018).	Positive
Health **HEALTH**	Life expectancy at birth (years)	UNDESA (2017a). World Population Prospects: The 2017 Revision. New York. http://esa.un.org/unpd/wpp/. Accessed 10 May 2018.	Positive
Income/composition of resources **INCOM**	Gross domestic product (GDP), total (2011 PPP $ billions)	World Bank (2018a). World Development Indicators database. Washington, DC. http://data.worldbank.org. Accessed 6 July 2018.	Positive
Mobility and communication	Mobile phone subscriptions (per 100 people) **MOBCOM**	ITU (International Telecommunication Union) (2018). ICT Facts and Figures 2018. www.itu.int/en/ITU-D/Statistics/Pages/stat/. Accessed 18 July 2018.	Positive
Socio-economic sustainability	Rural population with access to electricity (%) **SOCECO**	World Bank (2018a). World Development Indicators database. Washington, DC. http://data.worldbank.org. Accessed 6 July 2018.	Positive
Trade and financial flows	Exports and imports (% of GDP) **TRADE**	World Bank (2018a). World Development Indicators database. Washington, DC. http://data.worldbank.org. Accessed 6 July 2018.	Positive
Work, employment and vulnerability **WORK**	Employment to population ratio (% ages 15 and older)	ILO (International Labour Organization) (2018a). ILOSTAT database. www.ilo.org/ilostat. Accessed 13 April 2018.	Positive

The database contains missing values. To balance it, the missing values have been replaced with the imputed values using predictive mean matching (PMM). Figure 4.1 shows that almost 89% of the samples are not missing any information. The worse single indicator contains 8% of missing values and the best none.

We observe that the density functions have not been altered by the imputation of missing values. For instance, Figure 4.2 shows magenta slope in the density of the imputed data of the 5th worse single indicators, while blue slope density shows the observed data.

Figure 2.1: Percentage and patterns of missing values for each indicator

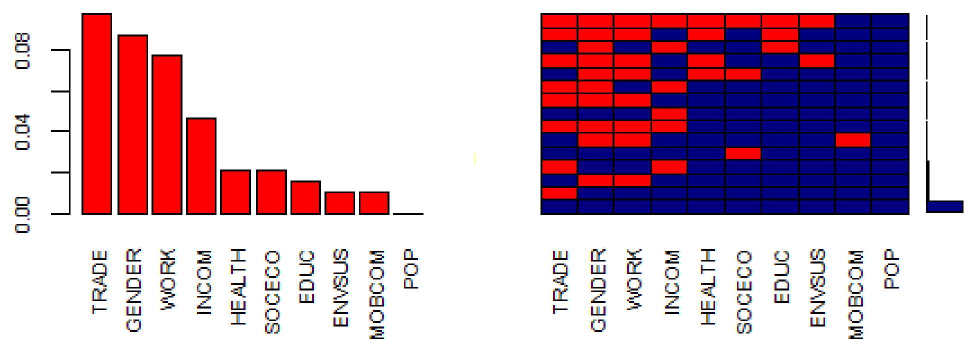
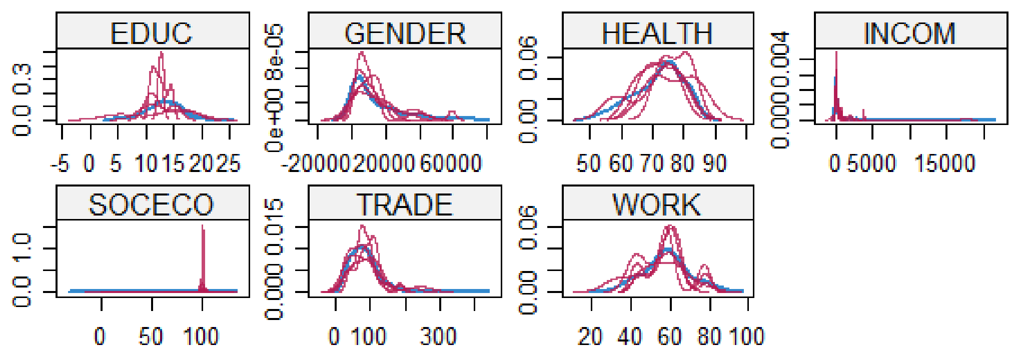


Figure 2.2: Empirical density function of variables (with and without missing)



The use of PCA, Distance P_2 is not suitable when the correlations between the indicators is very weak. Table 4.2 shows correlation among single indicators. The single indicators provided present high and medium correlation to avoid this problem.

Please note that the choice of single indicators for constructing this composite indicator is irrelevant, it is simply a reference framework to observe the possible divergences that can arise among the discussed methods. Hence, we calculate the composite index with the selected single indices using the four procedures analysed. Kendall rank correlation test is used to measure the ordinal association between two composite indicators (Kendall, 1938). This non-parametric statistic test assesses whether two variables may be regarded as statistically dependent by determining if there exist a monotonic relationship between them, namely, preserving the rank (τ statistic close to one) under the null hypothesis of independence. Intuitively, Kendall tau represents the difference between the probability that the analyzed data are in the same order for the two computed indicators versus the probability that the analyzed data are in different orders for the two computed indicators. In the case of being close to 0 there is no evidence that the ranks are equal. Table 4.3 shows the close order-relationship among the composite indicators, from which it follows that some of these procedures are unreliable for constructing composite indices if the goal

Table 2.3: Correlation among single indicators

	POP	EDUC	ENVSUS	GENDER	HEALTH	INCOM	MOBCOM	SOCECO	TRADE	WORK
POP	1.00	0.58	-0.51	0.60	0.59	0.13	0.42	0.51	0.23	-0.15
EDUC	0.58	1.00	-0.48	0.67	0.80	0.16	0.57	0.70	0.22	-0.12
ENVSUS	-0.51	-0.48	1.00	-0.38	-0.60	-0.14	-0.44	-0.73	-0.27	0.34
GENDER	0.60	0.67	-0.38	1.00	0.65 0,51	0.16	0.45	0.48	0.43	0.06
HEALTH	0.59	0.80	-0.60	0.65	1.00	0.16	0.54	0.83	0.27	-0.12
INCOM	0.13	0.16	-0.14	0.16	0.16	1.00	0.05	0.15	-0.19	0.01
MOBCOM	0.42	0.57	-0.44	0.45	0.54	0.05	1.00	0.56	0.35	-0.09
SOCECO	0.51	0.70	-0.73	0.48	0.83	0.15	0.56	1.00	0.20	-0.29
TRADE	0.23	0.22	-0.27	0.43	0.27	-0.19	0.35	0.20	1.00	-0.01
WORK	-0.15	-0.12	0.34	0.06	-0.12	0.01	-0.09	-0.29	-0.01	1.00

is to rank observations. The null hypothesis of independence was not rejected between the composite indicators constructed by Distance P_2 versus PCA and PCA versus MPI. On the other hand, independence was rejected between Distance P_2 versus DEA; Distance P_2 versus MPI; DEA versus MPI and DEA versus PCA, as the concordance pairs are very low. The negative sign for the MPI versus PCA test indicates a reverse order between these two indicators. These divergences are caused because each of the methods befits a specific measurement goal, as we discuss next. Is there a mistake in the p-values reported for the Kendall test between Dp2 and PCA? In fact, I don't quite get when independence is rejected and when is not.

Table 2.4: Results of Kendall's rank correlation test (p-value).

	DEA	PCA	MPI
DP2	0.4535($<0.0001***$)	-0.0181(0.7066)	0.3441($<0.0001***$)
DEA		0.2358($<0.0001***$)	0.2825($<0.0001***$)
PCA			-0.0911(0.0566)
p-value significance	$***\ p<0.001$, $**\ p<0.01$, $*\ p<0.05$,.0.1		

On the other hand, we analyze the behaviour when observations are deleted to the dataset being considered. The objective is to analyze if the ranking of the composite index is statistically altered. Starting from the original sample, we eliminate ten random observations in each iteration with reposition using Monte Carlo procedure and we compute the composite index. subsequently, we compute the composite indicator from original data from which the random sample from the previous step is removed. The Kendall rank correlation test is calculated for the resulting composite indices. In this case, Figure 4.3 show MCI is sensitive to deletion of observations, while the random elimination of observations leaves invariant the range in the rest of the methods.

2.4.3 Weighting schemes: indicators ranking

To understand these divergences between the four methods, we take the previous database as baseline. The weighting scheme of the different methods provides part of the answer. As briefly specified in the Methodology section, the way in which the weights are obtained differs significantly from one another.

1. **DEA-BoD** In our view, DEA-BoD constitutes a good tool to obtain good strategies or policies for each observation. DEA-BoD provides a collection of weights (different for each observation), endogenously determined. However the differential weighting inherent in the process prevents comparison among observations, (Greco et al. 2018). To overcome this shortcoming, Peiro et al. (2018) propose to combine DEA and Multi-Criteria-Decision-Making

(MCDM) techniques to achieve both a common set of weightings and allow for comparisons in the ranking of observations. The latter method gives a unique structure of weights that allows a ranking, although it is not possible to compare the results of the compound index between pairs, since the result obtained through DEA-BoD-MCDM is not a metric. Also, Spearman's rank correlations between the well-being scores was obtained with the DEA-BoD and different DEA-BoD-MCDM approaches. As a result, high rank and statistically significant at the 1% confidence level correlation was found. We continue with the DEA-BoD rank as reference, because there are no significant differences between the ranks of the latter approaches. As stated in section 2, DEA-BoD approach impose on each observation under evaluation the optimal set of weightings such that the observation achieves the best relative position with respect to all other observations. Therefore, for each observation, a whole weighs is computed. Thus, we have not an unique weight associated to each indicator but a collection of weights. In order to provide an approximation to this collections, Table 4.4 provide a result summaries (Max., 3rd Qu., Median, Mean, 1rd Qu., Min:) of each indicator[4]. The previous analysis shows that SOCECO is the most important indicator, being HEALTH and GENDER the less relevant.

Table 2.5: Summary of the weights for DEA

v1	v2	v3	v4	v5
Min. :0.00000	Min. :0.00000	Min. :0.0000	Min. :0.00000	Min. :0.00000
1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.00000
Median :0.00000	Median :0.00000	Median :0.0000	Median :0.00000	Median :0.00000
Mean :0.03888	Mean :0.02403	Mean :0.1958	Mean :0.01692	Mean :0.01335
3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.3097	3rd Qu.:0.00000	3rd Qu.:0.00000
Max. :1.00000	Max. :0.75824	Max. :1.0000	Max. :0.99835	Max. :0.93764

v6	v7	v8	v9	v10
Min. :0.0000	Min. :0.00000	Min. :0.0000	Min. :0.0000	Min. :0.00000
1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.00000
Median :0.0000	Median :0.00000	Median :1.0000	Median :0.0000	Median :0.00000
Mean :0.0845	Mean :0.08278	Mean :0.6019	Mean :0.1065	Mean :0.10195
3rd Qu.:0.0000	3rd Qu.:0.00000	3rd Qu.:1.0000	3rd Qu.:0.0000	3rd Qu.:0.05978
Max. :1.0000	Max. :0.89727	Max. :1.0000	Max. :1.0000	Max. :1.00000

2. **MPI** As described in the MPI Equation 2.6, for each observation (country or region), the arithmetic mean is calculated after normalization, therefore we can assume that all indicators have the same weight, to which a function of indicator variability called *penalty* is added in order to minimize a the duplication of information. Notice that, the weight scheme is not included in Table 6 because all indicators have the same relevance. In this way, the penalty could restore the unbalance produced by a poor performance in some indicators. A variant of the previous is the Adjusted Mazziotta?Pareto Index (AMPI). The AMPI summarises a set of indicators that are assumed to be non-substitutable, Mazziotta & Paretto (2018). However, as with DEA-BoD, this aggregation does not allow to compare observations with each other, and consequently the ranking provided has a difficult interpretation.

[4]This difficulty in understanding the relevance of weights and the contribution of each to the composite indicator is one of the weaknesses of this approach.

3. **PCA** PCA uses the factor loadings as indicators weights, for instance, to make HDI (Noorbakhsh, 1996). Since the number of components/factors to be retained must be chosen by the decision maker, subjectivity is introduced to a certain degree. It is common that in the context of composite indexes, the first component is chosen to make the composite index (Greyling & Tregenna 2016). Extensive PCA-related literature is provided by Greco et al. (2019). Although the first component alone only explains a portion of the variance of the indicators, we assume the factor loadings of the first component as the weights for the computed composite indicator. Table 2.6 shows the coordinates of the eigenvector[5] (weights) corresponding to the high eigenvalue $\lambda = 4.70347$ that explains the 47% of variance of the composite indicator. With this methodology, TRADE is the most relevant indicator, followed by HEALTH and SOCECO. Notice that PCA assumes a linear relationship between the indicators. Thus, in the current case, that TRADE, WORK and INCOM have low correlations with the rest of indicators can lead to unwanted results. In addition, though the reductionism of this approach is useful to avoid duplicity of information, the weights so endogenously obtained could not necessarily correspond to the real links among the indicators (Saisana & Tarantola 2002).

4. **Distance P2** Distance P_2 provides another rank distribution of indicators. After ordering the indicators through the correlation between Frechet's distance and the indicators, we compute the coefficient of determination in the multiple linear regression of x_j over $x_{j-1}, ..., x_1$ assuming $R_1^2 = 0$, then the weights $1 - R^2_{j,...1}$ are obtained. Notice that Distance P_2 measures a distance. Namely, let $X_i = (x_{i1}, ..., x_{im})$ denote the vector whose coordinates correspond to the indicators values for each observation $i \in \{1, ..., n\}$. Then Distance P_2 is symmetric $dp_2(X_i, X_j) = dp_2(X_j, X_i)$, positive $dp_2(X_i, X_j) > 0$ if $i \neq j$ and $dp_2(X_i, X_j) = 0$ if $i = j$, and satisfies the triangular inequality $dp_2(X_i, X_j) + dp_2(X_j, X_k) > dp_2(X_i, X_k)$. The latter property allows for comparisons between observations, and this is the only one out of the analysed methods that obeys this property. However, the method has a strong dependence on the linearity of the model, for which when the indicators have very weak correlations the results can be wrong. Table 2.6 shows GENDER as the most relevant indicator followed by INCOM and WORK, being SOCECO the least relevant indicator.

Table 2.6: Weights indicators DEA,DP2,PCA

Indicators	POP	EDUC	ENVSUS	GENDER	HEALTH	INCOM	MOBCOM	SOCECO	TRADE	WORK
DEA	0.0368	0.0256	0.1929	0.0160	0.0126	0.0800	0.0822	0.6082	0.1058	0.1009
DP2	0.5629	0.3258	0.4183	1	0.5812	0.8766	0.6470	0.2857	0.7529	0.8030
PCA	0.1167	0.1548	0.1187	0.1219	0.1703	0.0754	0.1055	0.1576	0.3577	0.1086

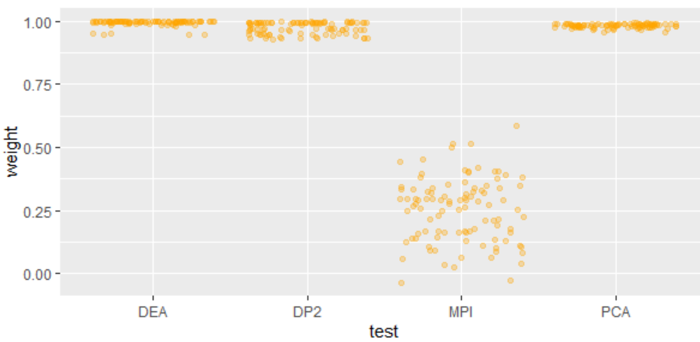
2.4.4 Sensitivity analysis

Sometimes, the observations should be chosen for further efficiency or deleted if they are contaminated by data errors (Wilson, 1995). In this way, an approach within sensitivity analysis studies responses with given data to manipulations -addition or substraction- in the number of observations. The objective is to analyze if the ranking of the composite index is statistically altered when several observations are, e.g., deleted. Starting from the original sample, we eliminate ten random observations in each iteration with reposition using Monte Carlo procedure and compute the composite index. Subsequently, we compute the composite indicator for original data from

[5]Eigenvector provides the direction for which the variance of the first component is maximized.

which a random sample from the previous step is removed. The Kendall rank correlation test is calculated for the resulting composite indices. Figure 2.3 shows that MPI is sensitive to deletion of observations, while the random elimination of observations leaves invariant the ranks in the rest of the methods. The Z-score normalization approach assumed in MPI notably provide a significance variation in the composite index computed when observations are deleted.

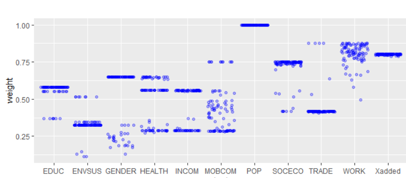
Figure 2.3: Tau outputs from Kendall rank correlation test for sensitivity analysis



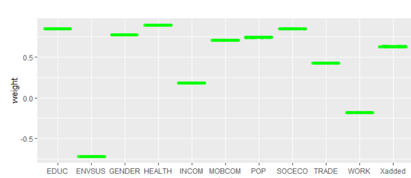
The high correlation of indicators can generate a duplication of information in the calculated composite index. The aggregation methods used should be able to overcome this difficulty in order to avoid overlaping of information. In this context, the weighting scheme plays an essential role. To test the behaviour of the methods presented in this work, we introduce an indicator which is a lineal combination of the other indicators (noise)in order to determine the extent to which they are capable of eliminating this redundant information. Through 100 random combinations of the 10 indicators, we compute the weights for Distance P_2, PCA and DEA-BoD[6]. Figure 2.4 shows the results of the Monte Carlo method used. Both DEA and Distance P_2 reasonably discriminate the noise introduced into the model. However, PCA is not able to discriminate through the weights the variable entered. This is due to the fact that the weights are actually the coordinates of the eigenvector corresponding to the highest eigenvalue of the covariance matrix, i.e. the value that maximises the variance of the first component, or rather, of the composite index, and are not strictly speaking weights that balance the model according to the relevance of the indicators.

Figure 2.4: Weighs indicators

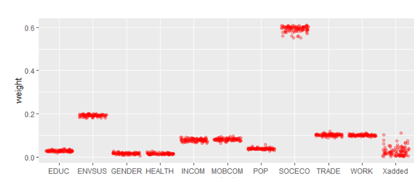
(a) Distance P_2



(b) PCA



(c) DEA-BoD



[6]MPI has not been included in this analysis because the weights are all the same for this procedure

2.5 Conclusions

Most socioeconomic phenomena are multidimensional, which renders a single indicator unable to capture the inherent complexity in, for example, development, poverty, well-being. The construction of composite indicators should follow a respectful methodological approach to ensure that the *big picture* fundamentally captures what it is meant to (OECD,2008). The methodological process to construct a composite indicator starts with the precise definition of the conceptual framework of measurement, (Maggino, 2017, p.87), which may be reflective or formative and conditions the selection of single indicators that (attempt to) measure the various dimensions of the concept. Indicators should be selected befitting the phenomena to be measured. Following the theoretical framework, the suitable normalization approach for individual indicators should be applied, and normalized indicators should be aggregated taken into consideration compensability issues and weighting schemes. Finally, the robustness of the composite index should be assessed.

In this paper we focus on four aggregations methods to analyze their weighting and aggregation approaches. The choice of these four approaches relies on their methodological differences in their aggregation schemes, despite all of them being data-driven techniques. First, a method based on imposing each observation under evaluation the optimal set of weights that rate it in the best relative position with respect to all other observations. The optimization of each individual observation for which you can compute a discrete piecewise frontier through the set of Pareto-efficient Decision Making Units (DMU) is called DEA-BoD. Secondly, a method whose fundamental virtue is reducing the dimensionality of the dataset when there are high correlations among indicators, PCA. Thirdly, a method that builds a metric that inherits analytical properties allowing observations to be compared, the Distance P_2. Finally, a method that produces a composite index that penalises substitutability among indicators, Mazziotta-Pareto Index (MPI).

From the 2017 Human Development database, we select ten individual indicators to construct a composite indicator that allows us to check the divergences of the four methods analysed. All indicators have been chosen with the same polarity to apply the four methods. Additionally the same normalisation approach has been applied, except for Distance P_2 for which a reference vector has been chosen in accordance with the above approaches.

The core of PCA's philosophy is to optimize the variation of the new components that reduce the dimensionality of the indicators, as long as there is high correlation between them. All indicators must have the same polarity and its use is more suitable for reflective models. The use of this methodology for the construction of composite indices outside this context can lead to important errors. It is not advisable to use PCA when indicators are poorly correlated. The procedure is stable when eliminating observations, maintaining the ranking of the calculated composite index. Additionally, when an indicator (linear combination of the principals) is added, PCA does not discriminate against it.

DEA-BoD is reasonable to use when the aim is to study the efficiency of each unit separately (observation). However, a composite index constructed following this methodology does not allow comparisons, which makes very complicated the interpretation of the results as a whole. The procedure is stable when eliminating observations, maintaining the ranking of the calculated composite index. Additionally, when an indicator (linear combination of the principals) is added, DEA assigns a weight that makes this new indicator irrelevant in the model.

Distance P_2 computes a composite index resulting from a metric, therefore, unlike the previous procedure, it allows a comparison between observations and provides a mathematical structure for the analysis of the whole composite indicator. However, given the dependence of the model on its linearity, when the correlations between variables are very poor, Distance P_2 does not behave efficiently as PCA. The procedure is stable when eliminating observations, maintaining the ranking of the calculated composite index. In addition, when an indicator (linear combination of the principals) is added, Distance P_2 assigns a weight by removing irrelevant information.

MPI is an aggregation method that bears the cost of compensability. Unlike the Distance P_2 or PCA, the indicators can be poorly correlated. However, the penalty defined in this approach will not always act as a catalyst for imbalances between indicators. This penalty, calculated for each observation, can also make the composite index difficult to understand, as it is the case with the DEA-BoD approach. The procedure is not stable when observations are eliminated, generating significant alterations in the ranking of the calculated composite index.

Chapter 3

Dealing with weighting scheme in composite indicators: an unsupervised distance-machine learning proposal for quantitative data

Abstract

There is increasing interest in the construction of composite indicators to benchmark units. However, the mathematical approach on which the most commonly used techniques are based does not allow benchmarking in a reliable way. Additionally, the determination of the weighting scheme in the composite indicators remains one of the most troubling issues. Using the vector space formed by all the observations, we propose a new method for building composite indicators: a distance or metric that considers the concept of proximity among units. This approach enables comparisons between the units being studied, which are always quantitative. To this end, we take the P2 Distance method of Pena Trapero as a starting point and improve its limitations. The proposed methodology eliminates the linear dependence on the model and seeks functional relationships that enable constructing the most efficient model. This approach reduces researcher subjectivity by assigning the weighting scheme with unsupervised machine learning techniques. Monte Carlo simulations confirm that the proposed methodology is robust.

3.1 Introduction

Composite indicators have clear advantages that justify their increasing use for summarising complex and multidimensional realities that are not directly measurable. For instance, they are used to support decision makers, make comparisons and assess the progress of units (companies, countries, regions, etc.) over time or facilitate communication with the general public (Maggino, 2017; Mazziotta and Pareto, 2017; Sánchez et al., 2018).

Composite indices developed by international organizations and institutions choose simplicity as the best methodological option. The most widely used aggregation method is the arithmetic mean. Some examples of indices that use the arithmetic mean include the United Nations Human Development Index from 1990 until 2010 when it was substituted by the geometric mean (UNDP, 2018); the Ease of Doing Business ranking, which studies business regulation at country level (World Bank, 2020); the Better Life Index developed by the Organisation for Economic Cooperation and Development (OECD, 2017) to visualise well-being; the Canadian Index of Wellbeing developed by the Canadian Research Advisory Group University of Waterloo (Canadian Index of Wellbeing; 2016); or the Sustainable Development Goals (SDG) Index supported by Cambridge University Press to assess where each country stands in achieving the SDGs (Sachs et al., 2018). Other institutions have combined the arithmetic mean and principal component analysis (PCA) in their indices, such as the World Economic Forum's Global Competitiveness Index since 2008 (Schwab and Porter, 2008) and the European Commission's European Regional Competitiveness Index from 2009 to 2019 (Annoni and Dijkstr, 2019). PCA is applied to verify whether the indicators within each dimension are internally consistent and then aggregate them by an arithmetic mean in a second step. In addition to the composite indices developed by international organisations, it is worth highlighting empirical applications that use data envelopment analysis (DEA) in academia. DEA is a methodological approach to evaluate the performance of a set of observations referred to as decision making units (DMUs) that subsequently transform multiple inputs into multiple outputs. DEA methodology has the advantage that it does not depend on the method chosen to normalise the data or on the weights used for aggregation. Recent studies in this line have constructed composite indicators to assess the level of competitiveness of Costa Rican counties (Lafuente et al., 2020), evaluate the provision of local municipal services in Flanders (D'Inverno and De Witte, 2020) or to assess the performance of public hospitals in Portugal from the perspective of users and providers (Alves Pereira et al., 2021).

These indices and rankings, some of which are more influential than others, are taken as a reference to make comparisons between companies, countries or regions and guide decision-making on public policies and in private companies. However, despite the popularity of these methods, none of them enables benchmarking because they do not provide a mathematical structure to analyse the results through a metric that permits comparisons between units. Benchmarking is one of the required characteristics in a composite indicator and is broadly defined as the capability to interpret results according to a specific frame (Maggino, 2017, pp. 106-107). Within the setting of the construction of composite indicators, a metric is the natural way to establish the proximity or distance between the analysed observations and therefore perform benchmarking in a rigorous and reliable way. Additionally, the weighting schemes of these methods have serious disadvantages that call into question the veracity of their results. Neither the arithmetic mean nor the geometric mean avoid the redundant or overlapping information of single indicators, and PCA only removes the linear redundant information (Becker et al., 2017; Greco et al., 2019; Jiménez-Fernández and Ruiz-Martos, 2020; Mazziotta and Pareto, 2019; OECD, 2008).

