EXTENDED ABSTRACT

Title: REVIEW OF SOME STATISTICAL METHODS FOR CONSTRUCTING COMPOSITE INDICATORS

Authors and e-mail of all:
María J. Ruiz-Martos (ruizmartosmaria@ugr.es)*
Eduardo Jiménez-Fernández (edjimfer@ugr.es)**
Ángeles Sánchez (sancheza@ugr.es)***

Department: *Economía Internacional y de España
****Teoría e Historia Económica**
*****Economía Aplicada**

University: Universidad de Granada

Subject area: S03 *Challenges in the construction of composite indicators for monitoring public policies*

Abstract: *(minimum1500 words)*

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 favors a multi-indicator approach. Composite indicators, which synthesize the information conveyed by a wide range of indicators, constitute a popular alternative. The most well-known example of composite indicators of human well-being is the Human Development Index (HDI) by the United Nations Development Programme (UNDP 1990, 2001, 2010).

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). As noted by Maggino (2017), measuring in social sciences requires a robust conceptual definition of the target, a consistent collection of observations and a subsequent analysis of the relationship between observations and defined concepts. The relationship between target of measurement and indicators determines the model of measurement and conditions the construction process of the composite indicator, especially the aggregation method (Maggino, 2017, p.97).

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 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).

In particular, 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, which is the independent variable, so that changes in the latent variable trigger changes in the indicators). Hence, indicators should be highly correlated and the approach should reduce dimensionality by a factor or scaling model such as factor analysis or principal components analysis (Maggino, 2017; Mazziota and Pareto, 2019). A typical example of a reflective model is the measurement of intelligence through a questionnaire (Mazziota and Pareto, 2019). Conversely, in a formative model, the latent variable depends on the indicators: changes in the latent variable do not necessarily imply changes in all the indicators (Mazziota and Pareto, 2018). In this case, indicators should not be correlated (those correlated may be redundant) and the latent variable is estimated by taking a weighted average of the indicators (Mazziota and Pareto, 2019). Data envelopment analysis (DEA), Distance *P*2 (DP2) and Mazziota Pareto Index (MPI) are examples of formative models (Jiménez-Fernández and Ruiz-Martos, 2020), as well as Distance-Learning (DL2) (Jiménez-Fernández et al., *Socio-Economic Planning Sciences*, https://doi.org/10.1016/j.seps.2022.101339). An accepted example of a formative model is the measurement of human well-being (among others, Mazziota and Pareto, 2019).

There are different aggregation approaches for constructing composite indicators. We can distinguish between compensatory and non-compensatory methods. This refers to the possibility that low values in a single indicator may or may not be compensated by high values in another indicator. The appropriateness of the (degree of) compensability of the aggregation technique depends on the conceptual framework (Jiménez-Fernández and Ruiz-Martos, 2020). 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 ELECTRE and PROMETHEE methods. The downside of non-compensatory approaches is their computational complexity, which minimizes their popularity (Greco et al., 2019).

We review, first, the methodological steps in the construction of a composite indicator. Secondly, we discuss some popular aggregation methods to construct composite indicators of human well-being, which are characterized by eliciting weights based on statistical methods (*data-driven techniques*, Decanq and Lugo, 2013, p. 19 in Greco et al., 2019): DEA, MPI, PCA, DP2 and DL2. A more detailed discussion is devoted to the DP2 and the recent DL2 method that improves the former by eliminating its crucial linear dependence weakness (Jiménez-Fernández et al., *Socio-Economic Planning Sciences*, https://doi.org/10.1016/j.seps.2022.101339). The DP2 main methodological contribution was the introduction of a metric in the construction of composite indicators. Within the construction of composite indicators, a metric is the natural way to establish the proximity or distance between countries or regions and therefore perform benchmarking in a rigorous and reliable way to guide decision making on public policies ((Jiménez-Fernández et al., *Socio-Economic Planning Sciences*, https://doi.org/10.1016/j.seps.2022.101339). Benchmarking, broadly defined as the capability to interpret results according to a specific frame (Maggino, 2017), is, therefore, one of the required characteristics in a composite indicator. The DL2 method uses quantitative data and a partially compensatory aggregation method based on the mathematical concept of distance or metric, which addresses DP2 weaknesses by using machine learning (ML) techniques. More specifically, the DL2 composite indicator is the outcome of a weighted $\ell 2$ metric, where the weights are computed using unsupervised ML algorithms. DL2 makes several notable contributions. Firstly, it measures 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 desirable mathematical properties.