This study presents a new method for building composite indicators that can be applied in formative measurement models (Coltman et al., 2008; Diamantopoulos et al., 2008), using quantitative data and a partially compensatory aggregation method based on the mathematical concept of distance or metric. To this end, we draw on the P2 Distance method (hereafter DP2) developed by Pena Trapero (1977) that has been widely used in very recent applications (see, for instance, Cuenca et al., 2019; Martín et al., 2020; Sánchez et al., 2018; Sánchez and Ruiz-Martos, 2018).

Taking into account the available tools when Professor Pena Trapero developed the DP2 method, this method constituted a great methodological advance, particularly due to the introduction of a metric in the construction of composite indicators. Nonetheless, DP2 has some limitations. In this study we address these weaknesses using machine learning (ML) techniques.

More specifically, our composite indicator is the outcome of a weighted ℓ^2 metric, where the weights are computed using unsupervised ML algorithms. Our proposal makes several notable contributions. Firstly, our composite indicator is able to measure distances to perform benchmarking between the units studied in a rigorous way. Secondly, it efficiently eliminates the redundant information provided by the single indicators, so that the weights of the single indicators properly reflect their relative importance. Thirdly, it satisfies a sufficiently large number of mathematical properties to be considered a reliable method. Finally, our composite indicator has passed a robustness analysis.

To achieve our goal, the rest of the paper is structured as follows. In section 2, the DP2 method is reviewed and its positive aspects and limitations are highlighted. Section 3 introduces the `DL2`, the methodology proposed in this study. Specifically, we analyse the formula to calculate `DL2` composite indicators, select the best set of disjoint polynomials between the composite indicator and the set of single indicators, estimate the weights of the single indicators in the calculation of the composite indicator, remove redundant information, and the iterative method or algorithm for calculating the values of the `DL2` composite indicators in each unit. Section 4 analyses the mathematical properties the `DL2` methodology satisfies and its goodness of fit. Section 5 examines three strategies for checking the robustness of the proposed composite indicator. Section 6 compares the properties that the DP2 and `DL2` methods satisfy. Finally, section 7 concludes.

3.2 Drawing on P2 Distance

As in any empirical analysis in which data are used to test a theory or estimate a relationship between variables, the construction of composite indicators requires performing a series of stages to ensure a reliable result. This is a complex task that involves, at least, the following steps: (1) defining the phenomenon to be measured (latent construct), which in turn requires identifying the nature and direction of the structural relationships between the latent construct and the observed variables; (2) selecting a group of variables or single indicators that represent the phenomenon to be studied according to the conceptual framework; (3) normalising the single indicators; (4) weighting and aggregating the normalised indicators using a mathematical method (compensatory, partially compensatory or non-compensatory) and (5) validating the composite index (Jiménez-Fernández and Ruiz-Martos, 2020; Maggino, 2017; Mazziotta and Pareto, 2017; OECD, 2008). Likewise, to maximise the robustness and validity of a composite indicator, the most appropriate methodological choices must be made in each of the previous steps.

In this section, we review how the DP2 method responds to these stages, which will allow us to highlight its strengths and identify some weaknesses. Thus, our methodological proposal for constructing composite indicators focuses on overcoming the drawbacks of the DP2 method while taking advantage of its strengths.

Let us first introduce some formal technical concepts and definitions regarding the metric or distance. Let Λ be a nonempty set and let $\mathbb{R}$ be the set of real numbers. A function $d : \Lambda \times \Lambda \to \mathbb{R}^+$ is said to be a metric or distance if for all $A, B, C \in \Lambda$ the following statements are satisfied:

1 $d(A,B) \geq 0$; $d(A,B) = 0$ if and only if $A = B$,

2 $d(A,B) = d(B,A)$,

3 $d(A,B) \leq d(A,C) + d(C,B)$ (triangular inequality).

Let X be an $n \times m$-dimension matrix where columns $X_{\cdot 1}, \ldots, X_{\cdot m}$ represent the single indicators and the rows of X refer to the i observations or units (regions, countries, etc.). Let $X_{i\cdot} = (x_{i1}, \ldots, x_{im})$ be an m-dimension row vector associated to the i-observation and let $X_{*\cdot} = (x_{*1}, \ldots, x_{*m})$ be a hypothetical unit we call the target vector or baseline. For instance, the vector $X_{*\cdot}$ may represent the best or worst case scenario for each of the single indicators depending on the phenomenon to be measured[1]. Let $d_{ij} = |x_{ij} - x_{*j}|$ be the distance from the i-observation $i \in \{1,...,n\}$ to the j-coordinate of the target vector. For each $j \in \{1,...,m\}$, $R^2_{j,...1}$ represents the coefficient of determination in the multiple linear regression of $X_{\cdot j}$ over the preceding indicators $X_{\cdot j-1}, ..., X_{\cdot 1}$ assuming $R^2_1 = 0$. Let $\omega_j = 1 - R^2_{j,...1}$ be the weights computed following an iterative process explained below.

Distance $DP2$ can be defined as follows:

$$DP2(X_{i\cdot}, X_{*\cdot}) = \sum_{j=1}^{m} \frac{d_{ij}}{\sigma_j}\omega_j \tag{3.1}$$

where σ_j is the standard deviation of the j-single indicator $j \in \{1,...,m\}$ subject to the standard deviation $\sigma_j \neq 0$.

Next, we review how the DP2 solves the main issues in the construction of composite indicators and the drawbacks detected in the DP2 method.

1. Under the DP2 framework, the $\frac{|x_{ij} - x_{*j}|}{\sigma_j}$ term transforms the single indicators into dimensionless numbers because the units of measure in the numerator and denominator are cancelled. However, this transformation does not ensure that the scale of measurement is the same for all the indicators, since the transformed indicators have a minimum value of zero but the maximum value is not limited. The maximum value depends on the specific distributions of the indicators. To correct this drawback, the indicators must be normalised.

2. Generally, the single indicators that make up the composite indicators are correlated, as they provide information from the same constructor. Hence, to perform a composite indicator of a latent phenomenon, a partially compensatory or non-compensatory aggregation technique of the single indicators should be chosen. In the case of DP2, several studies have tested the results reached with DP2 using other composite indicators developed under a

[1] Some single indicators may be positively correlated with the latent variable (positive polarity), whereas others may be negatively correlated with it (negative polarity). For instance, investment in R&D would be positively associated with sustainable development (latent variable), whereas CO2 emissions would be negatively associated. In this case, the target vector would be formed by the worst case scenario in all single indicators: the minimum value of the indicators with positive polarity and the maximum value of the single indicators with negative polarity. Thus, the greater the distance of one unit from the target vector, the higher the value of the composite indicator (i.e. the higher the level of sustainable development).

multi-criterion approach (with a double reference point) and concluded that the results are very similar to those of weak and mixed indices (see Cuenca-García et al., 2019; Luque et al., 2017). This indicates that DP2 can be considered a partially compensatory approach. Regarding this, the method should provide an adequate treatment to avoid the overlapping of information. Unlike other methods (see Jiménez-Fernández and Ruiz-Martos, 2020), DP2 introduces the coefficient factor $1-R^2$ for this purpose. Nevertheless, the coefficient of determination R^2 only detects the linear correlations between single indicators. This is one of the limitations of the DP2 method that we try to overcome with our proposal.

3. DP2 provides an iterative method to objectively assign weights to the single indicators in the composite indicator. To do so, the Fréchet distance (FD) is taken as a starting point. The FD corresponding to the i-observation is defined as follows

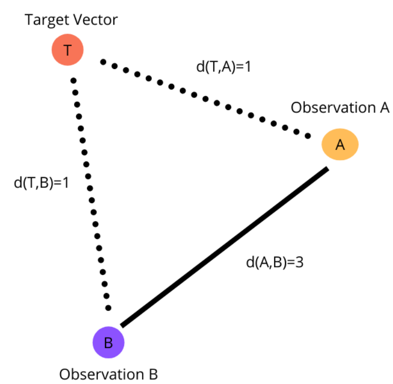
$$FD_i(X_{i\cdot}, X_{*\cdot}) = \sum_{j=1}^{m} \frac{d_{ij}}{\sigma_j} \tag{3.2}$$

All the single indicators in the FD have the same weight or importance. In a first step, DP2 computes the pairwise correlation coefficients between each single indicator and the FD and then sorts the indicators from highest to lowest according to the absolute values of the pairwise correlation coefficient. The indicators are introduced in the model following the previous rank and the weights are calculated according to this criterion. The process continues iteratively until the difference between two average adjacent $DP2$s is less than a fixed threshold. This criterion does not guarantee convergence in the order of the value of the composite indicator of the units or observations. In other words, the DP2 criterion can choose some values of the composite indicator so that the average of the difference between these values and the previous ones is very small but with large differences in the ranking of the units. This is why the criterion to reach the final value of the composite indicator in each unit should take into account the order convergence of units rather than the convergence in mean.

4. DP2 allows comparing observations and provides a mathematical structure to analyse the results through a metric, except when there are collinearity problems among single indicators. More specifically, the procedure will not provide a satisfactory result when a single indicator is a linear combination of other indicators. In this case, for some $j \in \{1,\ldots,m\}$ the weight ω_j is equal to zero and, therefore, the formula $DP2$ is not a distance or metric (statement (1) of a metric is not satisfied) because the defined weights are not always strictly positive. This is an essential aspect in order to inherit the rich properties of a metric. For example, let us assume that function d defines a composite indicator to measure socio-economic status. Let T be the vector of the worst observations of a given set of regions, that is, the hypothetical region with the lowest socio-economic status. Let us assume that A, B denote two different regions belonging to the same set such that $d(T,A) = d(T,B) = 1$ and $d(A,B) = 3$. Figure 3.1 shows the measures of the distance between observations where the triangle inequality is not satisfied. Accordingly, $3 = d(A,B) > d(T,A) + d(T,b) = 1+1$. A metric space is a set with an associated distance function. This function allows us to establish the concept of proximity, so that for any pair of points of the set we can know the distance between them and therefore perform a range according to this function. In relation to this, the function defined in Figure 3.1 is not a distance since it does not satisfy the triangular inequality. Therefore, we cannot know which units have a higher socio-economic status than

others.

Figure 3.1: Triangle inequality is not satisfied



3.3 A new proposal: the DL2 composite indicator

In this section we present our proposal for constructing a composite indicator, Distance-Learning or DL2. This composite indicator is the outcome of a weighted ℓ^2 metric[2] in which the weights are computed using iterative ML algorithms. The technique is based on the following concepts. Firstly, the measurement model is formative and works with quantitative data. Secondly, it is based on benchmarking and, thirdly, it is partially compensatory.

Like the review of the DP2 method in the previous section, this section is divided into six parts where we describe how to apply the (DL2) method to determine the values of the composite indicator and overcome the drawbacks of DP2.

To clarify the proposed methodology, the pseudo-code of the proposed algorithm is described in what follows.

- **Normalisation**

  In a first stage, we pre-process the set of data X. The quantitative variables are normalised by re-scaling (or Min-Max) that converts single indicators into a common scale, namely into the interval $[0,1]$. This prevents some indicators from weighing more than others in the composite indicator (DL2). Let $M_j = max(x_{ij})$, $m_j = min(x_{ij})$ denote the maximum and minimum corresponding to each j-single indicator $j \in \{1,\dots,m\}$, then

$$z_{ij} = \frac{x_{ij} - m_j}{M_j - m_j} \tag{3.3}$$

[2]The metric induced by the norm $\| X_i \| = \sqrt{\sum_{i=1}^{\infty} |x_i|^2}$ associated to ℓ^2 spaces is the generalised way of expressing the classical Euclidean distance.

Algorithm 1: Computation of the composite indicator $D(Z_{s\cdot}, Z_{*\cdot})$ with respect to a reference vector $Z_{*\cdot}$.

Data: (Inputs)
$\ell = 1$,
$error = \alpha$,
$\max\_iterations > 1$,
$p - value_1 = 1$,
τ,
$weights = rep(1, n)$,
Result: (0utput)
Composite indicator $D(Z_{s\cdot}, Z_{t\cdot})^{(\ell)}$
Initialisation: Compute $D(Z_{s\cdot}, Z_{t\cdot})^{(0)}$ and the weights $\{\omega_1^{(0)}, ..., \omega_m^{(0)}\}$;
repeat
  Compute $D(Z_{s\cdot}, Z_{t\cdot})^{(\ell)}$ using the weights $\{\omega_1^{(\ell-1)}, ..., \omega_m^{(\ell-1)}\}$
  Compute the next weights $\{\omega_1^{(\ell)}, ..., \omega_m^{(\ell)}\}$ of the composite indicator $D(Z_{s\cdot}, Z_{t\cdot})^{(\ell)}$
  Apply Kendall correlation to $D(Z_{s\cdot}, Z_{t\cdot})^{(\ell)}$ $D(Z_{s\cdot}, Z_{t\cdot})^{(\ell-1)}$
  Compute $p - value^{\ell}$ (null hypothesis of no association)
  Compute Kendall $\widehat{\tau}^{(\ell)}$
  $\ell = \ell + 1$;
until $\ell \leq$ max_iterations or ($\widehat{\tau}_\ell \leq \tau$ and $p - value_\ell < error$);

if the single indicator has positive polarity and

$$z_{ij} = \frac{M_j - x_{ij}}{M_j - m_j} \tag{3.4}$$

if the single indicator has negative polarity. Let Z denote the $n \times m$ matrix whose columns are the single standardised indicators.

- **DL2 formula**

Definition 3.1 *Let $D : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^+ \cup \{0\}$ be a map and $\omega_j \in \mathbb{R}^+$ for all $j \in \{1, \ldots, m\}$. We define DL2 as follows:*

$$D(Z_{s\cdot}, Z_{t\cdot}) = \left(\sum_{j=1}^{m} |z_{sj} - z_{tj}|^2 \omega_j \right)^{1/2} \tag{3.5}$$

where s and t are two compared units or observations[3].

- **Fréchet distance.** In a third stage, we compute the FD with respect to the vector reference $Z_{*\cdot} = (z_{*1}, ..., z_{*m})$ as follows:

$$FD(Z_{s\cdot}, Z_{*\cdot}) = \left(\sum_{j=1}^{m} |z_{sj} - z_{*j}|^2 \right)^{1/2} \tag{3.6}$$

[3] This formula provides the distance between two observations for any normalisation of data. Note that with the normalisation chosen in this article, z_{*j} is the null vector.

The FD does not take into account the overlap of information that may exist between the single indicators; however, it provides a first approximation to the final composite indicator. The FD depicts the initial seed of our supervised algorithm. Note that if the single indicators were independent of each other, the information they contribute to the composite indicators would not overlap in conceptual terms. In such a case, the methodology presented in this article would simply consist of calculating the FD.

- **Selecting the best set of disjoint polynomials.** In a fourth stage, when the iteration is equal to one, the algorithm searches functional relationships between the set of single indicators and the FD. To perform this task, we use multivariate adaptive regression splines (MARS) (Friedman, 1991).MARS estimates different regression slopes at different intervals for each predictor and selects the best set of disjoint polynomials between the composite indicator and the set of single indicators.

Unlike $DP2$, which selects weights using ordinary linear regression (OLS) as a hinge to establish relationships between the composite indicator and single indicators, MARS is a non-parametric method that extends the model by looking for non-linear interactions between the single indicators and the composite indicator. Moreover, the algorithm is insensitive to the basic assumptions of linear regression, which enables it to detect irrelevant indicators in the model (Kuhn, 2008). To perform this task, the model explicitly includes polynomial parameters or step functions.

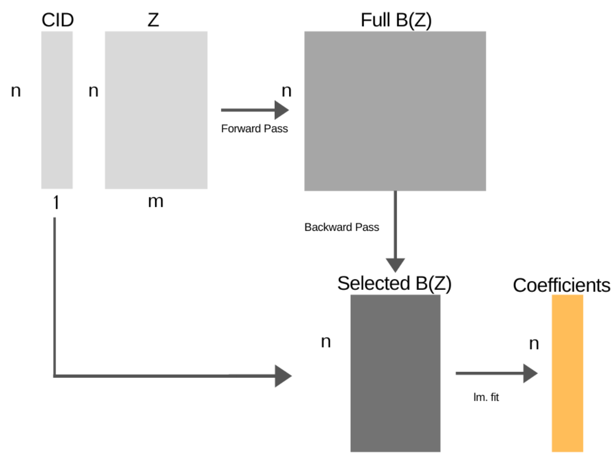
Let us shorten $D_i^{\ell-1} = D(Z_{i\cdot}, Z_{*\cdot})$ as the composite indicator `DL2` computed in the $\ell - 1$ iteration $\ell \in \{1, \ldots, m\}$ (for more details, see Algorithm to determine `DL2` below) associated to the i observation or unit and the target vector. The set of disjoint polynomials $B(z_{ih})$ are functions that depend on the respective variables z_{ih}, where each $B(z_{ih})$ $h \in \{1, \ldots, H\}$ can be written as $B(z_{ih}) = \max(0, z_{ih} - c)$ or $B(z_{ih}) = max(0, c - z_{ih})$, where c is a threshold value and H represents the number of explanatory indicators, which includes interactions of the predictor variables. The final model is a combination of the generated set of disjoint polynomials including possible interactions between predictors. The MARS model can be written as follows:

$$D_i^{\ell-1} = \beta_0 + \beta_1 B(z_{i1}) + \ldots + \beta_m B(z_{iH}) + \varepsilon_i \tag{3.7}$$

where the coefficients β_j are estimated by minimising the sum of squared errors and the error term ε follows a normal distribution $N(0, \sigma^2)$.

To select the best model, an ML algorithm is used. A first step is to start with a model containing only the β_0 intercept and iteratively add disjoint polynomials to the model. During the training process, MARS selects new disjoint polynomials that minimise the sum of squared (residual) errors (SSE) using OLS. The forward step continues until a matrix of predictors $B(Z)$ is completed. In general, at the end of this process, $B(Z)$ has a much larger number of columns than the original single indicators $H > m$ (Figure 4.1). The second phase of this algorithm uses the backward stepwise process. The functions that contribute least to the fit are removed through 10-fold cross-validation (CV) (Craven and Wahba, 1979; Friedman and Silverman, 1989) until the best sub-model is found. The entire procedure is executed using the R EARTH package (“Notes on the EARTH package”, Stephen Milborrow, personal notes, September 15, 2020). The steps explained above are illustrated in Figure 4.1.

Figure 3.2: Overview of EARTH steps



- **Computing the weights of DL2.** In a fifth stage, we compute the weights of the DL2 (ω_j in Equation 3.5) corresponding to the variable importance (VI) function using the simple feature importance ranking measure (FIRM) of the VIP package (Greenwell et al., 2018). This tool provides a standardised model-based approach for measuring a single indicator's importance across the growing spectrum of supervised learning algorithms. This function allows us to classify the single indicators in terms of their relative influence on the predicted DL2. Roughly speaking, VI provides a measure of the strength of the relationship between single indicators. Thus, VI quantifies the relative "flatness" of the effect of each feature z_j with respect to the other indicators $\{z_1, \dots, z_{j-1}, z_{j+1}, z_{j+1}, \dots, z_k\}$, considering that the estimation is evaluated on the functional relationship $\widehat{B}$ obtained in the previous step. The normalised scores provided by VI have a similar role as the correction factor $1 - R^2_{j,\dots,1}$ in the DP2 methodology.

  For example, if VI assigns a value close to 0 to the first single indicator, the indicator will contribute very little information to the model. When a score is equal to zero, we assign the value $min(\omega_j)/m$ to the corresponding weight, where $min(\omega_j)$ is the minimum nonzero weight and m is the number of single indicators [4]. Conversely, if VI assigns a value of 1 to a single indicator, then this indicator will be the most relevant in the model and the assigned weight will be 1.

- **Algorithm to determine DL2.** We now analyze the iterative method of calculation. To compute the composite indicator, we begin by calculating the FD, for which all single indicators have the same relevance, i.e. $\omega_1 = \dots = \omega_m = 1$. The FD will be the first composite indicator for iteration one and is called $D^{(0)}$. Assuming $D^{(0)}$ as a response variable, we calculate the set of disjoint polynomials $\widehat{B}(Z)$ that best approximates the single standardised indicators. We then compute the variable importance with respect to $D^{(0)}$ and obtain the first weights $\{\omega_1^{(1)}, \dots, \omega_m^{(1)}\}$. The metric or distance (Equation 3.5) allows us to calculate a new composite indicator with the previous weights we call $D^{(1)}$. This iterative process generates a succes-

[4] Note that the weigths are not completely eliminated to ensure our DL2 continues to be a metric or distance.

sion of weights $\{\omega_1^{(\ell)}, \ldots, \omega_m^{(\ell)}\}_{\ell=1}^m$ and composite indicators $\{D^{(\ell)}\}_{\ell=1}^m$.

Note that the weights of iteration ℓ have been calculated with respect to the composite indicator of iteration $\ell - 1$. Therefore, it is necessary to decide when the algorithm should be stopped. Each composite indicator induces a rank into the observations. We use the non-parametric hypothesis test (the Kendall rank correlation coefficient or Kendall's τ coefficient) as a measure to compute the ordinal association between $D^{(\ell-1)}$ and $D^{(\ell)}$. This tool provides the rank similarity of the two composite indicators. In the case that there are no ties between the indicator $D^{(\ell-1)}$ and $D^{(\ell)}$, the correlation coefficient is expressed as follows

$$\tau = \frac{\sum_{i<j}(sign(D_j^{(\ell)} - D_i^{(\ell)}) * sign(D_j^{(\ell-1)} - D_i^{(\ell-1)}))}{D}$$

where $D = n(n-1)/2$. In the case of ties, the expression D is somewhat more complex (see Kendall, 1976, chapter 3). Thus, under the null hypothesis of no association, the τ statistic provides. Consequently, a τ in the interval $[0.9, 1]$ indicates strong agreement between two consecutive composite indicators. Moreover, the algorithm stops when we reject the null hypothesis and the τ statistic is greater than 0.9. Intuitively, this result will confirm that the weights added to the model have not changed the ranking of the composite indicator.

3.4 Properties of the DL2

The methodology presented in this article is based on a weighted distance metric. The mathematical properties of this type of structure hold and are listed below.The proofs of all statements can be found in the Appendix.

i Map D defines a metric or distance. Firstly, the distance between two observations is positive or zero. In the latter case, the two units must be the same. Secondly, the distance from observation A to observation B is the same as the distance from B to A. Thirdly, according to Figure 3.1, the distance from B to T plus T to A must be greater or equal than the distance from B to A.

ii Map D is well defined. The composite indicator assigns a unique interpretation or value for each unit or observation.

iii Monotonicity. If a single indicator (with positive polarity) increases (decreases) while keeping the others constant, the computed index should increase (decrease).

iv Invariance by origin and scale changes. The standardisation is invariant by origin and scale changes.

v Transitivity. The composite indicator satisfies that if A is at least as great as B, and B is at least as great as C, then A is at least as great as C.

vi Homogeneity. A proportional increase (decrease) in all the single indicators generates a proportional increase (decrease) of equal magnitude in the composite indicator.

vii Symmetry. Permutations of the simple indexers lead to the same result.

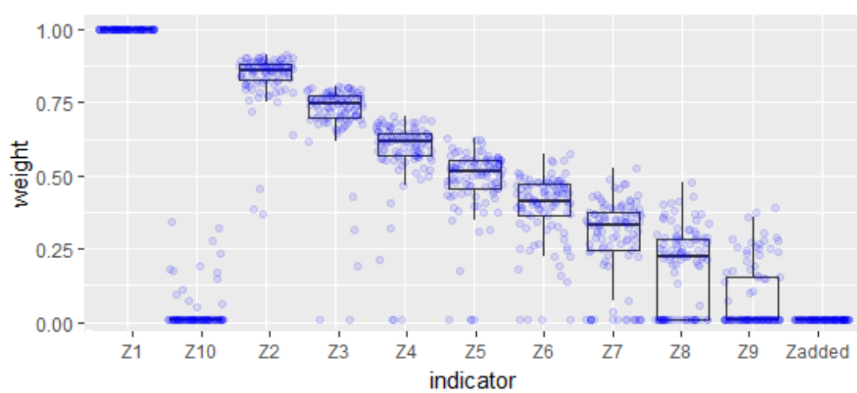
Teorema 3.2 *Let $D : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be a map. The composite indicator distance defined as Equation 3.5 satisfies the properties listed above.*

3.5 Robustness of the DL2

This section focuses on the robustness assessment of the DL2 in terms of its capacity to produce correct and stable measures. We develop three strategies to test whether the composite indicators built with the DL2 method are able to deal with adversities that may arise due to the selection of single indicators, the way single indicators are introduced in the model, and changes in the number of units analysed over time. To this end, we use Monte Carlo simulations. We perform a data set Z where 10 single indicators $\{Z_{\cdot 1}, \ldots, Z_{\cdot 10}\}$ are analysed and 400 random uniform observations are generated for each single indicator ($i \in \{1, \ldots, 400\}$).

The first strategy to validate the DL2 is to examine the selection of single indicators. The selection of single indicators is a fundamental step that is closely related to the concept to be modelled and the choice must be supported by the theoretical and empirical literature (Maggino, 2017). Nevertheless, some of the single indicators could be a perfect or almost perfect linear combination of the rest. We want to know how this situation could affect the values of the DL2 composite indicator. Let $\{Z_{\cdot 1}, \ldots, Z_{\cdot 10}\}$ be the single indicators generated through uniform random variables and $Z_{\cdot Added} = \sum_{j=1}^{10} \alpha_j Z_{\cdot j}$ a linear convex combination. We built a Monte Carlo procedure on 100 random convex linear combinations. For each linear combination, we calculated the corresponding weights using the proposed methodology. Figure 4.3 shows that the *Zadded* single indicator is irrelevant for the computed iterations.

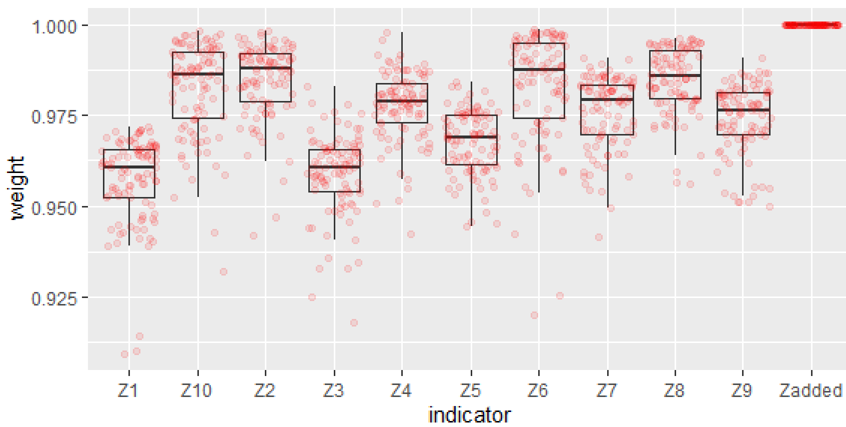
Figure 3.3: Weights of single indicators in 100 iterations for DL2 ($n = 400$)



However, we observe that if we perform the same computation, but this time for the distance P2, the variable Zadded is not eliminated and is also the most relevant (Figure 4.4).

The second strategy to validate the DL2 involves checking whether the way single indicators are introduced in the model alters the values of the composite indicator in the units. Given that our proposal relies on VI to estimate the functional relationships between the composite indicator

Figure 3.4: Weights of single indicators in 100 iterations for DP2 ($n = 400$)

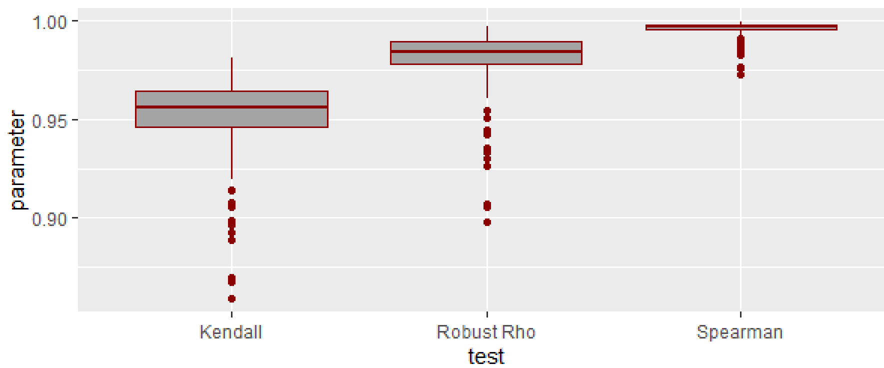


(DL2) and the set of single indicators, it is reasonable to test whether permutations of the single indicators change their weights or importance in the composite indicator. Therefore, we calculated the DL2 corresponding to 100 different permutations according to the rank of the single indicators $Z_{\cdot j}$ $j \in \{1,\dots,10\}$. The results provided the same weighted scheme, producing equal values of DL2. Therefore, the algorithm is invariant to how single indicators are introduced in the model and the calculated weights are not altered by these permutations (results are available upon request).

Finally, the third strategy checks to what extent an uncertain number of units (in the future) would affect the units? rank. This is a situation that can arise when companies that are part of a database in one year leave the market in other years (e.g. due to closure), when there are administrative changes that affect the number of units (e.g. recognition of new municipalities) or simply when for one year or several years the information of the single indicators is not available for some of the units analysed. When studying the robustness of a method for constructing stable composite indicators, it is essential to check if the method will still be able to provide a reliable and comparable measurement over time in the event any of these situations occur.

To check whether composite indicators built with the DL2 method are able to deal with such situations, we analyse the variability in the range of scores of the DL2 when some units are removed at random. To perform this test, we again use a Monte Carlo procedure. Firstly, 10 random observations or units are deleted from the database and the composite indicator DL2 is calculated with the remaining 390 observations. Secondly, we use the original database with the 400 units to calculate the DL2. Once computed, we remove the DL2 values corresponding to the same 10 observations that were eliminated in the previous step. As a result of the two steps, two composite indicators are obtained and compared using Spearman's, Kendall's and robust range correlation statistics. Assuming that the probability of a type I error is $\alpha = 0.05$, a 100 data set was analysed, for which 10 observations were randomly deleted. Figure 4.5 shows the three correlation tests computed. All the tests show evidence of a strong correlation between the two composite indicators. Thus, the random elimination of observations does not produce significant differences in the units' ranking in the composite indicator. Consequently, the composite indicators constructed with the DL2 method are stable over time even if there are changes in the units analysed.

Figure 3.5: Results of simulations of DL2 when 10 observations are randomly deleted ($n = 390$)



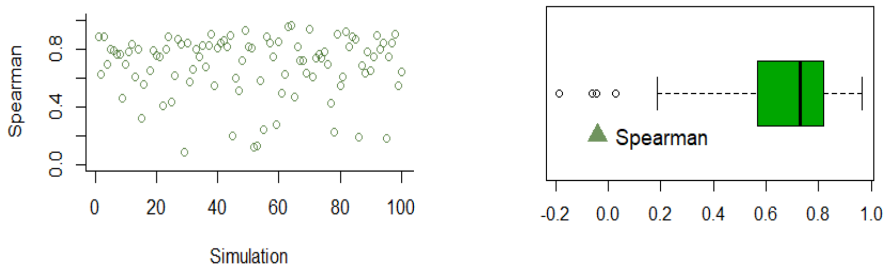
3.6 Comparison between DL2 and DP2 methods

In this section, we compare the DL2 and DP2 methods in terms of the results that would be achieved in an empirical application and the mathematical properties that verify the composite indicators constructed with both methodologies.

For the first comparison, a Monte Carlo algorithm was designed to analyse the ranking of the units in terms of the composite indicator values using both methods. One hundred databases with 400 observations or units and 10 single indicators were generated. The single indicators follow a normal distribution with randomly chosen mean and variance parameters. In addition, each database was designed to have strong, intermediate and weak correlations between the single indicators to ensure that the algorithm presents extreme cases and to highlight the differences in the methodologies. Figure 4.6 shows the results of the comparison.

The boxplot of Spearman's test (part b of Figure 4.6) indicates that 75% of the simulations have a Spearman's correlation between the values of both composite indicators (DL2 and DP2) greater than 0.6, but only 25% of them showed correlations above 0.85. Two plausible explanations can be given for these divergences in the ranking of both composite indicators, which is why Spearman's correlations appear so low. Firstly, as analysed in a previous section, the methodology presented in this paper (DL2) introduces improvements that allow solving the collinearity problems that may occur in some databases, whereas DP2 is unable to detect single indicators with multicollinearity problems. Secondly, some simulations in both methods can reach the maximum number of iterations introduced to stop the algorithm and, therefore, the optimal solution is not reached. In these cases of non-convergence, the rankings of the two methods show a larger difference.

Additionally, for the second comparison between the DL2 and DP2 methods, we consider the bigsalary data set (Baser and Pema, 2003) available in the Wooldridge R software package. The total number of observations is 246. The variables considered for this comparison are the identifier of each faculty member (id), the annual salary measured in dollars (salary), an indicator of publications (publiindex), and the standardised total article pages (artpages). Figure 3.7 shows high correlations among the variables analysed. These correlations make it necessary to eliminate redundant information to avoid overestimates in the calculated composite indicator. We compute DL2 and DP2. It is worth highlighting that DP2 calculates weights that are not a convex linear com-



(a) Simulations Spearman test (b) Boxplot Spearman test

Figure 3.6: Montecarlo simulation to compare DL2 and DP2 composite indicators

bination. If this were the case, the weights would correspond to 0.3102296, 0.3623602, 0.3274103, that is, they would be very similar to those provided by DL2.

Figure 3.7: Correlogram with ggpairs ($n = 246$)

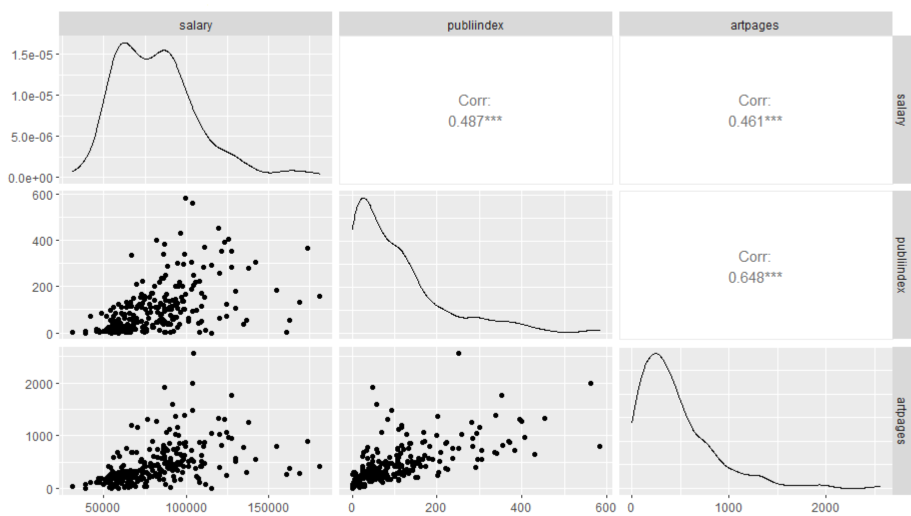


Table 3.1: Weights of aggregation methods

	Salary	Pubiindex	artpages
DL2	0.3684520	0.2296465	0.4019016
DP2	0.7685526	0.8976993	0.8111156

We also tested for associations between the DL2 and DP2 indicators using Pearson's product-moment correlation coefficient, Kendall's τ or Spearman's ρ. The null hypothesis for which the parameter corresponding to each test is zero is rejected, obtaining correlations of 0.9544713, 0.8417787 and 0.9659782, respectively. Table 4.1 shows the first 10 observations where $dell2$ denotes the DL2 composite indicator, $dp2$ denotes the $DP2$ composite indicator and $dell2\_rank$ and $dp2\_rank$

denote the rank obtained from the worst numbering in each of the methods.

Table 3.2: First 10 observations ranked by DL2

id	salary	publiindex	artpages	dell2	dell2_rank	dp2	dp2_rank
154	30813.67	2.24	29.00	0.05	1	0.06	1
17	38770.00	0.00	0.00	0.19	2	0.23	2
139	38934.50	4.86	65.00	0.22	3	0.38	3
182	45533.33	6.87	105.00	0.39	4	0.64	4
146	45319.00	15.19	110.50	0.40	5	0.72	8
54	47113.00	4.29	105.00	0.43	6	0.66	5
148	41748.33	71.44	112.00	0.45	7	1.14	35
164	49890.00	7.51	71.00	0.47	8	0.72	9
48	46492.00	10.67	193.50	0.49	9	0.84	15
122	50971.00	9.38	129.00	0.53	10	0.86	17

The third comparison between the DP2 and the proposed DL2 methodologies focuses on the mathematical properties a composite indicator should fulfil to assess its goodness of fit. Table 4.2 summarises the comparison between DL2 and DP2 in terms of these properties.

In a previous section we showed that since the DL2 is a weighted metric distance, it satisfies the following seven properties: distance, well-defined, monotonicity, invariance by origin and scale changes, transitivity, homogeneity and symmetry. To these seven properties, we add the concept of exhaustiveness. Exhaustiveness refers to the fact that a composite indicator should take full advantage, and in a useful way, of the information provided by the single indicators. In this vein, a composite indicator is better than another one if it provides more useful information about the phenomenon being studied, but it must also be able to eliminate duplicate information (Pena Trapero, 1977, 2019; Zarzosa Espina, 1996). Composite indicators built with DL2 are exhaustive because the weights of the single indicators are computed according to their relevance through VI scores, and the model is able to avoid overlapping information.

As regards the composite indicators built with the DP2 method, it should be noted that they fulfil all the properties except for distance (in all cases), existence and homogeneity. These weaknesses have been analysed in a previous section. Composite indicators constructed with DP2 would not be a metric or distance if there were a collinearity problem. In such a case, $R^2_{j,\dots,1} = 1$ for some $j \in \{1,\dots,m\}$, thus indicating that statement (i) in section 3.2 is not satisfied. If the standard deviation of a single indicator were equal to zero (i.e. the single indicator has the same value in all the units), problems of existence could occur since Equation 3.1 would not be well defined. Let α be a positive real number. The non-homogeneity would be due to the fact that $DP2(\alpha X_{i\cdot}) = DP2(X_{i\cdot})$. Lastly, it is worth noting that DP2 would be considered exhaustive, although it only detects relationships between single and composite indicators of a linear nature.

Table 3.3: Comparison of aggregation methods

	Distance	Existence	Monotonicity	Invariance
DP2	✖	✖	✔	✔
DL2	✔	✔	✔	✔
	Transitivity	**Homogeneity**	**Symmetry**	**Exhaustiveness**
DP2	✔	✖	✔	✔
DL2	✔	✔	✔	✔

3.7 Conclusions

The increasing use of composite indicators in economics and the social sciences is warranted because they allow comparisons to be made between units (companies, territories, etc.) and assess the progress or evolution of units over time (Maggino, 2017; Mazziota and Pareto, 2017; Sánchez et al., 2018). Consequently, for composite indicators to be effective, they must be developed with robust methods that ensure that the two objectives of benchmarking and stability over time are achieved. However, the mathematical approach in which the most widely used techniques to build composite indicators are grounded (i.e. arithmetic and geometric means, PCA and DEA) does not enable addressing these issues in a reliable way. Indeed, the weighting and aggregation aspects of these techniques have received much criticism in the most recent literature (see Becker et al., 2017; Greco et al., 2019; Jiménez-Fernández & Ruiz-Martos, 2020; Keogh et al., 2021).

In addition to weighting scheme and aggregation, in this paper we showed that none of these techniques provide a mathematical structure for analysing results through a metric or distance. Consequently, it is unfeasible to establish proximity or distance between the units analysed, so that both quantitative and ordinal comparisons (unit rankings) lack a solid basis. This paper has proposed a new method for building composite indicators called `DL2`. This method is based on the mathematical concept of distance or metric that enables comparisons between the units being studied. Its application is intended for quantitative data in formative measurement models in the context of compensatory aggregation.

Our proposal took as a starting point the Distance P2 or DP2 method developed by Pena Trapero (1977) given its remarkable advantages studied in a previous section: it provides a mathematical structure (except in extreme cases) that enables the units to be ranked according to distance, it avoids linear overlapping information between single indicators and the composite indicator, it fulfils most of the mathematical properties required to assess goodness of fit, and it is quite versatile, thus allows the analysis of a wide range of multidimensional phenomena (see, for instance, Cuenca-García et al., 2019; Martin et al., 2020; Montero et al., 2010; Sánchez et al., 2018; Sánchez and Ruiz-Martos, 2018).

Likewise, we identified the limitations of the DP2 method and improved them by taking advantage of ML techniques, as well as the growing computational capacity. Our improvements included, firstly, correcting cases in which DP2 is not a metric since the defined weights are not always strictly positive. Secondly, because the DP2 method inherently relies on linear models, it does not behave efficiently when single indicators are poorly correlated with the composite indicator. The proposed `DL2` method corrects this weakness by means of unsupervised ML algorithms.

From the composite indicator generated by the unweighted metrics, the algorithm optimizes the best functional relationship (not necessarily linear) between the composite indicator and the single indicators. By means of ML, `DL2` ranks the single indicators in order of importance by assigning weights to the metric based on this relationship. The algorithm stops when the order of the units (in terms of composite indicator values) remains unchanged.

We also analysed the mathematical properties of our proposed method to study its goodness of fit and concluded that it is a distance, it is well-defined, and it satisfies the properties of monotonicity, invariance by origin and scale changes, transitivity, homogeneity and symmetry. To the best of our knowledge, this kind of analysis has been scarcely addressed in the literature. Furthermore, we compared the DP2 to our method and identified the properties that DP2 might not fulfil in some cases. In this regard, the method we have proposed overcomes these weaknesses of the DP.

Finally, the Monte Carlo simulations and real data set confirm that the proposed methodology `DL2` is robust for building composite indicators. The results of the composite indicator or unit rankings remain stable over time, even when the number of analysed units changes. The method detects and eliminates multicollinearity problems among the single indicators. Likewise, the weighting scheme is not altered by permutations in the order in which the single indicators are computed. The requirement of robustness should be a mandatory step in any composite indicator proposal, since the results can guide public decisions regarding the allocation of economic resources, which can be scarce and susceptible to alternative uses. Nevertheless, little attention has been paid to this step in empirical applications (Greco et al., 2019).

3.8 Appendix

Proofs of the mathematical properties listed in 3.8.

i Map D defines a metric or distance. Firstly, the distance between two observations is positive or zero. In the latter case, the two units must be the same. Secondly, the distance from observation A to B is the same as the distance from B to A. Thirdly, according to Figure 3.1, the distance from B to T plus T to A must be greater or equal than the distance from B to A. Verification is immediate for all statements except for triangular inequality (statement 3).
Proof i

$$
\begin{aligned}
D(Z_{s\cdot}, Z_{p\cdot}) &= \left(\sum_{j=1}^{m} |z_{sj} - z_{pj}|^2 \omega_j\right)^{1/2} = \left(\sum_{j=1}^{m} |z_{sj}\omega_j^{1/2} - z_{pj}\omega_j^{1/2}|^2\right)^{1/2} \\
&= \left(\sum_{j=1}^{m} |z_{sj}\omega_j^{1/2} - z_{tj}\omega_j^{1/2} + z_{tj}\omega_j^{1/2} - z_{pj}\omega_j^{1/2}|^2\right)^{1/2} \\
&\overset{*}{\leq} \left(\sum_{j=1}^{m} |z_{sj} - z_{tj}|^2 \omega_j\right)^{1/2} + \left(\sum_{j=1}^{m} |z_{tj} - z_{pj}|^2 \omega_j\right)^{1/2} \\
&= D(Z_{s\cdot}, Z_{t\cdot}) + D(Z_{t\cdot}, Z_{p\cdot}).
\end{aligned}
$$

in which the inequality (*) is obtained through Minkowski and Holder inequality.

ii Map D is well defined. The composite indicator assigns a unique interpretation or value for each unit or observation.
Proof ii
For any $Z_{i\cdot}$, $i \in \{1, \ldots, n\}$, map D exists and belongs to $\mathbb{R}^{+} \cup 0$.

iii Monotonicity. If a single indicator (with positive polarity) increases (decreases) while keeping the others constant, the computed index should increase (decrease). `DL2` is monotone.
Proof iii
Let $D(Z_{i\cdot}) = (z_{i1}, \ldots, z_{im})$ be the m-vector of an observation corresponding to the i observation and let $D(Z_{*\cdot})$ be the vector reference. Without loss of generality, let us assume that for $j = 1$ $z_{*1} < z_{i1} < z_{i1} + \varepsilon$, where z_{*1} is the best scenario and $varepsilon$ is some positive constant, then $|z_{i1} - z_{*1}| < |z_{i1} + \varepsilon - z_{*1}|$.

$$
\begin{aligned}
D(Z_{i\cdot}, Z_{*\cdot}) &= \left(\sum_{j=1}^{m} |z_{ij} - z_{*j}|^2 \omega_j \right)^{1/2} \\
&= \left(|z_{i1} - z_{*1}|^2 \omega_1 + |z_{i2} - z_{*2}|^2 \omega_2 + \ldots + |z_{im} - z_{*m}|^2 \omega_m \right)^{1/2} \\
&< \left(|z_{i1} + \varepsilon - z_{*1}|^2 \omega_1 + |z_{i2} - z_{*2}|^2 \omega_2 + \ldots + |z_{im} - z_{*m}|^2 \omega_m \right)^{1/2} \\
&= D(Z'_{i\cdot}, Z_{*\cdot})
\end{aligned}
$$

where $Z'_{i\cdot}$ is the m-dimension vector whose first coordinate corresponds to the increment $z_{i1} + \varepsilon$.
Conversely, assuming that the Z_{1*} single indicator has negative polarity, $z_{i1} < z_{i1} + \varepsilon < z_{*1}$, where z_{*j} is the worst scenario is some positive constant, then $|z_{i1} - z_{*1}| > |z_{i1} + \varepsilon - z_{*1}|$.

$$
\begin{aligned}
D(Z_{i\cdot}, Z_{*\cdot}) &= \left(\sum_{j=1}^{m} |z_{ij} - z_{*j}|^2 \omega_j \right)^{1/2} \\
&= \left(|z_{i1} - z_{*1}|^2 \omega_1 + |z_{i2} - z_{*2}|^2 \omega_2 + \ldots + |z_{im} - z_{*m}|^2 \omega_m \right)^{1/2} \\
&> \left(|z_{i1} + \varepsilon - z_{*1}|^2 \omega_1 + |z_{i2} - z_{*2}|^2 \omega_2 + \ldots + |z_{im} - z_{*m}|^2 \omega_m \right)^{1/2} \\
&= D(Z'_{i\cdot}, Z_{*\cdot})
\end{aligned}
$$

On the other hand, an increase in a single indicator with negative polarity must generate a decrease in the composite indicator, responding positively to a positive change in any indicators and negatively to a negative change. The proof is symmetric with respect to the positive polarity.

iv Invariance by origin and scale changes. The standardisation tool is invariant by origin and scale changes.
Proof iv
Let $M_j = max(x_{ij})$, $m_j = min(x_{ij})$ denote the maximum and minimum corresponding to each j-single indicator, then

$$
z_{ij} = \frac{x_{ij - m_j}}{M_j - m_j} \tag{3.8}
$$

for each $j \in \{1, \ldots, m\}$. It is sufficient to check that the standardisation is invariant by change of scale for each single indicator. Let $v_{ij} = \alpha x_{ij} + \beta$ denote a origin and scale changes, where

α is a positive real number and β any real number. For all $j \in \{1,\dots,m\}$

$$\min\{v_{ij}\} = \min\{\alpha x_{ij} + \beta\} = \alpha \min\{x_{ij}\} + \beta.$$
$$\max\{v_{ij}\} = \max\{\alpha x_{ij} + \beta\} = \alpha \max\{x_{ij}\} + \beta.$$

hence,

$$\frac{v_{ij} - \min\{v_j\}}{\max\{v_j\} - \min\{v_j\}} = \frac{\alpha x_{ij} + \beta - \min\{\alpha x_{ij} + \beta\}}{\max\{\alpha x_{ij} + \beta\} - \min\{\alpha x_{ij} + \beta\}} = \frac{x_{ij-m_j}}{M_j - m_j} = z_{ij} \tag{3.9}$$

Therefore, it is immediate to check that the standardisation is invariant by origin and scale changes.

v Transitivity.
Proof v
Let $Z_{s\cdot}, Z_{t\cdot}, Z_{\ell\cdot} \in \mathbb{R}^+$ be three observations and let $Z_{*\cdot} \in \mathbb{R}^+$ be the vector reference. We assume that $D(Z_{s\cdot}, Z_{*\cdot}) < D(Z_{t\cdot}, Z_{*\cdot})$ and $D(Z_{t\cdot}, Z_{*\cdot}) < D(Z_{\ell\cdot}, Z_{*\cdot})$ then $D(Z_{s\cdot}, Z_{*\cdot}) < D(Z_{\ell\cdot}, Z_{*\cdot})$.

vi Homogeneity. A proportional increase (decrease) in all the single indicators generates a proportional increase (decrease) of equal magnitude in the composite indicator.
Proof vi
Let α be a real positive constant. Then

$$D(\alpha Z_{s\cdot}, \alpha Z_{t\cdot}) = \left(\sum_{j=1}^{m} |\alpha z_{sj} - \alpha z_{tj}|^2 \omega_j\right)^{1/2} = \alpha D(Z_{s\cdot}, Z_{t\cdot}). \tag{3.10}$$

vii Symmetry.
Proof vii
In the proposed method, the value of the composite indicator does not depend on the rank of the indicators introduced in the $n \times m$ X matrix data.

Chapter 4

European Union Cohesion Policy: socio-economic vulnerability of the regions and the COVID-19 shock

Abstract

The European Union Cohesion Policy for the period 2021-2027 focuses on five goals to make the European Union smarter, greener, more connected, more social and closer to citizens. However, a macroeconomic index is proposed as the predominant criterion for allocating the Structural Funds among regions. In this paper, we hypothesise that it is possible to take into account new, complementary criteria that better reflect citizens? quality of life. To that end, we build a composite index of socio-economic vulnerability for the 233 regions. The results show that following our multidimensional approach for allocating the Structural Funds, there are remarkable differences in the maps of priority regions. In addition, the COVID-19 pandemic represents a threat to well-being. Are all regions equally exposed to COVID-10 in terms of their socio-economic vulnerability? To address this issue, we estimate multilevel models which indicate that country characteristics interact with regions? characteristics to alter patterns of vulnerability. More specifically, increases in government expenditures in education and an improvement in political stability would reduce the regional vulnerability or foster the capacity for resilience, whereas increases in poverty would be associated with greater vulnerability. Likewise, more vulnerable regions would be the most exposed to the negative socio-economic effects of COVID-19. However, it is remarkable that several regions of Sweden and Finland would be among the group of regions whose socio-economic vulnerability would be the most negatively affected.