We focus on these statistical 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, their approaches to the computation of weights are intrinsically different, which results in severely dissimilar measures and makes each one of them appropriate for a specific measurement exercise. We review the desired properties of an aggregation method and the properties verified by the five methodologies. Finally, we compare these methods with respect to their weighting schemes; and perform robustness tests by studying the consequences of eliminating observations and adding noise (introducing an indicator which is a lineal combination of the other indicators).

Main conclusion is that the selection among these aggregation methods requires a refinement of the conceptual framework that specifically defines the ultimate purpose of the measurement exercise. That is, it does not suffice to state the targeted multidimensional concept, e.g., human well-being. It is necessary to establish how exactly we aim to measure human well-being. If the research goal is to produce a ranking of observations (countries, regions, etc.) regarding, e.g., human well-being then PCA, DP2 and DL2 should be applied, with an increasing preference order (see Mazziota and Pareto, 2019; Jiménez-Fernández and Ruiz-Martos, 2020; and, Jiménez-Fernández et al., *Socio-Economic Planning Sciences*, https://doi.org/10.1016/j.seps.2022.101339). If the research goal is, however, to determine which dimension/s (or individual indicator/s) is/are more efficient to maximize human well-being 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.

References:

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.

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

GRECO, S., ISHIZABA, A., TASIOUR, 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.

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.

JIMÉNEZ-FERNÁNDEZ, E., SÁNCHEZ, A., ORTEGA-PÉREZ, M. *Socio-Economic Planning Sciences*, https://doi.org/10.1016/j.seps.2022.101339.

MAGGINO, F. (2017). *Complexity in Society: From Indicators Construction to their Synthesis*. Cham, Switzerland: Springer International Publishing.

MAZZIOTTA, M., & PARETO, A. (2018). *Measuring Well-Being Over Time: The Adjusted Mazziotta–Pareto Index Versus Other Non-compensatory Indices*. Social Indicators Research, 136 (3), 967-976.

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

OECD & JOINT RESEARCH CENTRE (2008). Handbook on constructing composite indicators: methodology and user guide. Paris: OECD.

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.

SANCHEZ, A., & RUIZ-MARTOS, M. (2018). Europe 2020 Strategy and Citizens' Life Satisfaction. Journal of Happiness Studies.

SOMARRIBA, N., & PENA, B. (2009). Synthetic indicators of quality of life in Europe. Social Indicators Research, 94(1), 115-133.

UNDP. (1990). *Human development report 1990: Concept and measurement of human development.* Oxford: Oxford Universiry Press.

UNDP. (2001). *Human development report 2001*. New York: Oxford University Press.

UNDP. (2010). *Human development report 2010.* New York: Palgrave MacMillan.

YANG, F.-C., KAO, R.-H., CHEN, Y.-T., HO, Y.-F., CHO, C.-C., & HUANG, S.-W. (2018). "A Common Weight Approach to Construct Composite Indicators: The Evaluation of Fourteen Emerging Markets". Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement, Springer, vol. 137(2), pages 463-479, June.

Keywords: (máximum 6 words) Indicadores compuestos, Ponderación, Agregación, DEA-BoD, PCA, Distancia P2, Índice de Pareto de Mazziotta.
JEL codes: I30, C13, C51, C55