4.1 Introduction

The main objective of the European Union (EU) Regional Policy, or Cohesion Policy, is to reduce the disparities between the levels of development of the regions and the backwardness of lagging regions. The EU Cohesion Policy for the 2021-2027 Multiannual Financial Framework aims at fostering a modernised regional development and cohesion policy focusing on five political goals so that the EU becomes: (1) smarter, through innovation and digitisation, (2) greener, (3) more connected, (4) more social and (5) closer to citizens (European Commission, 2018). The EU will dedicate 34% of its budget over 2021-2027 to cohesion and values, that is, economic, social and territorial cohesion and investment in competitiveness, people and values (European Commission, 2020a). This is the item that will receive the highest amount of commitment appropriations. Structural Funds are the main source of funding to implement the EU Cohesion Policy.

These guidelines represent significant challenges for the design of the regional development policies within the scope of "beyond GDP", according to which the European Commission should develop several indicators that complement the gross domestic product (GDP) to support policy decisions through more comprehensive information (Commission of the European Communities, 2009). The EU opts for the increasingly accepted train of thought, stressing that GDP is insufficient to analyse the overall development and progress of society, and the measurement of regional development has to struggle with the multidimensional nature of well-being (O'Donnell et al., 2014; Stiglitz et al., 2018; Van den Bergh, 2009). However, a single macroeconomic index is again proposed as the predominant criterion for allocating the Structural Funds among the regions in 2021-2027.

In this paper, we hypothesise that new complementary criteria could be taken into account in line with the five goals of EU Cohesion Policy outlined above in order to better reflect the reality on the ground of the regions. With this in mind, the first aim of this paper is to construct a composite indicator of socio-economic vulnerability (SEVI) that synthesises the position of each EU region (NUTS-2 of the 27 Member States) in 2017 with respect to the five goals of the EU Cohesion Policy for 2021-2027[1]

In addition, the COVID-19 pandemic caused by the SARS-CoV-2 coronavirus represents a threat to people's well-being and new public policy challenges. Worldwide, the COVID-19 pandemic is a serious threat to the achievement of the Sustainable Development Goals since it is pushing tens of millions of people back into extreme poverty, putting years of progress at risk (United Nations, 2020a). In the context of the EU, it is foreseeable that COVID-19 will negatively affect the socio-economic development of the regions, as well as the quality of life of people since COVID-19 is impacting on a wide range of aspects: health and subjective well-being, social capital, human capital, product markets, financial markets and public finance (Bittmann, 2021; Bonaccorsia et al., 2020; Fasani & Mazza, 2020; Fetting, 2020; Giovanis & Ozdamar, 2022; Giovannini et al., 2020; Shek, 2021; United Nations, 2020b).

Faced with this situation, in its first annual strategic foresight report, the European Commission describes the first lessons of the COVID-19 crisis and introduces resilience as a new compass for the development of EU policies (European Commission, 2020b). In the report, the European Commission presents resilience dashboards in the socio-economic, green and digital dimensions and proposes further discussion to explore the feasibility of developing a synthetic resilience index. At the financial level, the EU has approved the Next Generation EU (Euro 750 billion) to build

[1]Nomenclature of territorial units for statistics level 2 (NUTS-2) is the classification used in regional statistics and funding allocation which subdivides Member States into regions according to existing national administrative subdivisions and the population thresholds from 800,000 to 3 million inhabitants. The EU Cohesion Policy is designed and monitored at NUTS-2 level.

a more resilient, sustainable and fair Europe through large-scale financial support for investment and reforms. The majority of funds (Euro 672.5 billion) will be allocated to the Recovery and Resilience Facility programme to support public investments and green and digital projects in the crucial first years of the recovery after the pandemic.

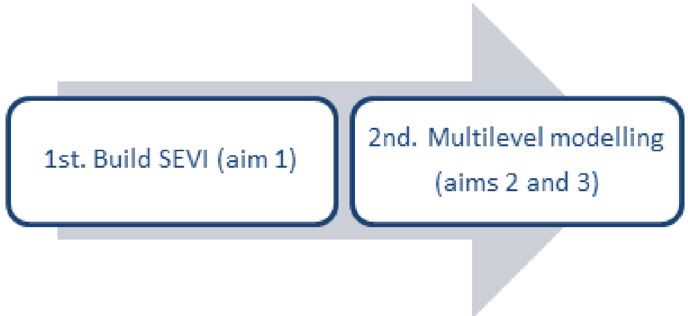
In this scenario, assessing how changes in the environment or covariates of the regions could affect their socio-economic vulnerability is key for the planning of Cohesion Policy in order to determine actions that can increase the resilience of different territories. Accordingly, the second aim of this paper is to check whether country characteristics interact with regions' characteristics to alter patterns of vulnerability. That is, we check if the structure of regions' socio-economic vulnerability is hierarchical and causes a "country effect" or if the socio-economic vulnerability of the regions differs across countries. If this interaction or country effect were confirmed, the third aim of this paper would be to analyse both the idiosyncratic and covariate shocks that COVID-19 might represent for the vulnerability of EU regions. In the context of the EU, formulating objectives 2 and 3 is highly significant due to the existence of a multilevel governance system with central, state and local governments in most of the Member States that assume different competences in matters of public policy affecting citizens' quality of life. If the hierarchical structure were confirmed, multilevel modelling would be a suitable approach to address these two aims since standard estimation techniques could lead to incorrect conclusions (see Goldstein, 2011; Snijders & Bosker, 2012).

To sum up, this paper aims to achieve three objectives. Firstly, in a first stage, we build a composite indicator to study the socio-economic vulnerability of the EU regions in terms of the 2021-2027 Cohesion Policy. Once we have an instrument (SEVI) to analyse the socio-economic vulnerability of EU regions, in a second stage we estimate mixed effects models or multilevel models with SEVI as the dependent variable, which allows us to achieve aims 2 and 3. Figure 1 indicates the two stages of our work.

The main contributions of our paper are twofold. First, we take the concept of vulnerability from other fields, such as poverty and economics, where it is studied at the individual level, and apply it to the level of regions. To this end, we follow a multidimensional approach to identify the factors driving socio-economic fragility and resilience in terms of the 2021-2027 EU Cohesion Policy goals and take into account findings from the first studies on the social and economic effects of COVID-19. Secondly, we exploit the probably little-known potential of multilevel models to identify the regional and country-level characteristics and/or public policies associated with resilient behaviour (via random intercept models) and examine the impact of a shock such as COVID-19 on regional vulnerability (via the random or stochastic part of the models).

The rest of this paper is structured as follows. Section 2 focuses on studying the conceptual framework of the 2021-2027 EU Cohesion Policy and the socio-economic vulnerability, the first step in the construction of a composite indicator, which in turn provides the basis for the selection and aggregation of single indicators. Section 3 presents and justifies the dataset and single indicators used both to build the composite index and to develop the multilevel models. Section 4 describes the empirical strategy to build the composite index SEVI, as well as the multilevel modelling approach to study the shock that COVID-19 could represent. Section 5 presents the main results of our analysis and examines some implications for public policies. Lastly, conclusions are drawn in section 6.

Figure 4.1: Two stages in the study of socio-economic vulnerability of the EU regions



4.2 Conceptual framework

4.2.1 European Union Cohesion Policy, 2021-2027

The five objectives of the Cohesion Policy for the period 2021-2027 are framed in the political guidelines for a strategic long-term vision to achieve the transition towards a green, digital and fair Europe. To do this, the EU must continue to develop as a social market economy, as outlined in the Europe 2020 Strategy (European Commission, 2020b). The social market economy is an integrated social, economic and political order characterised by having a market economic policy and a social policy. In turn, the social policy regulates the market economic policy. The latter is configured as its greatest difference from neoliberalism (European Commission, 2010).

Specifically, the five policy objectives drive investments to foster (European Commission, 2018):

1. A Smarter Europe through innovation, digitisation, economic transformation and support to small and medium-sized businesses.

2. A Greener, carbon free Europe, implementing the Paris Agreement and investing in energy transition, renewables and the fight against climate change.

3. A more Connected Europe, with strategic transport and digital networks.

4. A more Social Europe, delivering on the European Pillar of Social Rights and supporting quality employment, education, skills, social inclusion and equal access to healthcare.

5. A Europe closer to citizens, by supporting locally-led development strategies and sustainable urban development across the EU.

The underlying assumption is that these five priorities are mutually reinforcing: to improve education levels and increase investment in R&D, innovation and digitisation will improve competitiveness and economic growth in a sustainable way, thereby fostering job creation and reducing social exclusion. As is customary in the EU Cohesion Policy, the objectives of economic growth and job creation carry great weight, probably on the erroneous basis that social cohesion will follow from them (Sánchez & Ruiz-Martos, 2018). Accordingly, in the 2014 2020 period, the financial weight of the allocation criteria of the Structural Funds was 86% for relative wealth (per capita GDP) and 14% for labour market, education and demographic factors. However, a qualitative change was introduced for the period 2021-2027. In addition to the above criteria, youth unemployment, migration and greenhouse gas emissions will also be considered for the first time in the distribution of Structural Funds. More specifically, per capita GDP accounts for 81% of regional allocations; 15% of labour market, education and demographics allocations; 3% of migration allocations and 1% of climate change allocations (European Court of Auditors, 2019). The five goals are reviewed below.

Goals 1 and 3 are a continuity of previous planning periods, especially since the 2000 2006 period, when emphasis was placed on investment in R&D (Romer, 1994), human capital (Lucas, 1993), industrial innovation (Grossman & Helpman, 1994) and the provision of infrastructure or public capital (Aschauer, 1989) as drivers of economic growth. These models, inspired by the EU Cohesion Policy, integrated endogenous growth theory and argued that investment in these special categories of capital increased the productivity of all factors and therefore promoted economic growth. Subsequently, the concept of infrastructure was extended to research and innovation. Thus, the Horizon 2020 programme (financial instrument of the Europe 2020 Strategy to develop EU innovation policy since 2014) introduced the concept of research infrastructure (European Commission, 2011). Research Infrastructures are facilities that provide resources and services for research communities to conduct research and foster innovation. This concept aims to integrate research and innovation to promote market-related activities, which leads to a direct economic stimulus (European Commission, 2020c).

Goal 2 focuses on sustainable growth, which was introduced in the Europe 2020 Strategy as one of the pillars of the EU. Compared to other strategies, such as the Lisbon Strategy, Europe 2020 constituted a step forward. Since the publication of the Brundtland Report in 1982, there has been growing awareness of the importance of achieving a balance between the economic, social and environmental subsystems. Sustainable economic growth is understood as a growth rate that can be maintained without creating other significant problems, such as the depletion of resources or environmental problems, especially for future generations. This goal is rooted in the EU?s objective of competitive sustainability and cohesion through a new growth strategy: the European Green Deal. The key aim is to shift towards a sustainable and inclusive economic model, enabled by a broader diffusion and uptake of digital and clean technologies (European Commission, 2021a).

Due to the negative effects of the economic crisis on certain groups (the elderly, youth, women, migrants and lower-skilled workers), goal 4 of the Cohesion Policy focuses on fostering inclusive growth by promoting the European Pillar of Social Rights. In turn, in 2021, and given that the effects of COVID-19 affected these groups more, a new “social rulebook” has been introduced in the European Pillar of Social Rights to enhance social rights and strengthen the European social dimension across all policies of the Union (European Commission, 2021c). The main lines of action that should guide policy decisions in the Member States and their regions, including the programming of the 2021-2027 Cohesion Policy and the national recovery and resilience plans (European Commission, 2021c, p. 10), are aimed at reducing the gender employment gap, decreasing the rate of youth unemployment, reducing early school leaving and fostering higher education. The underlying idea is that special attention needs to be paid to young people and the low skilled (including migrants in both categories), who are more vulnerable to labour market fluctuations. Likewise, the demographic trends of the EU, marked by an ageing society, represent challenges for the principles of the Pillar of Social Rights, which focus on promoting health and care and ensuring that everyone in old age has the right to resources that ensure living with dignity.

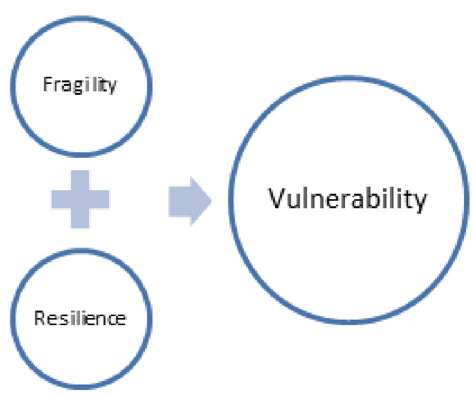
Finally, goal 5 aims at promoting locally-led development strategies and sustainable urban development, with the objective of satisfying local objectives and needs and contributing to the smart, sustainable and inclusive growth of the EU. This local development strategy has also been a key factor in the EU Cohesion Policy since the 2000-2006 period. The approach is largely inspired by local development theories whose basic idea is to identify and enhance competitiveness factors at the local level (see Scott & Garofoli, 2007). COVID-19 has highlighted the importance of strengthening the resilience of urban areas to promote the well-being of inhabitants with challenges such as sustainable mobility and consumption, the treatment of urban waste through recycling, or the need for housing for new urban dwellers (European Commission, 2020b).

4.2.2 Regional socio-economic vulnerability in the European Union

Studies on vulnerability have been carried out in a range of fields. The fields that have probably received the most attention are poverty (Acconcia et al., 2020; Azeem et al., 2016; Gallardo, 2020), climate change, and physical vulnerability to natural disaster (Halkos et al., 2020; Marulanda Fraume et al., 2020) and financial or economic vulnerability (Alessi et al., 2020). Vulnerability is defined in a various ways in the literature (for a review, see Acconcia et al., 2020; Gallardo, 2018; Mina & Imai, 2016), so a crucial step of this study is to define the conceptual framework of socio-economic vulnerability. This is also important because the conceptual framework will determine the empirical strategy of our study.

In general terms, vulnerability refers to the propensity or predisposition to be adversely affected together with the difficulty of reacting. The most recent vulnerability studies encompass a variety of concepts grouped into two broad forms: sensitivity or fragility to suffer harm, and the capacity to cope and adapt or resilience (Azeem et al., 2016; Halkos et al., 2020; Marulanda Fraume et al., 2020). Figure 2 shows this idea.

Figure 4.2: Components of socio-economic vulnerability



Under this framework, socio-economic fragility refers to the predisposition to suffer harm from the disadvantageous conditions and relative weaknesses related to social and economic factors (Cardona, 2004). In this vein, the 2020 Strategic Foresight Report (European Commission, 2020b) identifies groups and areas that have suffered the effects of the pandemic most and face greater difficulties in coping with the effects of the COVID-19 shock. For example, residential care facilities and support services for older people and persons with disabilities were structurally fragile and unprepared to cope with and control the spread of the coronavirus. Other groups that have shown to be more fragile are students from disadvantaged backgrounds because they were less likely to benefit from online learning and lower skilled workers that were more likely to be employed in “contact jobs” with greater exposure to the virus.

On the other hand, resilience is the ability to face shocks and persistent structural changes (e.g. digital transformation, globalisation and climate change) that affect people and society in such a way that current societal well-being or quality of life is preserved (Alessi et al., 2020; Benczur et. al., 2020). Therefore, a resilient society aims to sustain its level of individual and societal well-being in an intergenerational fair distribution, that is, by ensuring current well-being without seriously compromising that of future generations (Manca et al., 2017, p. 6). Adaptation and transformation are key to bouncing forward. In this regard, the 2020 Strategic Foresight Report (European Commission, 2020b) highlights that the EU’s social and economic resilience rests on its population and its unique social market economy. Among the key points to enhance resilience against COVID-19 are access to education and social protection, flexible work arrangements and a highly skilled workforce. Consequently, in the context of the EU 2021-2027 Cohesion Policy, the degree of a region’s socio-economic vulnerability might be estimated by a composite indicator built from a system of single indicators able to take into account these policy goals. At the same time, this system of indicators should allow identifying the socio-economic weaknesses of the regions, as well as defining the social and economical dimensions related to how a region is able to respond to the pressure from these dimensions, and whether it is capable of adapting to those pressures to deliver well-being in a sustainable way. Under this framework, the situation of a region with a greater degree of socio-economic vulnerability might be understood as having greater obstacles or found in a worse position to achieve the Cohesion Policy goals (2021-2027). In short, our premise is that socio-economic vulnerability is a latent variable, since it is a concept or construct which cannot be measured or estimated directly, but rather indirectly using collectable social and economic indicators.

Once we achieve an instrument to analyse the socio-economic vulnerability of EU regions in terms of the 2021-2027 Cohesion Policy goals, the next step is to study how a situation of economic and social stress such as the COVID-19 pandemic could affect regional vulnerability. In this vein, societies that are more resilient to disturbances will also be able to ensure a higher level of well-being or quality of life as the shock will have a less severe impact on them (Alessi et al., 2020; Manca et al., 2017). Taking into account the magnitude and duration of the COVID-19 effects, especially compared to previous experiences such as the SARS outbreak of 2003 (see for instance Lee & McKibbin, 2004; Keogh-Brown & Smith, 2008), it is reasonable to hypothesise that regions’ socio-economic vulnerability will not only be affected by their particular variables or characteristics, but also by the country’s characteristics (for example, public policies at the country level). This region-country interrelation may determine the degree to which a region is affected by the COVID-19 shock. Econometric multilevel modelling is a proper quantitative method to address these issues since it allows incorporating observed variables at both the regional and country levels among the explanatory variables.

4.3 Data and variables

4.3.1 Data and single indicators to build the SEVI

To develop the socio-economic vulnerability index (SEVI) in the EU regions, we use the official statistics of EUROSTAT and OECD at the NUTS-2 level which is the basic unit for the application of regional policies. We work with the most recent regional territorial classification, known as NUTS 2016, which entered into force on 1 January 2018 in accordance with the Commission Regulation (EU) 2016/2066. The overseas NUTS-2 territories have not been taken into account in this study (Ceuta and Melilla in Spain; and Guadeloupe, Martinique, Guyane, La Réunion and Mayotte in France). A total of 233 EU regions or NUTS-2 territories are studied.

In order to develop a system of indicators capable of representing how a region is able to respond to the pressures and challenges of the 2021-2027 Cohesion Policy, we selected 16 single indicators. For the system of indicators to be balanced, eight representative indicators of the socio-economic weakness or fragility of the regions and eight representative indicators of the capacity of the regions to face challenges or structural changes have been chosen. The eight single indicators of fragility have positive polarity, which means that an increase in the indicator could also lead to an increase in socio-economic vulnerability. Conversely, the eight single indicators of resilience have negative polarity, which means that an increase in the indicator could lead to a reduction in vulnerability. Appendix A presents the definitions and technical information of the single indicators. The values of the single indicators have been obtained as the average of the last two available years, including in all cases (except R&D) the year 2017, as is the usual practice in matters of EU Cohesion Policy.

The selection of single indicators has essentially been guided by the five goals set by the European Commission (2018) for the Cohesion Policy 2021-2027 mentioned above, as well as by plans, strategies and projects of the EU approved in the context of the COVID-19 crisis that also use monitoring indicators in areas related to the five goals of the EU Cohesion Policy. Appendix B displays the 16 single indicators with the EU official documents that guided our choice of single indicators indicated in the right column. In any case, our selection has been determined by the availability of statistical information, which is quite scarce at the NUTS-2 level in several areas such as climate change, income inequality and self-reported measures. Table 4.1 shows the descriptive statistics of the single indicators.

Table 4.1: Descriptive statistics of socio-economic vulnerability indicators for the EU27 regions in 2016-2017 (N = 233 NUTS-2)

	Mean	SD	Min	Max	CV	Region-Baseline
Early leavers	10.21	4.87	1.35	27.35	47.72	HR03 - Jadranska Hrvatska
PM2.5	12.89	4.27	4.40	28.28	33.16	PT20 - Regiao Autónoma Acores
Elderly people	9.48	2.11	4.57	15.52	22.28	NL23 - Flevoland
Male unemployment	7.94	4.80	1.85	24.15	60.52	CZ01 - Praha
Female unemployment	8.78	6.95	1.90	39.25	79.22	DE22 - Niederbayern
Youth unemployment	20.39	12.89	3.60	57.15	63.21	DE93 - Lüneburg
Migrant	3.23	4.36	0	38.80	134.91	Regions with negative rate
Assault & crime	0.74	0.52	0.08	4.24	70.61	FRC2 - Franche-Comté
R&D business	0.98	0.96	0	8.06	97.31	DE91 - Braunschweig
R&D state	0.61	0.44	0	2.52	71.31	DEB2 - Trier
Tertiary education	29.12	9.00	11.80	55.00	30.89	PL91 - Warszawski stoleczny
Human resources in technology	31.92	8.38	13.95	54.70	26.24	PL91 - Warszawski stoleczny
Registered community designs	3,591.21	4,000.26	0	24,813.07	111.39	ITH4 - Friuli-Venezia Giulia
Internet	97.33	2.69	87.50	100.00	2.77	Several regions with 100
E-Administration	51.42	20.09	4.50	92.00	39.08	DK01 - Hovedstaden
GDP-Gini	19,782.30	7,905.24	5,459.60	52,512.45	39.96	LU00 - Luxembourg

Note. HR is Croatia, PT Portugal, NL Netherlands, CZ Czech Republic, DE Germany, FR France, PL Poland, IT Italy, DK Denmark, LU Luxemburg.

Next, we discuss the rationale that justifies the relationship between each single indicator and the composite indicator of socio-economic vulnerability (that is, the polarity). We start with the indicators of fragility and then illustrate why they might be considered indicators of fragility based on the literature and EU official documents and reports.

Dropping out of school has negative effects both for individuals and society (unemployment, less lifetime earning, more risk of poverty, higher public spending for social protection, etc.), hence the reduction of the percentage of people who dropped out of primary and secondary studies until a maximum of 10% is a target set out in the European 2020 Strategy in order to attain social cohesion in the EU (European Commission, 2010). The new "social rulebook" of the European Pillar of Social Rights (European Commission, 2021c) identifies the reduction of early school leaving as one of the priorities of the 2021-2027 Cohesion Policy to foster inclusive growth. In the same vein, the prototype dashboard for social and economic resilience (European Commission, 2020b) considers early school leavers as a factor of socio-economic vulnerability in the category of social distress.

Inhaling PM2.5 has negative effects for health, among them respiratory and cardiovascular morbidity and lung cancer (World Health Organization, 2013). Moreover, this higher incidence of illnesses also puts greater pressure on public finances through health programmes and social benefits (sick leave, for example). PM2.5 air pollution is considered an indicator of vulnerability in the prototype dashboards for the geopolitical, green and digital dimensions of resilience because it constitutes an environmental threat (European Commission, 2020b).

Overall, older people who have left the labour market have lower average incomes and are more exposed to poverty than the rest of population (Marical et al., 2008; Peichl et al., 2012). Due to the uncertainty caused by the pandemic, even lower birth rates are expected in the EU and greater population ageing. Because people older than 65 constitute a healthcare burden and are at greater risk of poverty, they are considered a factor of vulnerability in the prototype dashboard for social and economic resilience (European Commission, 2020b). The European Pillar of Social Rights Action Plan also considers that special attention needs to be devoted to older people to promote health and care and ensure they live in dignity (European Commission, 2021c).

Higher levels of unemployment offset economic development processes since these are linked with lower standards of living and social problems (for example, robberies, crimes, etc.), so that unemployment reduces life satisfaction of the wider population (Chadi, 2014; Helliwell & Huang, 2014). In addition, during the COVID-19 pandemic, people not working for involuntary reasons were at greater risk of suffering mental disorders (Yao & Wu, 2021) and more likely to self-report more physical and mental health problems (Ikeda et al., 2021). It is also convenient to include women?s unemployment and youth unemployment because in the EU27 they reached values above men?s unemployment in 2019 (6.9%, 15.3% and 6.3% respectively, Eurostat information), and because they add specific aspects of fragility to the regions. Female unemployment is one of the social conditions most strongly correlated with income inequality (Kollmeyer, 2013; Sánchez & López-Corral, 2018) and is an explanatory factor for the higher incidence of risk of poverty in older women than in older men (Dessimirova & Bustamante, 2019). For its part, youth unemployment contributes to deteriorating their resilience, optimism, autonomy and overall life satisfaction (Merino et al., 2019). Unemployment rate is considered an indicator of economic vulnerability in the prototype dashboard for social and economic resilience (European Commission, 2020b). The European Pillar of Social Rights Plan distinguishes unemployment rates by groups of people and indicates, among its objectives, the reduction of the gap in male and female employment rates (European Commission, 2021c). It also defends that special efforts need to be devoted to young people who are more vulnerable to labour market fluctuations.

Migrants are likely to be one of the most vulnerable population groups, whether displacement is due to economic reasons or forced by violence. Migration has negative effects on quality of life because people's family and social ties break down and they are more exposed to poverty (Sánchez Mójica, 2013). Within the context of the COVID-19, migrant workers in the EU are very vulnerable because they are more likely to be in temporary employment, earn lower wages and have jobs that are less amenable to teleworking (Fasani & Mazza, 2020). In the same vein, the European Pillar of Social Rights Action Plan states that the 2021-2027 Cohesion Policy should pay special attention to migrants since they are more vulnerable to fluctuations in the labour market (European Commission, 2021c). Additionally, as we indicated in a previous section, for the first time, migration will receive 3% of the Structural Funds in the 2021-2027 Cohesion Policy (European Court of Auditors, 2019). Under this approach, positive migration ratios are considered a factor of socio-economic fragility and a greater pressure on public finances, and in those regions where the migration ratio is negative its values have been replaced by zero.

A prevalence of assaults and criminal activities creates unstable environments and deters investment in productive activities, is negatively related to quality of life and slows down sustainable urban development (Chica-Olmo et al., 2020). As a consequence of COVID-19 economic hardships have worsened, so this situation may also lead to higher exposure to organised crime and a rise in corruption (European Commission, 2020b, pp. 10-11). Crime and assault rates are a factor of fragility that increase socio-economic vulnerability directly in the cities or towns where they are registered.

We now examine why the rest of the indicators are considered indicators of resilience. Gramillano et al. (2018) analysed the indicators most frequently used by the Directorate-General for Regional and Urban Policy of the EU to assess the effectiveness in achieving the innovation and digitalization priorities of the previous EU Cohesion Policy period (2014-2020). They concluded that private investment in research and innovation, as well as enterprises receiving support from research institutions, can measure the networking activity and be proxies for potential technological transfer and knowledge exchange. As widely used indicators of innovation, the authors highlight the number of enterprises that introduce new services, products or processes. The five indicators of innovation and digitisation of our system are considered as proxies of a region's intellectual assets in the Regional Innovation Scoreboard for 2021 developed by the European Commission to assess innovation performance, namely the relative strengths and weaknesses of European regions (European Commission, 2021b). The idea is that innovation and a highly educated and well-trained workforce are critical to the development of a competitive, smart and knowledge economy. Education and innovation capacity, including product creativity and design ? as a link between innovation and the market ? are key factors in determining the recovery of regions before a shock (economic crises, for example). In this vein, the European 2020 Strategy set targets for Members States in terms of R&D investment (3% of GDP) and tertiary educational attainment (minimum 40% of the population aged 30-34) (European Commission, 2010). Expenditure on R&D, both private and public, is considered an indicator of economic growth and innovation that fosters socio-economic resilience in the prototype dashboard for social and economic resilience (European Commission, 2020b). Likewise, registered community designs per billion GDP is one of the outcome indicators of goal 1 of the 2021-2027 Cohesion Policy (European Commission, 2018).

The use of the Internet is an increasingly crucial factor for competitiveness and economic security, as it determines the capacity of territories to compete in and benefit from the knowledge-based economy. Studies with a territorial approach have shown that the availability of high-speed networks is a key determinant of quality of life because it facilitates economic, educational and social connections (Sánchez et al., 2018). On the contrary, the lack of Internet could represent a digital divide that increases levels of economic and social inequality. The COVID-19 crisis underscored the importance of households having internet access. During the lockdown, people relied more on online communication via the Internet for attending schools, buying daily necessities and working from home (Shek, 2021). Thus, the prototype dashboards for the geopolitical, green and digital dimensions include digital skills, teleworking capacity and e-health among the capacity indicators of digital resilience (European Commission, 2020b). Additionally, the percentage of individuals who use the Internet for interactions with public authorities or the e-Administration is considered an indicator of digital capacity that fosters regional resilience in the prototype dashboards for these same dimensions (European Commission, 2020b).

Per capita GDP is the main indicator considered by the European Commission (2018) for the allocation of the Structural Funds because it is the most neutral measure and reliable indicator and reflects the needs and disparities of the regions and Member States (European Court of Auditors, 2019). Under the scope of resilience, the ability to save is key for helping families and companies cope with adverse situations (Alessi et al., 2020; Benczur et al., 2020; European Commission, 2020b; Le Blanc, 2020). Taking into account the negative social and economic effects of income inequality (for a review, see Sánchez & Pérez-Corral, 2018; Sánchez Per capita GDP is the main indicator considered by the European Commission (2018) for the allocation of the Structural Funds because it is the most neutral measure and reliable indicator and reflects the needs and disparities of the regions and Member States (European Court of Auditors, 2019). Under the scope of resilience, the ability to save is key for helping families and companies cope with adverse situations (Alessi et al., 2020; Benczur et al., 2020; European Commission, 2020b; Le Blanc, 2020). Taking into account the negative social and economic effects of income inequality (for a review, see Sánchez

& Pérez-Corral, 2018; Sánchez & Ruiz-Martos, 2018), we consider the regional indicator proposed by Sen (1976), that is, GDP adjusted by the Gini index of each country (the Gini for NUTS-2 is not available). & Ruiz-Martos, 2018), we consider the regional indicator proposed by Sen (1976), that is, GDP adjusted by the Gini index of each country (the Gini for NUTS-2 is not available).

4.3.2 Variables for the multilevel modelling

The explanatory variables of the multilevel models come from level 1 or region and level 2 or country. More specifically, we consider monetary poverty at regional level, and government expenditure in education and political stability at country level (see Appendix C). According to the conceptual framework of this study, the choice of these three variables has been guided by the assumption that socio-economic vulnerability can be induced and/or explained by the sensitivity or fragility to harm and adaptive capacity or resilience. Several works have studied resilience and the impact of COVID-19 in the EU and conclude that one of the main ways to deal with a shock such as falling income is to use one's own savings (Alessi et al., 2020; Giovannini et al., 2020; Le Blanc, 2020; Manca et al. 2017). That is, family savings can act as financial buffers for households in the wake of the COVID-19 crisis. In addition, these works highlight that being at the bottom of the income distribution and/or living in a poor neighbourhood increases the chances of not knowing how to cope with a situation of distress. In this vein, the percentage of people in a region with an income below 60% of the region's median income (variable Poverty in our models) could be a proxy for the degree to which a region would be adversely affected by the pressure of the pandemic, as well as the capacity to deal with the shock.

The two country-level variables we have chosen (Education and Stability) aim to account for the role of the public sector in regional vulnerability (SEVI). The literature referred to in the previous section highlights the importance of human and social capital as drivers of resilient behaviour and an adaptive capacity to deal with shocks. Thus, government expenditure in education as a merit good that fosters citizen participation (more democratic societies), equal opportunities and lower income inequality (Sánchez & López-Corral, 2018) could favour a society's adaptation capacity, as well as promote the opportunity to bounce forward. Lastly, in situations of market economy stress, political stability and good governance ensuring compliance with contracts are essential to guarantee the functioning of the markets (Chang, 2011). Likewise, in a crisis context such as the COVID-19 pandemic, people can be more resilient when they trust in the institutions and live in a society that provides a safe and prosperous environment (Bittmann, 2021; Giovannini et al., 2020). Table 4.2 shows the descriptive statistics of the variables.

Table 4.2: Descriptive statistics of socio-economic vulnerability index (SEVI) and variables of multilevel modelling EU27, 2018

	Mean	SD	Min	Max	CV	Sample size
SEVI	1.25	0.23	0.66	1.74	18.43	233 NUTS-2
Poverty	16.44	5.83	4.10	41.40	35.44	233 NUTS-2
Education	4.84	0.96	3.00	6.80	19.94	27 Member States
Stability	0.69	0.36	0.06	1.37	51.32	27 Member States

4.4 Empirical strategy

The choice of mathematical method for aggregating the single indicators into a composite indicator will depend on the kind of measurement model that best fits the phenomenon being analysed (Maggino, 2017). The conceptual framework to analyse the socio-economic vulnerability of EU regions, provided in a previous section, led us to develop our model under the scope of a formative model. In formative measurement models, causality flows from the single indicators to the latent variable, since single indicators are viewed as causes of the latent variable (see Diamantopoulos et al., 2008; Jiménez-Fernández & Ruiz-Martos, 2020). For instance, in our case, the socio-economic vulnerability index (SEVI) of a region includes indicators of innovation, education, unemployment, pollution, etc. Any change in one or more of these components (even if the other factors do not change) is likely to cause a change in a region?s SEVI score (the latent construct). However, if a region' SEVI decreases, it would not necessarily be accompanied by an improvement in all of the components (single indicators).

eeping this in mind, we applied an iterative distance methodology based on the Distance P2 introduced by Pena Trapero (1977) and applied in several works (see Cuenca et al., 2018; Sánchez et al., 2018; Sánchez & Ruiz-Martos, 2018; Zarzosa Espina & Somarriba Arechavala, 2013). We use the metric structure in the $\mathbf{R}^m$ vector space, where m is the number of single indicators. This allows us to obtain a composite indicator that measures distances to perform benchmarking between the units studied in order to develop the socio-economic vulnerability indicator (SEVI) of the 233 European regions or NUTS-2. In our case, the composite indicator represents a weighted Euclidean metric that is defined as follows (see Jiménez-Fernández et al., 2022):

$$SEVI_i = \left(\sum_{i=1}^{m} |x_{ij} - x_{*j}|^2 \omega_j \right)^{1/2} \tag{4.1}$$

where m is the number of single indicators, x_{ij} is the value of the $j-th$ indicator in the $i-th$ region, x_{*j} is the $-th$ value in the reference vector $X_* = (x_{*1}, \ldots, x_{*m})$ and ω_j is the weight of the $j-th$ single indicator.

Given that the single indicators often have different measurement units, the single indicators $\{X_1, \ldots, X_j\}$ have been normalised using Min-Max normalisation in order to make them comparable. That is, unlike the method proposed by Pena Trapero, we normalise the indicators before introducing them in the formula.

Our method considers in its calculation formula the distance between each individual indicator and the most desirable situation taken from the reference vector. The reference vector (X^*) is like a hypothetical region that, in the set of all EU regions, registers the best values of all single indicators. Thus, we take into account the complete empirical distribution in the 233 EU regions. More specifically, for single indicators with positive polarity, we select the minimum value of the indicator in the entire sample. For instance, early leavers is a single indicator with positive polarity: the higher the early leavers rate is in a region, the greater the region's vulnerability. The hypothetical best region (the least vulnerable) will register the lowest rate of early leavers, that is, the minimum value of all the regions. For single indicators with negative polarity, the reference value is the maximum value of the sample. For example, for R&D investment, the higher the value is in a region, the less vulnerable it is. In this case, the hypothetic best region or the least vulnerable will register the maximum value in R&D investment. Proceeding in this way, the SEVI composite indicator will take higher values, the greater the distance it is with respect to the most desirable values of the individual indicators. That is, the greater the SEVI, the more vulnerable or the worse the performance of a region in the different indicators studied. Consequently, we can quantify and compare all the regions under analysis.

The weights of the single indicators (ω_j) are computed using unsupervised machine learning algorithms. More specifically, we use multivariate adaptative regression splines (MARS) to identify the best functional relationships between the composite indicator and the set of single indicators. In this way, ω_i denotes the importance of each indicator according to its contribution to the SEVI and avoids potential multicollinearity issues. For a more detailed approach to this methodology and its properties, see Jiménez-Fernández et al. (2022).

4.4.1 The impacts of shocks on socio-economic vulnerability of EU regions

The second aim of this paper is to check whether country characteristics interact with regions? characteristics to alter patterns of socio-economic vulnerability. In other words, we consider the possibility that two regions randomly selected from the same country will register a more similar level of socio-economic vulnerability than two regions randomly selected from different countries. This would mean that we assume no independence among regions belonging to the same country. To test this hypothesis, multilevel models should be used. In a classical one-level model it is assumed that the observations are independent, and the error is treated as noise, so the estimate should minimise the error. However, when the data is nested, the correlation between observations within a group could be different from the correlation between groups, resulting in two types of errors. An advantage of multilevel models is that they analyse what part of the random error is due to the effect of level 2 (country) and what part is due to level 1 (regions) (see Goldstein, 2011; Snijders & Bosker, 2012). That is, multilevel modelling allows us to determine what part of the variability in the regions' socio-economic vulnerability can be explained by country characteristics.

Likewise, multilevel modelling distinguishes between the fixed or deterministic part of the model and the random or stochastic part, thus enabling a two-directional analysis. Firstly, by estimating the signs and values of the model parameters (fixed part of the model), we can study how changes at the regional (level 1) and country level (level 2) influence $SEVI$, as well as identify the regional and country-level characteristics associated with resilient behaviour. Secondly, the random part of the models could inform us on how a shock, such as COVID-19, would impact on regional vulnerability (third objective of this paper). In turn, in the random part of the model, we can analyse the possible idiosyncratic and covariate shocks caused by COVID-19. That is to say, we can identify what proportion of the variability in vulnerability ($SEVI$) not explained by the model (stochastic or random effects) is attributable to regional-level characteristics (idiosyncratic effects) or to the interaction between the country characteristics and the regional characteristics (covariate effects).

Next, we present two different specifications to estimate multilevel models which will allow us to check the aims or hypotheses 2 and 3 of this study.

4.4.2 Specification 1: multilevel random intercept model

We consider a two-level structure where regions i (level 1) are nested or hierarchised into countries j (level 2). The random intercept model accounts for country differences in $SEVI$. In this specification, the intercept varies randomly between the countries, but the slope is the same for all of them. Let $SEVI_{ij}$ be the value of the socio-economic vulnerability index in region i and country j, where $i \in \{1,\ldots,233\}$ and $j \in \{1,...,27\}$. For each observation located in the j-country, the model can be written as follows:

$$SEVI_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij} \tag{4.2}$$

where $\beta_{0j} = \beta_0 + u_j$, $x_{ij} \in X$ a $n \times m$-dimensions matrix of observed explanatory variables both at regional and country level, and β_1 its associated parameters. For country j, the intercept is β_{0j} which may be smaller or larger than the intercept of population β_0. The country random effects are denoted by u_j and the regional residuals (with nxm dimensions) are denoted by e_{ij}. The residuals u_j are assumed to have a normal distribution of zero mean and variance . In order to identify the fixed and random parts of the model, Equation (2) can be written as:

$$SEVI_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij} \tag{4.3}$$

In this equation, the fixed part of the model shows the relationship between the mean of $SEVI$ and the explanatory variables ($\beta_0 + \beta_1 x_{ij}$ with parameters β_0, β_1), and the random part captures the residuals from different levels ($u_j + e_{ij}$ with variances σ_u^2, σ_e^2).

Following this specification, we estimated the null model (without explanatory variables) and Model 1, which includes the variables Poverty, Education and Stability. The null model allows us to check if the structure of socio-economic vulnerability in the EU regions is nested, that is, whether there is an interaction between the regional-level and country-level variables (objective 2). If a nested structure is confirmed, multilevel modelling would be a suitable approach because one-level modelling could lead to incorrect conclusions (see Goldstein, 2011; Snijders & Bosker, 2012).

4.4.3 Specification 2: random slope model for Poverty variable

Specification 2 is an extension of the random intercept model which also considers that the slope for the variable Poverty varies randomly among the different countries. Let $SEVI_{ij}$ be the variable that indicates the value of the socio-economic vulnerability index in region i of country j, where $i \in \{1,\ldots,233\}$ and $j \in \{1,\ldots,27\}$. For each observation located in j-country, the model can be written as follows:

$$SEVI_{ij} = \beta_{0j} + \beta_1 p_{ij} + \beta_2 x_{ij} + e_{ij} \tag{4.4}$$

where $\beta_{0j} = \beta_0 + u_{0j}$ and $\beta_{1j} = \beta_1 + u_{1j}$; $x_{ij} \in X$, being X a $n \times m$-dimensions matrix of observed explanatory variables at both regional and country levels, and β_2 its associated parameters. The variable Poverty is denoted by p. The average regression for Poverty has slope ?1 and the slope for each country is β_{1j}. The random errors u_{0j} and u_{1j} are assumed to have a normal distribution of zero mean and variance σ^2_{u0} and σ^2_{u1} , respectively. Model 2 is estimated following this specification. Developing equation 4, we can identify the fixed and random parts of the model:

$$SEVI_{ij} = \beta_0 + \beta_1 p_{ij} + \beta_2 x_{ij} + u_{0j} + p_{ij} u_{1j} + e_{ij} \tag{4.5}$$

In the Equation 4.5, the fixed part of the model shows the relationship between the mean of SEVI and the explanatory variables ($\beta_0 + \beta_1 p_{ij} + \beta_2 x_{ij}$ with parameters $\beta_0, \beta_1, \beta_2$), and the random part captures the residuals from different levels ($u_{0j} + c_{ij} u_{1j} + e_{ij}$ with the parameters σ^2_{u0}, σ^2_{u1}, σ^2_e ; where $p_{ij} u_{1j}$ is the interaction between the country and Poverty).

Following this specifications, we estimate Model 4.2. Following this specifications, we estimate Model 2, where all the explanatory variables are included (Poverty, Education and Stability)., where all the explanatory variables are included (Poverty, Education and Stability).

4.4.4 Idiosyncratic effects and covariate effects

Changes in socio-economic vulnerability caused by a shock such as COVID-19 can be introduced and analysed throughout the random or stochastic part of the multilevel models (see, for instance, Halkos et al., 2020). In turn, within the random part, we can distinguish what proportion of the variability in vulnerability (SEVI) not explained by the model is attributable to regional-level (idiosyncratic effects) or country-level effects and the interrelation between country and regional levels (covariate effects). In this vein, the interclass correlation (ICC) informs what part of the random effects would be explained by covariate effects. That is, the ICC informs us how changes in the environment or covariates of the regions could affect their socio-economic vulnerability.

In specification 1 with random intercept, the ICC can be calculated as follows:

$$ICC = \frac{\sigma^2_{u0}}{\sigma^2_{u0} + \sigma^2_e} \tag{4.6}$$

where σ^2_e is the residual variance and σ^2_{u0} the variance between groups (countries).

In specification 2 with random slope, the variance between groups depends on the value of the variable Poverty (p) in each region; hence the ICC takes different values in each region. This information is interesting since it allows us to obtain a map of EU regions that shows the intensity of the effects of COVID-19 on their socio-economic vulnerability. The formula for calculating the ICC can be expressed as follows:

$$ICC_i = \frac{Var(u_{0j} + p_{ij} u_{1j})}{Var(u_{0j} + p_{ij} u_{1j}) + \sigma^2_e} \tag{4.7}$$

where $p_{ij} u_{1j}$ is the interaction between country j and the variable Poverty (p) at regional level, and where

$$Var(u_{0j}+p_{ij}u_{1j}) = Var(u_{0j})+p^2_{ij}Var(u_{1j})+2p_{ij}Cov(u_{0j},u_{1j}) = \sigma^2_{u0}+p^2_{ij}\sigma^2_{u1}+2p_{ij}\sigma_{u_{0j},u_{1j}} \tag{4.8}$$

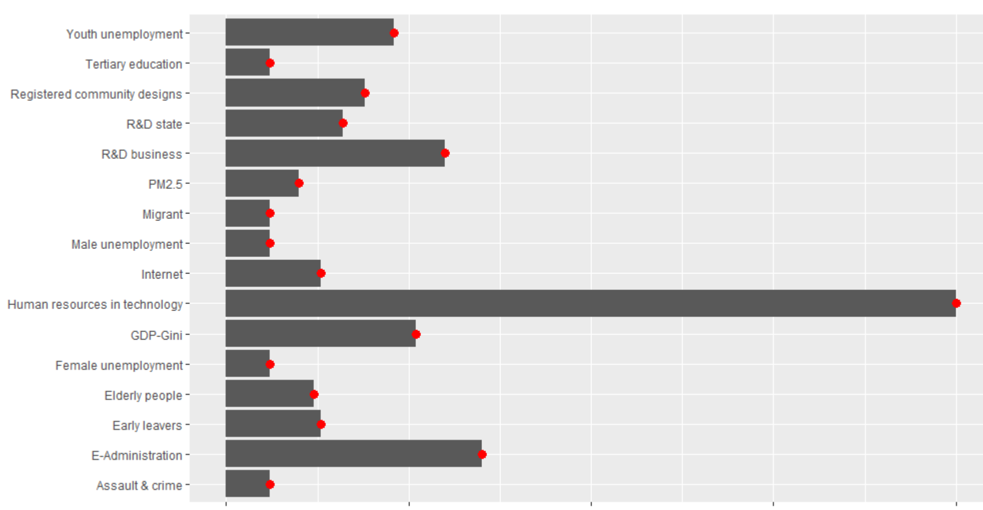
4.5 Results

4.5.1 The socio-economic vulnerability of EU regions

Focusing on the descriptive statistics for the 16 indicators of socio-economic vulnerability we have analysed (Table 4.1), the values of Pearson's coefficient of variation indicate that the largest territorial differences arose in the objective of fostering an innovative and smart economic transformation (especially in the indicators registered community designs, R&D business and R&D state), as well as in social rights (especially in migrant and female unemployment). The last column of Table 1 shows the regions that rank highest in each single indicator, namely our reference vector to build the SEVI. In other words, from a socio-economic viewpoint, the best theoretical region in the EU (the least fragile and most resilient) should register the values of the last column. The further a region is from this hypothetical region, the greater its socio-economic vulnerability and therefore the greater the value of its SEVI. Overall, we observe that two regions of eastern European countries (Croatia and Poland) register the best positions in the three indicators of human capital (early leavers, tertiary education and human resources in technology). Four regions of Germany invest the most in innovation (business R&D and state R&D) and have the lowest unemployment rates for both women and youth.

The average value of the SEVI is 1.25. Hovedstaden in Denmark is the least socio-economically vulnerable region in the EU27 as it has the lowest SEVI (0.66), while Dytiki Makedonia in Greece is the most vulnerable (maximum SEVI value = 1.74) (Table 4.2). From a statistical viewpoint, SEVI is a variable that follows a normal distribution (Shapiro-Francia test, $z = 1.552$, $p = 0.06027$, $N = 233$; see Appendix D). Figure 4.3 shows the weights assigned to each indicator; specifically, the proportion in which each indicator contributed to the metric and therefore to the SEVI. Resilience indicators, especially innovation and digitisation (goal 1 of the EU 2021-2027 Cohesion Policy), have the highest weights. Among the fragility indicators of the regions, youth unemployment, early leavers, elderly people and PM2.5 register the largest weights.

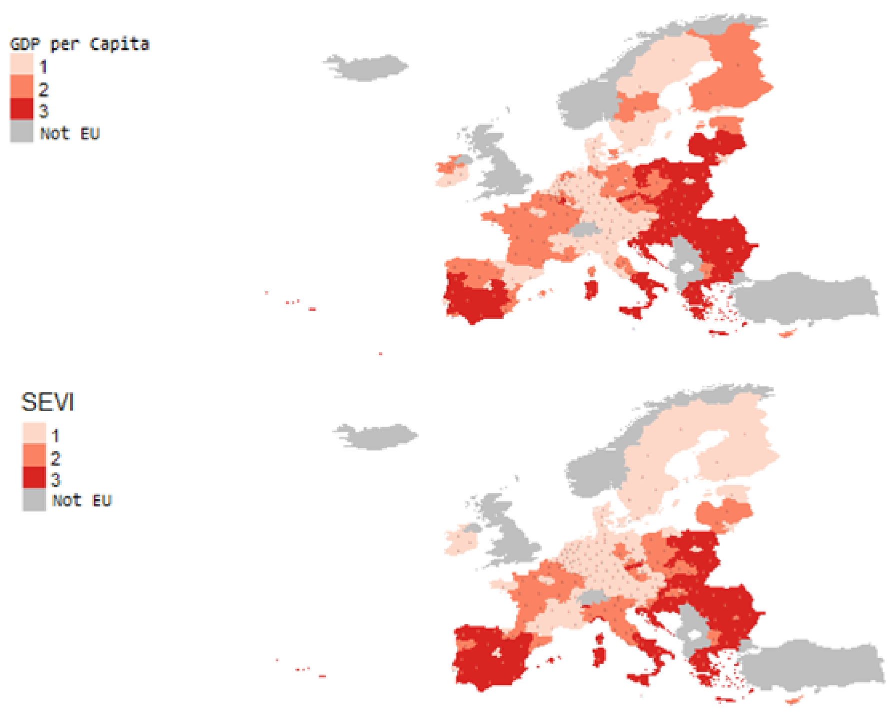
Figure 4.3: Weights of the single indicators of SEVI



Following the European Commission's proposal (2018) for the distribution of the Structural

Funds, that is, taking as a reference the GDP per capita (average 2016-2017) and the population (average 2016-2018), the 233 NUTS-2 could be grouped into three blocks: 47% of the population of the EU27 would reside in regions where the GDP per capita is above the GDP per capita for the whole of the EU27, 25% of the population in regions with a GDP per capita between 75% and 100% of the EU27, and the remaining 28% of the population in regions where GDP per capita is less than 75% of the EU27. In order to analyse the implications for the 2021-2027 Cohesion Policy while maintaining the same budgetary effort, we take these population percentages as a reference to divide the EU regions into three groups according to the SEVI. Figure 4.4 displays the results of the SEVI grouped into the three types of regions analysed.

Figure 4.4: Classification of EU regions according GDP per capita and socio-economic vulnerability



The regions that are in the most disadvantaged situation to face the challenges of the 2021-2027 Cohesion Policy, namely group 3 which represents 27.25% of the EU27 population are: all regions of Greece and Croatia; all regions of Romania, Bulgaria and Slovakia except the regions where their respective capitals are located; all regions of Hungary and Portugal except two; more than half of the territory of Spain and Poland; and the regions located in southern Italy. In contrast, the regions in the best position or group 1, which represent 46.35% of the EU27 population, are: Estonia and Malta; all regions of Denmark, Finland, Ireland, the Netherlands and Sweden; and all regions of Austria, Belgium and Germany except one. The rest of the regions (group 2) represent 26.40% of the EU27 population and are located mainly in France, northern Italy and the Czech Republic.

At a first glance, it might seem that the European Commission criterion for allocating the Structural Funds and our multidimensional proposal (SEVI) lead to similar results since the pairwise correlation between the GDP per capita and the SEVI for the whole set of 233 NUTS-2 is quite high ($r = -0.77$, $p < .001$). However, if we distinguish among the three groups of regions, the results are somewhat different. As Table 4.3 indicates, there is no correlation in the regions of group 2 between our proposal of socio-economic vulnerability and the one-dimensional criterion of the European Commission. Likewise, the correlation is low in group 3. Focusing on groups 2 and 3, we can identify which regions would be harmed in terms of the allocation of Structural Funds if the traditional criterion were applied. To do so, a single indicator can be taken as a reference of economic activity (GDP pc) instead of a set of indicators that complement the GDP and accurately reflect the socio-economic fragility and capabilities of the EU regions in order to face the challenges of the Cohesion Policy.

Table 4.3: Correlation between per capita GDP and SEVI by groups of EU regions

	Group 1	Group 2	Group 3
Correlation coefficient (r)	-0.6015	-0.1158	-0.3637
p-value	$< .001$	0.4280	0.0016
N	111	49	73
population	46.35	26.40	27.25

The most remarkable outcome in groups 2 and 3 is that 12 out of 21 Italian regions, 11 out of 17 Spanish regions, two regions in Portugal, Corse in France, Attiki in Greece and Bucuresti-Ilfov in Romania would be negatively affected following the European Commission criterion. In other words, despite the fact that these regions surpass the thresholds of per capita GDP, according to our multidimensional criterion, they are more vulnerable and less resilient and should therefore attract more financial attention under the EU Cohesion Policy for the period 2021-2027. On the other hand, 12 regions located in Member States of the previous eastern Europe turn out to be less vulnerable from a socio-economic standpoint than their relatively low position in per capita GDP reflects.

4.5.2 Multilevel analysis: socio-economic vulnerability and COVID-19

The results of the null model, the random intercept model (Model 1) and the random slope model (Model 2) are shown in Table 4. The results of the null model indicate differences in the socio-economic vulnerability of the regions across countries because the likelihood ratio (LR) test ($\chi^2(1) = 209.98$, $p < .001$), which contrasts the multilevel model against the one-level OLS model, is significant. In fact, the value of the intraclass correlation (ICC = 0.70) might be interpreted as meaning that 70% of the variability in socio-economic vulnerability is attributable to differences across countries. Thus, the estimation of multilevel models that take into account the "country" effect and the interaction between regional and country variables is justified.

Table 4.4: Multilevel modelling of the effects of regional and country characteristics on socio-economic vulnerability in the EU regions, 2017 ($N_{regions} = 233$; $N_{countries} = 27$)

	Null model	Model 1	Model 2
Fixed effects (p-value)			
Poverty (region level)		0.014(< 0.001)	0.014(< 0.001)
Education (country level)		-0.018(< 0.002)	-0.075(< 0.004)
Stability (country level)		-0.235(0.001)	-0.255(< .001)
Intercept	1.232(< .001)	1.543(< .001)	1.543(< 0.001)
Random effects (p-value)			
Variance intercept (σ^2_{u0})	0.03655	0.01337	0.00233
Variance poverty (σ^2_{u0})	—	—	0.00002
Covariance (u_{0j}, u_{1j})	—	—	0.00021
95% conf. interval covariance	—	—	(0.00004, 0.00038)
Variance residual (σ^2_e)	0.01576	0.0132	0.01008
Interclass correlation (ICC)	0.70	0.56	(0.31, 0.84) (a)
Model fit			
-2Log Lik	-234.79	-346.91	-349.14
LR test, χ^2(p-value)	209.98(< .001)	157.09(< .001)	159.31(< .001)
R^2m(fixed)	—	54.32%	69.71%
R^2c(fixed & random)	—	80.01%	75.45%

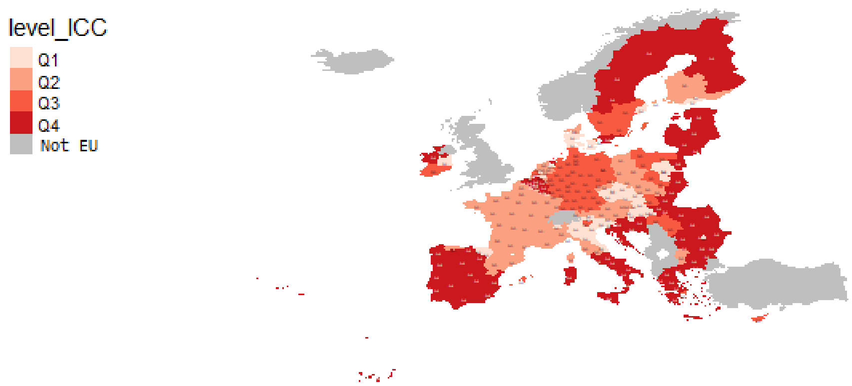
Note:(a) In the estimation with random intercept and random slope (Model 2), the variance (u_0, u_1) takes different scores for each value of the explanatory variable whose slope is considered to be random; thus the ICC yields different scores for each region.

Models 2 and 3 incorporate the three explanatory variables Poverty, Education and Stability. Several goodness measures of the model are reported at the bottom of Table 4. In the framework of multilevel models, the marginal R-squared (R^2m) represents the variance explained by fixed factors of the model and the conditional R-squared (R^2c) represents the variance explained by fixed and random factors (see Nakagawa & Schielzeth, 2013). The difference between the corresponding R^2c and R^2m values reflects the amount of variability in the random effects. Both models 1 and 2 present high R^2, the indicator -2 log likelihood decreases from model 1 to 2 and the result of the likelihood ratio (LR) test shows that Model 2 is an improvement over Model 1.

In both models (1 and 2), all the variables are statistically significant, and the signs of their estimated parameters are consistent with the literature. Namely, increases in regional monetary poverty would be associated with a rise in the socio-economic vulnerability of the regions, whereas increases in government expenditures in education and an improvement in self-reported political stability would lead to a reduction in vulnerability or foster the capacity for resilience. In addition, the results of Model 2, which analyses the relationship between vulnerability and monetary poverty for each country, indicate that increases in monetary poverty lead to greater socio-economic vulnerability in regions with a higher level of vulnerability (the covariance is positive and statistically significant).

Focusing on the random or stochastic part of Model 2 and applying Formulas 4.7 and 4.8, we calculated the ICC for each region. ICC provides the proportion of vulnerability variability not explained by the model that is attributable to changes in the environment or covariates of the regions, such as changes caused by the COVID-19 pandemic that could interact with each country?s characteristics and with each region?s poverty level. The ICC varies from 30.5% in Bucuresti-Ilfov (Romania) to 84.3% in Campania (Italy). Figure 4.5 illustrates the different degrees of exposure to the effects of COVID-19 on regional vulnerability, depending on the level of poverty. The ICC results are grouped into quartiles according to the number of regions. The regions in which socio-economic vulnerability would be most exposed to the effects of COVID-19 (fourth quartile, between 61.1% and 84.3%) would be all of Portugal, Greece, Croatia, Estonia, Latvia, and Lithuania; a large portion of the territory of Spain, Romania and Bulgaria; southern Italy and the eastern regions of Poland. It is worth noting that much of the territory of Sweden and Finland and a region of Ireland that occupied a better position in the SEVI would be among the regions most exposed to the covariate effects. On the other hand, most of the regions of Denmark, Slovakia, the Czech Republic, the Netherlands, Hungary, northern Italy, southern Finland and one of the three regions of Ireland register values in the first quartile (between 30.5% and 53.4%).

Figure 4.5: Covariate effects of COVID-19 on socio-economic vulnerability of EU regions



Note. ICC is intra-class correlation. Q1 (30.5%, 53.4%), Q2 (53.5%, 58.4%), Q3 (58.5%,61%), Q4 (61.1%,84.3%).

4.6 Conclusions and discussion

The EU Cohesion Policy for the period 2021-2027 focuses on five goals for the EU to become smarter, greener, more connected, more social and closer to citizens. However, a macroeconomic index (per capita GD) is proposed as the predominant criterion for classifying the regions and allocating the Structural Funds. We hypothesise that it is possible to consider new complementary criteria that better reflect citizens' quality of life. This approach is especially important because the COVID-19 has exposed the vulnerabilities within the EU in all the domains: jobs, education, economy, welfare systems and social life (European Commission, 2021c). On this basis, we have built a composite socio-economic vulnerability index (SEVI) for each of the 233 NUTS-2 of the EU in 2017 that synthesises the information on fragility and resilience factors in order to achieve the objectives of the 2021-2027 Cohesion Policy. The idea is that the higher the value of SEVI, the greater the difficulty in achieving these objectives compared to the rest of the regions.

By implementing the SEVI as an allocation mechanism of the Structural Funds rather than GDP per capita as proposed by the EU, and with an equivalent budgetary effort in terms of the benefited population, we obtain remarkable differences. Our main findings are that a large number of regions in Italy and Spain and some in Portugal, France and Greece which exceed the limit in terms of GDP should be in the group of the most benefited regions according to their socio-economic vulnerability. On the contrary, regions in Member States of the previous eastern Europe, which are historically characterised by low levels of GDP, reach relative positions of less socio-economic vulnerability in our multidimensional approach. These differences in the maps of priority regions could be a source of debate surrounding the introduction of new game rules for the EU Cohesion Policy, especially in the current context of economic and social changes where public policies should prioritise improving citizens' quality of life.

In a second stage, we study the effects of COVID-19 on regional vulnerability since it is foreseeable that the pandemic will trigger inequalities and increase poverty levels (Fetting, 2020; Giovannini et al., 2020; Shek, 2021; United Nations, 2020a, 2020b). The question is whether all regions will be equally exposed to COVID-19 in terms of their socio-economic vulnerability. To answer this question, we analyse both the idiosyncratic and covariate shocks that COVID-19 might represent by estimating multilevel models. Our findings indicate that increases in government expenditures in education and improving political stability would reduce the regional vulnerability or foster the capacity for resilience. On the other hand, increases in regional monetary poverty would be associated with increased vulnerability, causing bigger growth in the regions with a higher level of vulnerability. Even though regions with a larger SEVI would be the most exposed to the effects of COVID-19, it is remarkable that much of the territory of Sweden and Finland and the region of Ireland that ranked highest in the SEVI would be among the most exposed to the covariate effects. These results might have public policy implications; for example, to inform on how to distribute the European COVID-19 Recovery Funds.

The multidimensional character of our proposal, the study of regions' factors of vulnerability, fragility and resilience, fits into the mainstream view of economists and policymakers who argue that associating the notion of economic and social progress to a one-dimensional variable of economic activity, such as GDP or income, is debatable (Fetting, 2020; O'Donnell et al., 2014; Sánchez et al., 2018; Sánchez & Ruiz-Martos, 2018; Stiglitz et al., 2018). Our proposal is also in line with two plans or strategies that the European Commission has recently approved to continue advancing in the double green and digital transition and to recover from the COVID-19 crisis: the European Pillar of Social Rights Action Plan (European Commission, 2021c) and the 2020 Strategic Foresight Report (European Commission, 2020b). The objective of both plans is to promote resilience through the EU institutions so that Europe will recover faster and emerge stronger from the COVID-19 crisis and future crises . A key aspect is that the priorities identified in both plans must be taken into account in all EU policymaking, including the 2021-2027 Cohesion Policy.

In this paper, we have argued that, in terms of Cohesion Policy, there is still room to go "beyond GDP" and consider, in financial terms, the vulnerability and resilience factors that determine people's well-being. The two previous initiatives or strategies lead us to be hopeful and to think that the EU will continue to advance on the "beyond GDP" path by strengthening the principles of a social market economy. Likewise, the type of exercise carried out in the second stage of this paper can be useful to stimulate discussions regarding the guidelines on how to increase the resilience and reduce the fragility of the regions in order to cope with unforeseen shocks.

Lastly, we would like to point out that our approach to study socio-economic vulnerability differs from most of the studies carried out in this field in the following regards. Firstly, a large number of studies focus on defining vulnerability as the likelihood that, at a given time in the future, an individual will have a level of welfare (income, consumption, poverty, etc.) below some threshold established in a "normative" way. In contrast, under our methodological approach, the choice of normative or arbitrary thresholds is not required, thus overcoming one of the main criticisms of methods involving the elaboration of composite indicators and vulnerability analysis (see, for instance, Dutta et al., 2011; Gallardo, 2018; Nájera & Gordon, 2019; Povel, 2015). Secondly, we do not study risk by estimating the probability of occurrence of future events (for instance via probit/logit models) because our dependent variable in the multilevel models (SEVI) is expressed in a metric. Therefore, to express it as a categorical variable, it would be necessary to collapse the values into two categories, which means that a "normative" threshold would have to be set to establish the limit of the two values, as well as assuming an unnecessary loss of information.

4.7 Appendix A: Definitions and sources of the single indicators to build SEVI

Table 4.5

Definitions and sources of the single indicators to build SEVI	
Early leavers from education and training denotes the percentage of the population aged 18 to 24 having attained at most lower secondary education and not being involved in further education or training.	1.1 **Source:** Eurostat, Educational attainment level and transition from education to work (based on EU-LFS) (edat_lfse_04) 1.2 **Geographic level:**NUTS-1 for some regions in AT, DE, FI, FR, IT, PL, UK. NUTS-2 for all other countries. 1.3 **Date of data used:** Average 2016-2017.
PM2.5. Mean population exposure to fine particles PM2.5. Micrograms per cubic metre.	2.1 **Source:** OECD, Environment Database. Exposure to PM2.5. 2.2 **Geographic level:** NUTS-2, own elaboration. 2.3 **Date of data used:** Average 2016-2017.
Elderly people. Percentage of elderly people in population (75 years or over).	3.1 **Source:** Eurostat, Population change ? Demographic balance and crude rates at regional level (demo_r_pjangroup). 3.2 **Geographic level:** NUTS-2. 3.3 **Date of data used:** Average 2016-2017.
Male unemployment. Unemployment rate % from 20 to 64 years (male).	4.1 **Source:** Eurostat, Regional labour market statistics (lfst_r_lfu3pers). 4.2 **Geographic level:** NUTS-2. 4.3 **Date of data used:** Average 2016-2017.
Female unemployment. Unemployment rate % from 20 to 64 years (female).	5.1 **Source:** Eurostat, Regional labour market statistics (lfst_r_lfu3pers). 5.2 **Geographic level:** NUTS-2. 5.3 **Date of data used:** Average 2016-2017.
Youth unemployment rate % from 15 to 24 years (female + male).	6.1 **Source:** Eurostat, Regional labour market statistics (lfst_r_lfu3pers). 6.2 **Geographic level:** NUTS-1 for some regions in AT, DE, FI, HU, LT, PL, PT, UK. NUTS-2 for all other countries. 6.3 **Date of data used:** Average 2016-2017.

Table 4.6

Migrant. Crude rate of net migration plus statistical adjustment. Difference between the crude rate of population change and the crude rate of natural change; that is, net migration is considered as the part of population change not attributable to births and deaths expressed per 1,000 inhabitants. Only positive rates are considered, otherwise zero is assigned.	7.1 **Source:** Eurostat, Population change, Demographic balance and crude rates at regional level (demo_r_gind3). 7.2 **Geographic level:** NUTS-2. 7.3 **Date of data used:** Average 2016-2017.
Assault & crime. Number of deaths by assault and homicide divided by population and then multiplied by 100,000 (crude death rate).	8.1 **Source:** Eurostat, Causes of death, crude death rate by NUTS-2 region of residence (hlth_cd_acdr2). 8.2 **Geographic level:** NUTS-2. 8.3 **Date of data used:** Average 2016-2017.
R&D business. Intramural R&D expenditure Business enterprise sector (percentage of gross domestic product).	9.1 **Source:** Eurostat, Statistics on research and development (rd_e_gerdreg). 9.2 **Geographic level:** NUTS-1 for some regions in BE, LT, NL, PL. NUTS-2 for all other countries. 9.3 **Date of data used:** Average 2015-2016, except: BE 2014-2015; AT, DE, EL, IE 2015; NL 2014; FR 2013.
R&D state. Intramural R&D expenditure Government sector + higher education sector (percentage of gross domestic product).	10.1 **Source:** Eurostat, Statistics on research and development (rd_e_gerdreg). 10.2 **Geographic level:** NUTS-1 for some regions in BE, LT, NL, PL. NUTS-2 for all other countries. 10.3 **Date of data used:** Average 2015-2016, except: BE 2014-2015; AT, DE, EL, IE 2015; NL 2014; FR 2013.
Tertiary education. Individuals aged 25-64 who successfully completed tertiary education (levels 5-8 ISCED 2011) over the population with the same age (In %).	11.1 **Source:** Eurostat, Educational attainment level and transition from education to work (based on EU-LFS) (edat_lfse_04) 11.2 **Geographic level:** NUTS-2. 11.3 **Date of data used:** Average 2016-2017.
Human resources in technology. Persons employed in science and technology as percentage of active population.	12.1 **Source:** Eurostat, Human Resources in Science & Technology (hrst_st_rcat). 12.2 **Geographic level:** NUTS-2. 12.3 **Date of data used:** Average 2017-2018.
Registered community designs. Number of registered community designs per billion GDP purchasing power standards.	13.1 **Source:**Eurostat, Community design (ipr_dr_reg). 13.2 **Geographic level:** NUTS-2. 13.3 **Date of data used:** Average 2015-2016.

Table 4.7

Internet. Percentage of households with internet access at home.	14.1	**Source:**Eurostat, ICT usage in households and by individuals (isoc_r_iacc_h).
	14.2	**Geographic level:** NUTS-1 for some regions in DE, EL, PL, UK. NUTS-2 for all other countries.
	14.3	**Date of data used:** Average 2017-2018.
E-Administration. Percentage of individuals who used the Internet for interaction with public authorities (last 12 months).	15.1	**Source:**Eurostat, ICT usage in households and by individuals (isoc_r_gov_i).
	15.2	**Geographic level:** NUTS-1 for some regions in DE, EL, PL, UK. NUTS-2 for all other countries.
	15.3	**Date of data used:** Average 2017-2018.
GDP-Gini. Regional gross domestic product (GDP) purchasing power standard per inhabitant adjusted by the country Gini index of disposable household income [GDP per $capita^{*}$(1-Gini index)]. Disposable household income includes: all income from work (employee wages and self employment earnings), private income from investment and property, transfers between households, and all social transfers received in cash, including old-age pensions.	16.1	**Source:**Eurostat, Regional economic accounts (nam_10r_2gdp) and Income and living conditions (ilc_di12).
	16.2	**Geographic level:** NUTS-2.
	16.3	**Date of data used:** Average 2016-2017.

4.8 Appendix B: Selection of single indicators to build the socio-economic vulnerability index of European regions

Table 4.8

Fragility indicators	European Union plans, strategies and projects
Early leavers	European 2020 Strategy (European Commission, 2010). Prototype dashboard for social and economic resilience (European Commission, 2020b). The European Pillar of Social Rights Action Plan (European Commission, 2021c).
PM2.5	Prototype dashboard for the green resilience (European Commission, 2020b). Regional Innovation Scoreboard 2021 (European Commission, 2021a)
Elderly people	Prototype dashboard for social and economic resilience (European Commission, 2020b). The European Pillar of Social Rights Action Plan (European Commission, 2021c).
Male unemployment	Prototype dashboard for social and economic resilience (European Commission, 2020b). The European Pillar of Social Rights Action Plan (European Commission, 2021c).
Female unemployment. Unemployment rate % from 20 to 64 years (female).	Prototype dashboard for social and economic resilience (European Commission, 2020b). The European Pillar of Social Rights Action Plan (European Commission, 2021c)
Youth unemployment	Prototype dashboard for social and economic resilience (European Commission, 2020b). The European Pillar of Social Rights Action Plan (European Commission, 2021c).
Migrant. Crude rate of net migration plus statistical adjustment. Difference between the crude rate of population change and the crude rate of natural change; that is, net migration is considered as the part of population change not attributable to births and deaths expressed per 1,000 inhabitants. Only positive rates are considered, otherwise zero is assigned.	The European Pillar of Social Rights Action Plan (European Commission, 2021c). Allocation of Cohesion Policy Funding (European Court of Auditors, 2019).
Assault & crime.	Prototype dashboard for social and economic resilience (European Commission, 2020b).

Table 4.9

Resilience indicators	European Union plans, strategies and projects
R&D business	European 2020 Strategy (European Commission, 2010) Prototype dashboard for social and economic resilience (European Commission, 2020b). Regional Innovation Scoreboard 2021 (European Commission, 2021b).
R&D state	European 2020 Strategy (European Commission, 2010) Prototype dashboard for social and economic resilience (European Commission, 2020b). Regional Innovation Scoreboard 2021 (European Commission, 2021b).
Tertiary education	Prototype dashboard for social and economic resilience (European Commission, 2020b). Regional Innovation Scoreboard 2021 (European Commission, 2021b).
Human resources in technology	Regional Innovation Scoreboard 2021 (European Commission, 2021b).
Registered community designs	Regional Innovation Scoreboard 2021 (European Commission, 2021b). Cohesion Policy (European Commission, 2018).
Internet	Prototype dashboard for the digital resilience (European Commission, 2020b).

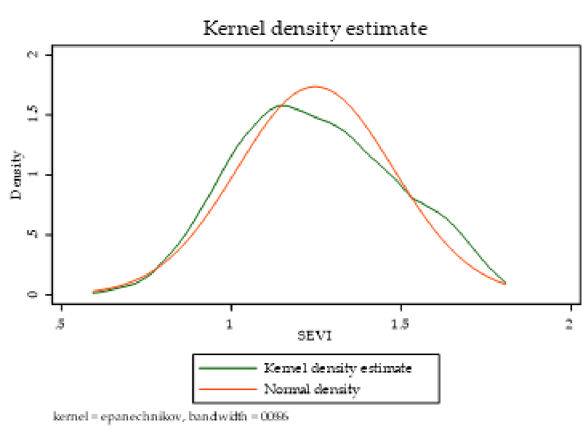
4.9 Appendix C: Definitions and sources of the variables for multilevel analysis

Table 4.10

Definitions and sources of the variables for multilevel analysis	
Poverty. Definitions and sources of the variables for multilevel analysis.	1.1 **Source:**Eurostat, Income and living conditions (ilc_mddd21) . 1.2 **Geographic level:** Country level for BE, DE and FR, NUTS-1 for PL and NUTS-2 for all other countries. 1.3 **Date of data used:** 2018.
Education. Government expenditure in education (percentage of gross domestic product).	2.1 **Source:**Eurostat, General government expenditure by function (COFOG) (gov_10a_exp). 2.2 **Geographic level:** NUTS-2. 2.3 **Date of data used:** Average 2017-2018.
Stability. Estimator of governance that measures the perceptions of political stability and absence of politically-motivated violence, including terrorism. Ranges from approximately -2.5 weak to 2.5 strong governance performance.	2.1 **Source:** OECD, Worldwide Governance Indicators. 2.2 **Geographic level:** Country. 2.3 **Date of data used:** 2018.

4.10 Appendix D: Density function of socio-economic vulnerability index (SEVI) of EU regions (N = 233)

Figure 4.6



Chapter 5

Conclusions

The purpose of this thesis is to model a new methodology based on a substantial improvement of the well-known Pena Trapero distance. Initially, a literature review is carried out in which 4 methodologies using very different mathematical tools (DEA, PCA, Mazziota Pareto Index y Distance P2) are analyzed. The literature already highlights the weaknesses and strengths of each of them. In the second chapter where these issues are addressed, a comparative study is carried out using the Regional Well-Being dataset as an example.

The results show that DEA methodology is stable when observations are eliminated, i.e. the ranking of observations remains unchanged when some observations are removed. On the other hand, the addition of indicators with high correlation with respect to the initial ones does not generate perturbations and this new indicator added is irrelevant in the model. However, this methodology was initially designed to study the efficiency of each observation separately. The assignment of different weights for each observation makes the result obtained not comparable. The main criticism of this aggregate approach is that it disregards the neutrality principle of social choice theory, i.e., that all observations (e.g., countries) should be treated equally and, therefore, it is not a good method for synthesizing indicators.

On the other hand, PCA is a methodology to reduce the dimensionality of multidimensional approaches. On this occasion, synthetic indicators called components are obtained, which are linear combinations of the initial indicators. All the components have the property that they are orthogonal to each other, and therefore, maximize the variance between them. However, researchers use those components that provide the most explained variability and omit those that provide the least information. This can be a problem because the composite indicator does not fully synthesize the concept to be measured. Another difficulty with this methodology is that negative weights appear (the weights are the eigenvectors of the covariance matrix calculated among all the indicators). Therefore, PCA can be used for cases in which the elicitation of weights is not the main goal. The use of this methodology is also restricted to indicators that have the same polarity. Although the elimination of observations does not generate alterations in the range, the introduction of highly correlated indicators does produce important perturbations in the weights of each indicator.

MPI indicator is another aggregation method for summarizing a set of individual indicators that are assumed to be not fully substitutable. The penalty defined in this approach will not always act as catalyst for imbalances between indicators, and sometimes can be difficult to understand. In this case, this methodology is sensitive to the elimination of observations.

Finally, the Distance P2 calculates the composite index from a metric. This metric is defined on the vector space of $\mathbf{R}^n$, where n represents the number of indicators to be synthesized. This structure is the most appropriate to compare observations since it satisfies the trinagular inequality, and therefore, allows to know which observations are closer to a given reference system. In our opinion, this methodology is the most appropriate for the construction of composite indicators. The reasons are multiple. Firstly, because of its metric nature. Secondly, because it eliminates the redundancy of information that can be collected by the composite indicator. The weights define the information attributed to each of the simple indicators that make up the composite indicator. Finally, it is a stable method when faced with the elimination of observations and the aggregation of highly correlated simple indicators.

However, the P2 Distance also has some limitations. The first is that it is not quite a metric. The possibility that any of the weights is zero, that is, that there is multicollinearity among the simple indicators, although highly improbable, does not meet one of the requirements for it to be a distance. The second is that the weights are calculated from linear models (OLS). This limitation explains other types of functional dependencies and restricts their use to the fulfillment of the hypotheses of an OLS. This is the fundamental motivation that raises the possibility of correcting these weaknesses.

In accordance with the foregoing, the second contribution of this thesis provides a new approach in order to obtain a composite indicator (`DL2`). Starting from a baseline observation, the proposed methodology proposes a metric whose weights are calculated through machine learning. In this case, the weights are always strictly positive, therefore, the weighted ℓ^2 formula 3.5 defines a metric o distance. On the second hand, the first step of this algorithm generates a composite indicator by the unweighted metric, which is used to optimizes the best functional relationship with respect to the single indicators. In addition, `DL2` ranks the single indicators in order of importance by assigning weights to the metric based on the previous relationships. This tool has the following properties of monotonicity, invariance by origin and scale charges, transitivity, homogeneity and symmetry. Monte Carlo simulations was implemented in order to confirm robust of the method. The method detects and delete multicollinearity problems among the single indicators. Likewise, the weighted scheme is not altered by permutations in the rank in which the single indicators are computed.

The final phase of this thesis aims to obtain an application of the proposed methodology. For this purpose, we analyze the socio-economic vulnerability of the European Union regions and the COVID-19 shock. The EU Cohesion Policy for the period 2021-2027 focuses on five goals for the EU to become smarter, greener, more connected, more social and closer to citizens. However, a macroeconomic index (per capita GD) is proposed as the predominant criterion for classifying the regions and allocating the Structural Funds. We hypothesise that it is possible to consider new complementary criteria that better reflect citizens' quality of life. This approach is especially important because the COVID-19 has exposed the vulnerabilities within the EU in all the domains: jobs, education, economy, welfare systems and social life). On this basis, we have built a composite socio-economic vulnerability index (SEVI) for each of the 233 NUTS-2 of the EU in 2017 that synthesises the information on fragility and resilience factors in order to achieve the objectives of the 2021-2027 Cohesion Policy. The idea is that the higher the value of SEVI, the greater the difficulty in achieving these objectives compared to the rest of the regions.

In a second stage, we study the effects of COVID-19 on regional vulnerability since it is foreseeable that the pandemic will trigger inequalities and increase poverty levels. The question is whether all regions will be equally exposed to COVID-19 in terms of their socio-economic vulnerability. To answer this question, we analyse both the idiosyncratic and covariate shocks that COVID-19 might represent by estimating multi-level models. Our findings indicate that increases in government expenditures in education and improving political stability would reduce the regional vulnerability or foster the capacity for resilience. On the other hand, increases in regional monetary poverty would be associated with increased vulnerability, causing bigger growth in the regions with a higher level of vulnerability. Even though regions with a larger SEVI would be the most exposed to the effects of COVID-19, it is remarkable that much of the territory of Sweden and Finland and the region of Ireland that ranked highest in the SEVI would be among the most exposed to the covariate effects. These results might have public policy implications; for example, to inform on how to distribute the European COVID-19 Recovery Funds.

The multidimensional character of the proposal, the study of regions? factors of vulnerability, fragility and resilience, fits into the mainstream view of economists and policymakers who argue that associating the notion of economic and social progress to a one-dimensional variable of economic activity, such as GDP or income, is debatable. Our proposal is also in line with two plans or strategies that the European Commission has recently approved to continue advancing in the double green and digital transition and to recover from the COVID-19 crisis: the European Pillar of Social Rights Action Plan (European Commission, 2021c) and the 2020 Strategic Foresight Report. The objective of both plans is to promote resilience through the EU institutions so that Europe will recover faster and emerge stronger from the COVID-19 crisis and future crises . A key aspect is that the priorities identified in both plans must be taken into account in all EU policymaking, including the 2021-2027 Cohesion Policy.

Finally, we have argued that, in terms of Cohesion Policy, there is still room to go “beyond GDP” and consider, in financial terms, the vulnerability and resilience factors that determine people’s well-being. The two previous initiatives or strategies lead us to be hopeful and to think that the EU will continue to advance on the “beyond GDP” path by strengthening the principles of a social market economy. Likewise, the type of exercise carried out in the second stage of this paper can be useful to stimulate discussions regarding the guidelines on how to increase the resilience and reduce the fragility of the regions in order to cope with unforeseen shocks.

The methodology presented in this thesis has some limitations. The first is that it does not admit indicators of a qualitative nature. This limitation is important and is currently being addressed in order to finally provide a complete methodology to synthesize all kinds of indicators. When the set of observations is not very large, the algorithm is able to obtain nonlinear functional relationships that minimize the error so that some simple indicators are excluded from the model. In some studies, this limitation may exclude variables or indicators that are empirically relevant to the experiment and as a result, the composite indicator obtained may provide unrealistic explanations of the phenomenon to be synthesized. This limitation is also being studied.

Synthesizing a multidimensional problem into a single indicator is a complex task. The complexity of the methodology presented in this study is the result of the intrinsic difficulty associated with synthesizing a given phenomenon into an indicator. In order to make the use of this method simple for a researcher, packages are being built for different software (R, Python and Stata).

Chapter 6

Bibliography

6.1 Introduction Bibliography

1. Bandura, R. (2008). *A survey of composite indices measuring country performance: 2008 update.* Technical report, Office of Development Studies, United Nations Development Programme (UNDP), New York.

2. Charnes, A., Cooper, W.,Lewin A. Y. & Seiford, L. M. (1997) *Data Envelopment Analysis Theory, Methodology and Applications*, Journal of the Operational Research Society, 48:3, 332-333, DOI: 10.1057/palgrave.jors.2600342

3. Decancq, K., & Lugo, M. A. (2013). *Weights in multidimensional indices of wellbeing: An overview. Econometric Reviews, 32*(1), 7-34.

4. De Muro, P., Mazziotta, M. & Pareto, A. (2011). Composite Indices of Development and Poverty: An Application to MDGs. *Social Indicators Research, 104*, 1-18.

5. Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the Methodological Framework of Cmposite Indices: A Review of the issues of Weighting, Aggregation, and Robustness. Social Indicators Research, 141, 61-94. DOI: 10.1007/s11205-017-1832-9

6. Greyling, T., & Tregenna, F. (2016). Construction and analysis of a composite quality of life index for a region of South Africa. *Social Indicators Research, 131*(3), 88-930.

7. Maggino, F. (2017). *Complexity in Society: From Indicators Construction to their Synthesis* . Cham, Switzerland: Springer International Publishing. DOI: 10.1007/978-3-319-60595-1

8. Mazziotta, M., & Pareto, A. (2018). *Measuring Well-Being Over Time: The Adjusted Mazziotta-Pareto Index Versus Other Non-compensatory Indices.* A. Soc Indic Res, 136 (3), 967-976. https://doi.org/10.1007/s11205-017-1577-5

9. Mazziotta, M., & Pareto, A. (2019). Use and Misuse of PCA for Measuring Well-Being. *Social Indicators Research, 142*, 451-476.

10. OECD (2008). *Handbook on Constructing Composite Indicators. Methodology and user guide.* OECD Publications, Paris.

11. Pena Trapero, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines (Una aplicación al caso español).* Madrid, Spain: INE.

12. Saisana, M., & Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development*. European Commission, Joint Research Centre, Institute for the Protection and the Security of the Citizen, Technological and Economic Risk Management Unit, Ispra, Italy.

13. Sánchez, A. & Ruiz-Martos, M. (2018). Europe 2020 Strategy and Citizens? Life Satisfaction. *Journal of Happiness Studies*. DOI: 10.1007/s10902-017-9928-0

14. Yang, F.-C., Kao, R.-H., Chen, Y.-T., Ho Y.-F., Cho, C.-C., & Hung, S.-W. (2017). A common weight approach to construct composite indicators: The evaluation of fourteen emerging markets. *Social Indicators Research*. https://doi.org/10.1007/s11205-017-1603-7.

6.2 Chapter 1 Bibliography

1. Bandura, R. (2008). *A survey of composite indices measuring country performance: 2008 update*. Technical report, Office of Development Studies, United Nations Development Programme (UNDP), New York.

2. Bodenhofer, U. & Klawonn, F. (2008). Robust rank correlation coefficients on the basis of fuzzy orderings: initial steps. *Mathware & Soft Computing 15*, 5-20.

3. Bandura, R. (2008). *A survey of composite indices measuring country performance: 2008 update*. Technical report, Office of Development Studies, United Nations Development Programme (UNDP), New York.

4. Bodenhofer, U. & Klawonn, F. (2008). Robust rank correlation coefficients on the basis of fuzzy orderings: initial steps. *Mathware & Soft Computing 15*, 5-20.

5. Cauchy AL. (1829) *Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes*. vol. 9. O'euvres Complètes (IIéme Série); Paris: Blanchard; 1829.

6. Charnes, A., Cooper, W.,Lewin A. Y. & Seiford, L. M. (1997) *Data Envelopment Analysis Theory, Methodology and Applications*, Journal of the Operational Research Society, 48:3, 332-333, DOI: 10.1057/palgrave.jors.2600342

7. Cherchye, L., Moesen, W., Rogge, N., & Van Puyenbroeck, T. (2007). *An introduction to "benefit of the doubt" composite indicator*. Social Indicators Research, 82(1), 111-145.

8. Decancq, K., & Lugo, M. A. (2013). *Weights in multidimensional indices of wellbeing: An overview. Econometric Reviews, 32*(1), 7–34.

9. Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*, 269-277.

10. Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). *Advancing formative measurement models. Journal of Business Research, 61*, 1203–1218.

11. Fattore, M. (2016). Partially ordered sets and the measurement of multidimensional ordinal deprivation. Social Indicators Research, 128(2), 835-858. DOI:10.1007/s11205-015-1059-6.

12. González, E., Cárcaba, A, & Ventura, J. (2010) The Importance of the Geographic Level of Analysis in the Assessment of the Quality of Life: The Case of Spain. *Social Indicators Research. 102*, 209-228.

13. Grabisch, M., Marichal, J.L., Mesiar, R. & Pap, E,(2011) Aggregation functions: Means. *Information Sciences, Elsevier, 181* (1), 1-22.

14. Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the Methodological Framework of Cmposite Indices: A Review of the issues of Weighting, Aggregation, and Robustness. Social Indicators Research, 141, 61-94. DOI: 10.1007/s11205-017-1832-9

15. Greyling, T., & Tregenna, F. (2016). Construction and analysis of a composite quality of life index for a region of South Africa. *Social Indicators Research, 131*(3), 88-930.

16. Sasan Karamizadeh, S., Shahidan M. Abdullah, S. M., Azizah Abd. Manaf, A Zamani, M. & Hooman, A.(2019) Principal Component Analysis, *Encyclopedia of Biometrics.*

17. Kendall, M., Gibbons, J. D. (1990) [First published 1948]. *Rank Correlation Methods.* Charles Griffin Book Series (5th ed.). Oxford: Oxford University Press. ISBN 978-0195208375.

18. Mazziotta, M., Pareto, A. (2012). A non-compensatory approach for the measurement of the quality of life. In F. Maggino & G. Nuvolati (Eds.), Quality of life in Italy: Research and reflections (pp. 27-40). New York: Springer.

19. Mazziotta, M., & Pareto, A. (2013). *Methods for constructing composite indices: One for all or all for one?* Rivista Italiana di Economia Demografia e Statistica, LXVII(2), 67-80.

20. Mazziotta, M. & Pareto, A. (2017). Synthesis of Indicators: The Composite Indicators Approach (Chapter 7). In F. Maggino (ed.) Complexity in Society: From Indicators Construction to their Synthesis (pp. 115-138). Cham, Switzerland: Springer International Publishing. DOI: 10.1007/978-3-319-60595-1

21. Mazziotta, M., & Pareto, A. (2018). *Measuring Well-Being Over Time: The Adjusted Mazziotta-Pareto Index Versus Other Non-compensatory Indices.* A. Soc Indic Res, 136 (3), 967-976. https://doi.org/10.1007/s11205-017-1577-5

22. Mazziotta, M., & Pareto, A. (2019). Use and Misuse of PCA for Measuring Well-Being. *Social Indicators Research, 142*, 451-476.

23. Maggino, F. (2017). *Complexity in Society: From Indicators Construction to their Synthesis* . Cham, Switzerland: Springer International Publishing. DOI: 10.1007/978-3-319-60595-1

24. Maggino, F. (2017a). Dealing with Syntheses in a System of Indicators (Chapter 5). In F. Maggino (ed.) Complexity in Society: From Indicators Construction to their Synthesis (pp. 115-138). Cham, Switzerland: Springer International Publishing. DOI: 10.1007/978-3-319-60595-1

25. Maggino, F. (2017b). Developing Indicators and Managing the Complexity (Chapter 4). In F. Maggino (ed.) Complexity in Society: From Indicators Construction to their Synthesis (pp. 87-114). Cham, Switzerland: Springer International Publishing. DOI: 10.1007/978-3-319-60595-1

26. Manca, A.R., Benczur, P., & Giovannini, E. (2017). Building a Scientific Narrative Towards a More Resilient EU Society, Part 1: a conceptual framework. EUR 28548 EN, *Publications Office of the European Union*, Luxembourg. https://doi.org/10.2760/635528

27. Marical, F., D'Ercole, M., Vaalavuo, M., & Verbist, G. (2008). Publicly-provided Services and the Distribution of Households. *OECD Economic Studies, 44*(1), 1-38.

28. Marulanda Fraume, M.C., Cardona, O.D., Marulanda Fraume, P., Carreño, .L., & Barbat, A.H. (2020). Evaluating risk from a holistic perspective to improve resilience: The United Nations evaluation at global level. *Safety Science, 127*, 104739. https://doi.org/10.1016/j.ssci.2020.104739

29. Merino, M.D., Privado, J., & Arnaiz, R. (2019). Is There Any Relationship between Unemployment in Young Graduates and Psychological Resources? An Empirical Research from the Conservation of Resources Theory. *Journal of Work and Organizational Psychology, 35*(1). 1-8. http://scielo.isciii.es/pdf/rpto/v35n1/1576-5962-rpto-35-1-0001.pdf

30. Mina, C., & Imai, K. (2016). Estimation of vulnerability to poverty using a multilevel longitudinal model: Evidence from the Philippines. *The Journal of Development Studies, 53*(12), 2118-2144. https://doi.org/10.1080/00220388.2016.1265942

31. Nájera, H., & Gordon, D. (2019). The Importance of Reliability and Construct Validity in Multidimensional Poverty Measurement: An Illustration Using the Multidimensional Poverty Index for Latin America (MPI-LA). *The Journal of Development Studies, 56*(9), 1763-1783. https://doi.org/10.1080/00220388.2019.1663176

32. Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133-142.

33. O'Donnell, G., Deaton, A., Durand, M., Halpern, D., & Layard, R. (2014). Wellbeing and Policy (Report). Commissioned by the Legatum Institute.

34. Peichl, A., Pestel, N., & Schneider, H. (2012). Does Size Matter? The Impact of Changes in Household Structure on Income Distribution in Germany. *Review of Income and Wealth, 58*(1), 118-141.

35. Pena Trapero, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines (Una aplicación al caso español)*. Madrid, Spain: INE.

36. Povel, F. (2015) Measuring exposure to downside risk with an application to Thailand and Vietnam. World Development, 71, 4-24.

37. Romer, P. (1994). The origins of endogenous growth. *The Journal of Economics Perspectives, 8*, 3-22.

38. Saisana, M., & Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development.* European Commission, Joint Research Centre, Institute for the Protection and the Security of the Citizen, Technological and Economic Risk Management Unit, Ispra, Italy.

39. Sánchez, A., & Ruiz-Martos, M. (2018). Europe 2020 Strategy and Citizens? Life Satisfaction. *Journal of Happiness Studies.* DOI: 10.1007/s10902-017-9928-0

40. Sanchez2018b Sánchez, A., Chica-Olmo, J., and Jiménez-Aguilera, J.D. (2018). A Space-Time Study for Mapping Quality of Life in Andalusia During the Crisis. Social Indicators Research, 135(2), 699-728. DOI: 10.1007/s11205-016-1497-9

41. Sánchez, A. & Lopez-Corral, A. (2018). Government Social Expenditure and Income Inequalities in the European Union. *Hacienda Pública Española/Review of Public Economics, 227*(4), 135-158. https://doi.org/10.7866/HPE-RPE.18.4.5

42. Sánchez Mojica, B.E. (2013). A City torn apart: Forced displacement in Medellín, Colombia. International Law.*Revista Colombiana de Derecho Internacional, 22*, 179-210.

43. Scott2007 Scott, A. J., & Garofoli, G. (eds.) (2007). Development on the Ground. London: Routledge. Sen, A. (1976). Real National Income. *Review of Economic Studies, 43*, 19-39.

44. Shek, D. (2021). COVID-19 and Quality of Life: Twelve Reflections. Applied Research in Quality of Life, 16, 1-11. https://doi.org/10.1007/s11482-020-09898-z

45. Snijders, T.A.B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modelling* (2nd ed.). London: Sage Publishers.

46. Stiglitz, J., Fitoussi, J., & Durand, M. (eds.) (2018). *For Good Measure: Advancing Research on Well-being Metrics Beyond GDP*. Paris: OECD Publishing.
https://doi.org/10.1787/9789264307278-en

47. United Nations (2020a). Progress towards the Sustainable Development Goals Report of the Secretary-General. Economic and Social Council.
https://sustainabledevelopment.un.org/content/documents/26158Final_SG_SDG_Progress_Report_14052020.pdf

48. United Nations (2020b). Shared responsibility, global solidarity: Responding to the socio-economic impacts of COVID-19.
https://www.un.org/sites/un2.un.org/files/sg_report_socio-economic_impact_of_covid19.pdf

49. Van den Bergh, J. (2009). The GDP paradox. *Journal of Economic Psychology, 30*(2), 117-135.

50. World Health Organization. (2013). *Health effects of particulate matter.*
https://www.euro.who.int/__data/assets/pdf_file/0006/189051/Health-effects-of-particulate-matter-final-Eng.pdf

51. Zarzosa Espina, P., & Somarriba Arechavala, N. (2013). An assessment of social welfare in Spain: Territorial analysis using a synthetic welfare Indicator. *Social Indicators Research, 111*(1),1-23.
https://doi.org/10.1007/s11205-012-0005-0

52. Yang, F.-C., Kao, R.-H., Chen, Y.-T., Ho Y.-F., Cho, C.-C., & Hung, S.-W. (2017). A common weight approach to construct composite indicators: The evaluation of fourteen emerging markets. *Social Indicators Research.*
https://doi.org/10.1007/s11205-017-1603-7.

53. Yao, R., & Wu, W. (2021). Mental Disorders Associated with COVID-19 Related Unemployment. *Applied Research in Quality of Life.*
https://doi.org/10.1007/s11482-021-09950-6

6.3 Chapter 2 Bibliography

1. Alves Pereira, M., Santos Camanho, A., Figueira, J.R., & Cunha Marques, R. (2021). Incorporating preference information in a range directional composite indicator: The case of Portuguese public hospitals. *European Journal of Operational Research, 294*(2), 633-650.
https://doi.org/10.1016/j.ejor.2021.01.045.

2. Annoni, P., & Dijkstr, L. (2019). *The EU Regional Competitiveness Index 2019*. Luxembourg: Publications Office of the European Union.

3. Baser, O., & Pema, E., (2003). The Return of Publications for Economics Faculty. *Economics Bulletin 1*, 1-13.

4. Becker, W., Saisana, M., Paruolo, P., & Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological Indicators, 80*, 12-22.
https://doi.org/10.1016/j.ecolind.2017.03.056

5. Canadian Index of Wellbeing. (2016). *How are Canadians Really Doing? The 2016 CIW National Report.* Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

6. Coltman, T., Devinney, T.M., Midgley, D.F., & Venaik, S. (2008). Formative versus reflective measurement models: two applications of formative measurement. *Journal of Business Research, 61*(12), 1250-1262.

7. Craven, P., & Wahba, G. (1979). Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik, 31*, 377-403.

8. Cuenca-García, E., Sánchez, A., & Navarro-Pabsdorf, M. (2019). Assessing the performance of the least developed countries in terms of the Millennium Development Goals. *Evaluation and Program Planning, 72*, 54-66.
https://doi.org/10.1016/j.evalprogplan.2018.09.009

9. D'Inverno, G., & De Witte, K. (2020). Service level provision in municipalities: A flexible directional distance composite indicator. *European Journal of Operational Research, 286*(3).
https://doi.org/10.1016/j.ejor.2020.04.012.

10. Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research, 61*, 1203-1218.

11. Friedman, J. H., & Silverman, B. W. (1989). Flexible parsimonious smoothing and additive modeling. *Technometrics, 31*(1), 3-21.

12. Friedman, J.H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics, 19*(1), 1-141.

13. 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*, 61-94.
https://doi.org/10.1007/s11205-017-1832-9

14. Greenwell, B.M., Boehmke, B.C., & McCarthy, A.J. (2018). A Simple and Effective Model-Based Variable Importance Measure, arXiv:1805.04755v1 [stat.ML].

15. Jiménez-Fernández, E., & Ruiz-Martos, M.J. (2020). Review of some statistical methods for constructing composite indicators. *Estudios de Economía Aplicada-Studies of Applied Economy, 38*(1), 1-15.
http://doi.org/10.25115/eea.v38i1.3002

16. Kendall, M.G. (1976). *Rank Correlation Methods.* 4th Ed. Griffin.

17. Keogh, S., O'Neill, S., & Walsh, K. (2021). Composite Measures for Assessing Multidimensional Social Exclusion in Later Life: Conceptual and Methodological Challenges. *Social Indicators Research.*
https://doi.org/10.1007/s11205-021-02617-7

18. Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software, 28*(5), 1-26.
http://dx.doi.org/10.18637/jss.v028.i05

19. Lafuente, E., Araya, M., & Leiva, J.C. (2020). Assessment of local competitiveness: A composite indicator analysis of Costa Rican counties using the "Benefit of the Doubt" model. *Socio-Economic Planning Sciences*, 100864.
https://doi.org/10.1016/j.seps.2020.100864.

20. Luque, M., Pérez-Moreno, S., Robles, J. A. & Rodríguez, B. (2017) Measuring Child and Maternal Health in Developing Countries: A Proposal of New Hybrid MDG Composite Indices. *Applied Research in Quality of Life, 12*(3), 737-758.

21. Maggino, F. (2017). Developing Indicators and Managing the Complexity (Chapter 4). In F. Maggino (ed.) *Complexity in Society: From Indicators Construction to their Synthesis* (pp. 87-114). Springer International Publishing. DOI: 10.1007/978-3-319-60595-1

22. Martín, J.M., Salinas Fernández, J.A., Rodríguez Martín, J.A., & Ostos Rey, S. (2020). Analysis of Tourism Seasonality as a Factor Limiting the Sustainable Development of Rural Areas. *Journal of Hospitality & Tourism Research, 44*(1).
https://doi.org/10.1177/1096348019876688

23. Mazziotta, M., & Pareto, A. (2017). Synthesis of Indicators: The Composite Indicators Approach (Chapter 7). In F. Maggino (ed.) *Complexity in Society: From Indicators Construction to their Synthesis* (pp. 159-192).

24. Mazziotta, M., & Pareto, A. (2019). Use and Misuse of PCA for Measuring Well-Being. *Social Indicators Research, 142*(2), 451-476.
https://doi.org/10.1007/s11205-018-1933-0

25. Montero, J.-M., Chasco, C., & 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.

26. OECD. (2008). *Handbook on Constructing Composite Indicators. Methodology and User Guide.* Paris: OECD publishing. DOI:10.1787/9789264043466-en

27. OECD. (2017). *How's Life? 2017: Measuring Well-being.* OECD Publishing. DOI: 10.1787/how-life-2017-en

28. Pena Trapero, J.B. (1977). *Problemas de la medición del bienestar y conceptos afines. Una aplicación al caso español (Problems of welfare measurement and related concepts. An application to the Spanish case).* Madrid, Spain: INE.

29. Pena Trapero, J.B. (2009). La medición del bienestar social: una revisión crítica(Measuring social welfare: a critical review). *Estudios de Economía Aplicada/Studies of Applied Economy, 27*(2), 299-324.

30. Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., & Fuller, G. (2018). *SDG indicator and Dashboards Report 2018.* New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).

31. Sánchez, A., Chica-Olmo, J., & Jiménez-Aguilera, J. D. (2018). A space-Time study for mapping quality of life in Andalusia during the crisis. *Social Indicators Research, 135*(2), 699-728.
https://doi.org/10.1007/s11205-016-1497-9

32. Sánchez, A., & Ruiz-Martos, M. (2018). Europe 2020 Strategy and Citizens' Life Satisfaction. *Journal of Happiness Studies, 19*, 2315?2338.
https://doi.org/10.1007/s10902-017-9928-0

33. Schwab, K., & Porter, M. (2008). *The Global Competitiveness Report 2008?2009.* Geneva, Switzerland: World Economic Forum.

34. Somarriba, N., & Pena, B. (2009). Synthetic indicators of quality of life in Europe. *Social Indicators Research, 94*(1), 115-133.

35. UNDP. (2018). *Human Development Indices and Indicators*. New York: United Nations Development Programme.

36. World Bank. (2020). *Doing Business 2020. Comparing Business Regulation in 190 Economies*. Washington, DC: International Bank for Reconstruction and Development / The World Bank.

37. Zarzosa Espina, P. (1996). *Aproximación a la medición del bienestar social (Approach to the measurement of social welfare)*. University of Valladolid.

6.4 Chapter 3 Bibliography

1. Acconcia, A., Carannante, M., Misuraca, M., & Scepi, G. (2020). Measuring Vulnerability to Poverty with Latent Transition Analysis. *Social Indicators Research, 151*, 1?31.
https://doi.org/10.1007/s11205-020-02362-3

2. Alessi, L., Benczur, P., Campolongo, F. et al. (2020). The Resilience of EU Member States to the Financial and Economic Crisis. *Social Indicators Research, 148*, 569-598. https://doi.org/10.1007/s11205-019-02200-1

3. Aschauer, D.A. (1989). Is public expenditure productive?. *Journal of Monetary Economics, 23*, 177-200.

4. Azeem, M., Mugera, A., & Schilizzi, S. (2016). Poverty and vulnerability in the Punjab, Pakistan: A multilevel analysis.*Journal of Asian Economics, 44*, 57-72.

5. Benczur, P., Joossens, E., Manca, A.R., Menyhert, B, & Zec, S. (2020). How resilient are the European regions: Evidence from the societal response to the 2008 financial crisis. *EUR 30352 EN, Publications Office of the European Union*, Luxembourg.
https://doi.org/10.2760/383460,JRC121554.

6. Bittmann, F. (2021). How Trust Makes a Difference: The Impact of the First Wave of the COVID-19 Pandemic on Life Satisfaction in Germany. *Applied Research in Quality of Life*.
https://doi.org/10.1007/s11482-021-09956-0

7. Bonaccorsia, G., Pierrib, F., Cinellic, M. et al. (2020). Economic and social consequences of human mobility restrictions under COVID-19. *PNAS, 117*(27), 15530-15535.
https://doi.org/10.1073/pnas.2007658117

8. Cardona, O.D. (2004). The Need for Rethinking the Concepts of Vulnerability and Risk from a Holistic Perspective: a Necessary Review and Criticism for Effective Risk Management. In G. Bankoff, G. Frerks, D. Hilhorst (Eds), *Mapping Vulnerability: Disasters, Development and People* (chapter 3). Earthscan Publishers.

9. Chadi, A. (2014). Regional unemployment and norm-induced effects on life satisfaction. *Empirical Economics, 46*, 1111-1141.
https://doi.org/10.1007/s00181-013-0712-7

10. Chang, H.J. (2011). Institutions and Economic Development: Theory, Policy and History. *Journal of Institutional Economics, 7*(4), 473-498.
https://doi.org/10.1017/S1744137410000378

11. Chica-Olmo, J., Sánchez, A. & Sepúlveda-Murillo, F.H. (2020). Assessing Colombia's policy of socio-economic stratification: An intra-city study of self-reported quality of life. *Cities, 97*, 102560.
https://doi.org/10.1016/j.cities.2019.102560

12. Commission of the European Communities (2009). Communication from the Commission to the Council and the European Parliament. *GDP and beyond. Measuring progress in a changing world*. COM(2009) 433 final. Brussels.

13. Cuenca-García, E., Sánchez, A., & Navarro-Pabsdorf, M. (2019). Assessing the performance of the least developed countries in terms of the Millennium Development Goals.*Evaluation and Program Planning, 72*, 54-66.
https://doi.org/10.1016/j.evalprogplan.2018.09.009

14. Despotis2005 Despotis, D. K. (2005a). A reassessment of the human development index via data envelopment analysis. *Journal of the Operational Research Society, 56*(8), 969-980.

15. Dessimirova, D., & Bustamante, M.A. (2019). The gender gap in pensions in the EU. European Parliament. Policy Department for Economic, *Scientific and Quality of Life Policies*.
https://doi.org/10.2861/20375

16. Diamantopoulos2008 Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). *Advancing formative measurement models. Journal of Business Research, 61*, 1203-1218.

17. Dutta, I., Foster, J., & Mishra, A. (2011). On measuring vulnerability to poverty. *Social Choice and Welfare, 37*(4), 743-761.

18. European Commission. (2010). *EUROPE 2020 a strategy for smart, sustainable and inclusive growth*. COM(2010) 2020. Brussels.

19. European Commission. (2011). *Horizon 2020-The Framework Programme for Research and Innovation*. COM(2011) 808 final. Brussels.

20. European Commission. (2018). *Proposal for a regulation of the European Parliament and of the Council on the European Regional Development Fund and on the Cohesion Fund. 2018/0197(COD)*. COM(2018)372 final. Strasbourg.

21. European Commission. (2020a). *Annex to the amended proposal for a council regulation laying down the multiannual financial framework for the years 2021 to 2027. COM(2020)443* final. Brussels, 28.5.2020

22. European Commission. (2020b). 2020 *Strategic Foresight Report. Charting the course towards a more resilient Europe. COM(2020)* 493 final. Brussels.

23. European Commission. (2020c). *Strategic Plan 2020-2024. DG Research and Innovation*.
https://ec.europa.eu/info/sites/default/files/rtd_sp_2020_2024_en.pdf

24. European Commission. (2021a). *Annual Sustainable Growth Strategy. COM(2020) 575 final*. Brussels.

25. European Commission. (2021b). *Regional Innovation Scoreboard 2021. Publications Office of the European Union*.
https://ec.europa.eu/docsroom/documents/45974

26. European Commission. (2021c). *The European Pillar of Social Rights Action Plan. Publication Office of European Union*.
file:///C:/Users/Usuario/AppData/Local/Temp/KE0921008ENN-1.pdf

27. European Court of Auditors. (2019). *Rapid case review Allocation of Cohesion policy funding to Member States for 2021-2027*. European Union
https://www.eca.europa.eu/lists/ecadocuments/rcr_cohesion/rcr_cohesion_en.pdf

28. Fasani, F., & Mazza, J., A. (2020). Vulnerable Workforce: Migrant Workers in the COVID-19 Pandemic. Join Research Center, *Publications Office of the European Union*. https://doi.org/10.2760/914810, JRC120730.

29. Fetting, C. (2020). *Impacts of the COVID-19 Pandemic on Sustainable Development and the SDGs in Europe*. ESDN Report, July 2020. Vienna: ESDN Office.

30. Gallardo, M. (2018). Identifying vulnerability to poverty: A critical survey. *Journal of Economic Surveys, 32*(4), 1074-1105. https://doi.org/10.1111/joes.12216

31. Gallardo, M. (2020). Measuring Vulnerability to Multidimensional Poverty. Social Indicators Research, 148, 67-103. https://doi.org/10.1007/s11205-019-02192-y

32. Gramillano, A., Celotti, P., Familiari, G., Schuh, B., & Nordstrom, M. (2018). *Development of a system of common indicators for European Regional Development Fund and Cohesion Fund interventions after 2020*. Publications Office of the European Union.

33. Grossman, G.M., & Helpman, E. (1994). Endogeneus innovation in the theory of growth. *The Journal of Economics Perspectives, 8*, 23-44.

34. Halkos, G., Skouloudis, A., Malesios, C., & Jones, N. (2020). A Hierarchical Multilevel Approach in Assessing Factors Explaining Country-Level Climate Change Vulnerability. *Sustainability, 12*, 4438. https://doi.org/10.3390/su12114438

35. Helliwell, J.F., & Huang, H. (2014). New measures of the costs of unemployment: Evidence from the subjective well-being of 3.3 million Americans. *Economic Inquiry, 52*(4), 1485-1502.

36. Ikeda, T., Igarashi, A., Odani, S., Murakami, M., & Tabuchi, T. (2021). Health-Related Quality of Life during COVID-19 Pandemic: Assessing Impacts of Job Loss and Financial Support Programs in Japan. *Applied Research in Quality of Life*. https://doi.org/10.1007/s11482-021-09918-6

37. Jiménez-Fernández, E., Sánchez, A., & Ortega-Pérez, M. (2022). Dealing with weighting scheme in composite indicators: an unsupervised distance-machine learning proposal for quantitative data. *Socio-Economic Planning Sciences, 101339* https://doi.org/10.1016/j.seps.2022.101339

38. Jiménez-Fernández, E.,& Ruiz-Martos, M.J. (2020). Review of some statistical methods for constructing composite indicators. *Estudios de Economía aplicada-Studies of applied economy, 38*(1), 1-15.

39. Sasan Karamizadeh, S., Shahidan M. Abdullah, S. M., Azizah Abd. Manaf, A Zamani, M. & Hooman, A.(2019) Principal Component Analysis, *Encyclopedia of Biometrics*.

40. Keogh-Brown, M.R., & Smith, R.D. (2008). The economic impact of SARS: How does the reality match the predictions?. *Health Policy, 88*(1), 110-120. https://doi.org/10.1016/j.healthpol.2008.03.003

41. Kollmeyer, C. (2013). Family Structure, Female Employment, and National Income Inequality: A Cross-National Study of 16 Western Countries. *European Sociological Review, 29*(4), 816-827.

42. Le Blanc, J. (2020). Financial buffers of households in the wake of the COVID-19 crisis. *JRC Science for Policy Report*, JCR120733.

43. Lee, J.W., & McKibbin, W.J. (2004). Globalization and Disease: The Case of SARS. Asian Economic Papers, 3(1), 113-131.
https://doi.org/10.1162/1535351041747932

44. Lucas, R.E. (1993). Making a miracle. *Econometrica, 61*(2), 251-272.

45. Maggino, F. (2017b). Developing Indicators and Managing the Complexity (Chapter 4). In F. Maggino (ed.) Complexity in Society: From Indicators Construction to their Synthesis (pp. 87-114). Cham, Switzerland: Springer International Publishing. DOI: 10.1007/978-3-319-60595-1

46. Manca, A.R., Benczur, P., & Giovannini, E. (2017). *Building a Scientific Narrative Towards a More Resilient EU Society*, Part 1: a conceptual framework. EUR 28548 EN, Publications Office of the European Union, Luxembourg.
https://doi.org/10.2760/635528

47. Marical, F., D'Ercole, M., Vaalavuo, M., & Verbist, G. (2008). Publicly-provided Services and the Distribution of Households. *OECD Economic Studies, 44*(1), 1-38.

48. Marulanda Fraume, M.C., Cardona, O.D., Marulanda Fraume, P., Carreño, .L., & Barbat, A.H. (2020). Evaluating risk from a holistic perspective to improve resilience: The United Nations evaluation at global level. *Safety Science, 127*, 104739.
https://doi.org/10.1016/j.ssci.2020.104739.

49. Merino, M.D., Privado, J., & Arnaiz, R. (2019). Is There Any Relationship between Unemployment in Young Graduates and Psychological Resources? An Empirical Research from the Conservation of Resources Theory. Journal of Work and Organizational Psychology, 35(1). 1-8.
http://scielo.isciii.es/pdf/rpto/v35n1/1576-5962-rpto-35-1-0001.pdf

50. Mina, C., & Imai, K. (2016). Estimation of vulnerability to poverty using a multilevel longitudinal model: Evidence from the Philippines. *The Journal of Development Studies, 53*(12), 2118-2144.
https://doi.org/10.1080/00220388.2016.1265942

51. Nájera, H., & Gordon, D. (2019). The Importance of Reliability and Construct Validity in Multidimensional Poverty Measurement: An Illustration Using the Multidimensional Poverty Index for Latin America (MPI-LA). *The Journal of Development Studies, 56*(9), 1763-1783.
https://doi.org/10.1080/00220388.2019.1663176

52. Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133-142.

53. O'Donnell, G., Deaton, A., Durand, M., Halpern, D., & Layard, R. (2014). *Wellbeing and Policy* (Report). Commissioned by the Legatum Institute.

54. Peichl, A., Pestel, N., & Schneider, H. (2012). Does Size Matter? The Impact of Changes in Household Structure on Income Distribution in Germany. *Review of Income and Wealth, 58*(1), 118-141.

55. Pena Trapero, J.B. (1977) *Problemas de la medición del bienestar y conceptos afines (Una aplicacóón al caso español)*, INE, Madrid.

56. Povel, F. (2015) Measuring exposure to downside risk with an application to Thailand and Vietnam. *World Development, 71*, 4-24.

57. Romer, P. (1994). The origins of endogenous growth. *The Journal of Economics Perspectives, 8*, 3-22.

58. Sánchez, A. & Lopez-Corral, A. (2018). Government Social Expenditure and Income Inequalities in the European Union. *Hacienda Pública Española/Review of Public Economics, 227*(4), 135-158. https://doi.org/10.7866/HPE-RPE.18.4.5

59. Sánchez, A., & Ruiz-Martos, M. (2018). Europe 2020 Strategy and Citizens? Life Satisfaction. Journal of Happiness Studies. DOI: 10.1007/s10902-017-9928-0

60. Sánchez, A., Chica-Olmo, J., and Jiménez-Aguilera, J.D. (2018). A Space?Time Study for Mapping Quality of Life in Andalusia During the Crisis. Social Indicators Research, 135(2), 699-728. DOI: 10.1007/s11205-016-1497-9

61. Sánchez Mojica, B.E. (2013). A City torn apart: Forced displacement in Medellín, Colombia. International Law. *Revista Colombiana de Derecho Internacional, 22*, 179?210.

62. Scott, A. J., & Garofoli, G. (eds.) (2007). Development on the Ground. London: Routledge. Sen, A. (1976). Real National Income. *Review of Economic Studies, 43*, 19-39.

63. Shek, D. (2021). COVID-19 and Quality of Life: Twelve Reflections. *Applied Research in Quality of Life, 16*, 1-11. https://doi.org/10.1007/s11482-020-09898-z

64. Snijders, T.A.B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modelling* (2nd ed.). London: Sage Publishers.

65. Stiglitz, J., Fitoussi, J., & Durand, M. (eds.) (2018). For Good Measure: Advancing Research on Well-being Metrics Beyond GDP. Paris: OECD Publishing. https://doi.org/10.1787/9789264307278-en

66. Yang, F.-C., Kao, R.-H., Chen, Y.-T., Ho, Y.-F., Cho, C.-C., & Huang, S.-W. (2017). *A common weight approach to construct composite indicators: The evaluation of fourteen emerging markets.* Social Indicators Research. http s://doi.org/10.1007 /s112 05-017-1603 -7. (advance online publication).

67. United Nations (2020a). Progress towards the Sustainable Development Goals Report of the Secretary-General. Economic and Social Council. https://sustainabledevelopment.un.org/content/documents/26158Final_SG_SDG_Progress_Report_14052020.pdf

68. United Nations (2020b). Shared responsibility, global solidarity: Responding to the socio-economic impacts of COVID-19. https://www.un.org/sites/un2.un.org/files/sg_report_socio-economic_impact_of_covid19.pdf

69. Van den Bergh, J. (2009). The GDP paradox. *Journal of Economic Psychology, 30*(2), 117-135.

70. World Health Organization. (2013). Health effects of particulate matter. https://www.euro.who.int/__data/assets/pdf_file/0006/189051/Health-effects-of-particulate-matter-final-Eng.pdf

71. Zarzosa Espina, P., & Somarriba Arechavala, N. (2013). An assessment of social welfare in Spain: Territorial analysis using a synthetic welfare Indicator. *Social Indicators Research, 111*(1), 1-23. https://doi.org/10.1007/s11205-012-0005-0

72. Yao, R., & Wu, W. (2021). Mental Disorders Associated with COVID-19 Related Unemployment. *Applied Research in Quality of Life.* https://doi.org/10.1007/s11482-021-09950-6

Chapter 7

First page of the journals included in this thesis

Estudios de Economía Aplicada

VOL. 38-1 2020 PP.1-15

REVIEW OF SOME STATISTICAL METHODS FOR CONSTRUCTING COMPOSITE INDICATORS

E. JIMENEZ-FERNANDEZ

Departament d'Economia, UNIVERSITAT JAUME I, SPAIN
E-mail: jimeneze@uji.es

MARIA J. RUIZ-MARTOS

Departamento de Teoría e Historia Económica, UNIVERSIDAD DE GRANADA, SPAIN
E-mail: ruizmartosmaria@ugr.es

Recibido: 31 de Diciembre de 2019
Aceptado: 26 de Enero de 2020

ABSTRACT

The methodology for the construction process of composite indicators is reviewed in a step-by-step approach ranging from the ex-ante definition of the latent variable that is intended to be measured through the aggregation process. We focus on comparing four statistical aggregations methods in terms of their weighting and aggregation approaches: Distance P_2, Principal Component Analysis, Data Envelopment Analysis and Mazziotta-Pareto Index. An empirical comparison among them is provided and the composite indicators divergences are discussed.

Keywords: Composite indicators, Weighting, Aggregation, DEA-BoD, PCA, Distance P2, Mazziotta Pareto Index.

RESUMEN

La metodología para el proceso de construcción de indicadores compuestos se ha examinado a través de un enfoque gradual que va desde la definición de la variable latente que se pretende medir hasta el proceso de agregación. En particular, nos centramos en la comparación de cuatro métodos de agregación estadística respecto de sus enfoques de ponderación y agregación: Distancia P2, Análisis de Componentes Principales, Análisis de Envolvente de Datos e Índice de Mazziotta-Pareto. Adicionalmente, se proporciona una comparación empírica entre ellos y se examinan las divergencias de los indicadores compuestos.

Palabras clave: Indicadores compuestos, Ponderación, Agregación, DEA-BoD, PCA, Distancia P2, Índice de Pareto de Mazziotta.

JEL Classification Codes: I30, C13, C51, C55

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps

Dealing with weighting scheme in composite indicators: An unsupervised distance-machine learning proposal for quantitative data

Eduardo Jiménez-Fernández [a], Angeles Sánchez [b,*], Mario Ortega-Pérez [c]

[a] Departamento de Teoría e Historia Económica, Campus Cartuja s/n. Universidad de Granada, 18071, Granada, Spain
[b] Departamento de Economía Aplicada, Campus Cartuja s/n. Universidad de Granada, 18071, Granada, Spain
[c] Institute for Research and Innovation in Bioengineering, Camino de Vera, s/n. Universitat Politècnica de València, 46022, València, Spain

ARTICLE INFO

Keywords:
Composite indicator
P2 distance
Unsupervised machine learning
Benchmarking
Weighting scheme
MARS
PACS
C02
C15
C44
C43

ABSTRACT

There is increasing interest in the construction of composite indicators to benchmark units. However, the mathematical approach on which the most commonly used techniques are based does not allow benchmarking in a reliable way. Additionally, the determination of the weighting scheme in the composite indicators remains one of the most troubling issues. Using the vector space formed by all the observations, we propose a new method for building composite indicators: a distance or metric that considers the concept of proximity among units. This approach enables comparisons between the units being studied, which are always quantitative. To this end, we take the P2 Distance method of Pena Trapero as a starting point and improve its limitations. The proposed methodology eliminates the linear dependence on the model and seeks functional relationships that enable constructing the most efficient model. This approach reduces researcher subjectivity by assigning the weighting scheme with unsupervised machine learning techniques. Monte Carlo simulations confirm that the proposed methodology is robust.

1. Introduction

Composite indicators have clear advantages that justify their increasing use for summarising complex and multidimensional realities that are not directly measurable. For instance, they are used to support decision makers, make comparisons and assess the progress of units (companies, countries, regions, etc.) over time or facilitate communication with the general public [21,23,31].

Composite indices developed by international organizations and institutions choose simplicity as the best methodological option. The most widely used aggregation method is the arithmetic mean. Some examples of indices that use the arithmetic mean include the United Nations Human Development Index from 1990 until 2010 when it was substituted by the geometric mean [35]; the Ease of Doing Business ranking, which studies business regulation at country level [36]; the Better Life Index developed by the Organisation for Economic Cooperation and Development [27] to visualise well-being; the Canadian Index of Wellbeing developed by the Canadian Research Advisory Group University of Waterloo [5]; or the Sustainable Development Goals (SDG) Index supported by Cambridge University Press to assess where each country stands in achieving the SDGs [30]. Other institutions have combined the arithmetic mean and principal component analysis (PCA) in their indices, such as the World Economic Forum's Global Competitiveness Index since 2008 [33] and the European Commission's European Regional Competitiveness Index from 2009 to 2019 [2]. PCA is applied to verify whether the indicators within each dimension are internally consistent and then aggregate them by an arithmetic mean in a second step. In addition to the composite indices developed by international organizations, it is worth highlighting empirical applications that use data envelopment analysis (DEA) in academia. DEA is a methodological approach to evaluate the performance of a set of observations referred to as decision making units (DMUs) that subsequently transform multiple inputs into multiple outputs. DEA methodology has the advantage that it does not depend on the method chosen to normalise the data or on the weights used for aggregation. Recent studies in this line have constructed composite indicators to assess the level of competitiveness of Costa Rican counties [19], evaluate the provision of local municipal services in Flanders [9] or to assess the performance of public hospitals in Portugal from the perspective of users and providers [1].

These indices and rankings, some of which are more influential than others, are taken as a reference to make comparisons between

* Corresponding author.
E-mail address: sancheza@ugr.es (A. Sánchez).

https://doi.org/10.1016/j.seps.2022.101339
Received 4 October 2021; Received in revised form 11 May 2022; Accepted 11 May 2022

Chapter 8

Documentation

UNIVERSITAT JAUME I
Escola de Doctorat · ED

(Nombre) María Jose Ruiz Martos....................., como coautor/ coautora doy mi **autorización** a (Nombre del doctorando/doctoranda) E.Jiménez Fernández...para la presentación de las siguientes publicaciones como parte de su tesis doctoral.

Relación de publicaciones:

Review of some statistical Methods for constructing composite indicators

Asimismo, **renuncio** a poder utilizar estas publicaciones como parte de otra tesis doctoral.

Y para que conste firmo el presente documento,

RUIZ MARTOS MARIA JOSE - 25672219X

Firmado digitalmente por RUIZ MARTOS MARIA JOSE - 25672219X
Nombre de reconocimiento (DN): c=ES, serialNumber=IDCES-25672219X, givenName=MARIA JOSE, sn=RUIZ MARTOS, cn=RUIZ MARTOS MARIA JOSE - 25672219X
Fecha: 2022.07.06 10:22:41 +02'00'

Lugar, fecha y firma

Todo ello, atendiendo al artículo 28 del Reglamento de los estudios de doctorado de la Universitat Jaume I de Castelló, regulados por el RD 99/2011, en la Universitat Jaume I (Aprobado en la sesión nº 8/2020 del Consejo de Gobierno de 02 /10/2020):

"(...)

4. En el caso de publicaciones conjuntas, todas las personas coautoras deberán manifestar explícitamente su autorización para que la doctoranda o doctorando presente el trabajo como parte de su tesis y la renuncia expresa a presentar este mismo trabajo como parte de otra tesis doctoral. Esta autorización se adjuntará como documentación en el momento del inicio de evaluación de la tesis.

UNIVERSITAT JAUME I
Escola de Doctorat · ED

(Nombre) Mario Ortega Pérez............................, como coautor/ coautora doy mi **autorización** a (Nombre del doctorando/doctoranda) E.Jiménez Fernández...para la presentación de las siguientes publicaciones como parte de su tesis doctoral.

Relación de publicaciones:

Dealing with weighted scheme in composite indicators: An unsupervised distance-machine learning proposal for quantitative data

Asimismo, **renuncio** a poder utilizar estas publicaciones como parte de otra tesis doctoral.

Y para que conste firmo el presente documento,

Lugar, fecha y firma
Valencia, 6 de Julio de 2022

Todo ello, atendiendo al artículo 28 del Reglamento de los estudios de doctorado de la Universitat Jaume I de Castelló, regulados por el RD 99/2011, en la Universitat Jaume I (Aprobado en la sesión nº 8/2020 del Consejo de Gobierno de 02 /10/2020):

"(...)

4. En el caso de publicaciones conjuntas, todas las personas coautoras deberán manifestar explícitamente su autorización para que la doctoranda o doctorando presente el trabajo como parte de su tesis y la renuncia expresa a presentar este mismo trabajo como parte de otra tesis doctoral. Esta autorización se adjuntará como documentación en el momento del inicio de evaluación de la tesis.

UNIVERSITAT JAUME I
Escola de Doctorat · ED

(Nombre).Angeles Sánchez Domínguez.........., como coautor/ coautora doy mi **autorización** a (Nombre del doctorando/doctoranda) E.Jiménez Fernández...para la presentación de las siguientes publicaciones como parte de su tesis doctoral.

Relación de publicaciones:

Dealing with weighted scheme in composite indicators: An unsupervised distance-machine learning proposal for quantitative data
EU Cohesion Policy: socio-economic vulnerability of the regions and the COVID-19 shock

Asimismo, **renuncio** a poder utilizar estas publicaciones como parte de otra tesis doctoral.

Y para que conste firmo el presente documento,

Lugar, fecha y firma

Todo ello, atendiendo al artículo 28 del Reglamento de los estudios de doctorado de la Universitat Jaume I de Castelló, regulados por el RD 99/2011, en la Universitat Jaume I (Aprobado en la sesión nº 8/2020 del Consejo de Gobierno de 02 /10/2020):

"(...)

4. En el caso de publicaciones conjuntas, todas las personas coautoras deberán manifestar explícitamente su autorización para que la doctoranda o doctorando presente el trabajo como parte de su tesis y la renuncia expresa a presentar este mismo trabajo como parte de otra tesis doctoral. Esta autorización se adjuntará como documentación en el momento del inicio de evaluación de la tesis.

Firma (1): MARÍA ÁNGELES SÁNCHEZ DOMÍNGUEZ
En calidad de: Personal Docente e Investigador UGR