Comparing Composite Indicators to measure Quality of Life: the Italian “Sole 24 Ore” case

Un confronto tra indicatori compositi per misurare la qualità della vita: il caso italiano de “Il Sole 24 Ore”

Gianna Agrò, Mariantonietta Ruggieri and Erasmo Vassallo

Abstract The measure of Quality of Life is still a topic widely discussed in the literature. In Italy, the newspaper "Il Sole 24 Ore" publishes a famous ranking that highlights a strong disparity between Northern and Southern Italy areas. In this paper, some methods are compared in order to show how different types of normalization and aggregation can influence the results.
Riassunto *La misura della qualità della vita continua ad essere un argomento ampiamente discusso in letteratura. In Italia, il quotidiano "Il Sole 24 Ore" pubblica una nota graduatoria che evidenzia una forte disparità tra le aree del Nord e del Sud Italia. In questo lavoro, si confrontano alcune procedure per mostrare come diversi tipi di normalizzazione e di aggregazione possano influenzano i risultati.*

Key words: Quality of Life, Composite Indicators, BoD approach, standardization

1 Introduction

The newspaper "Il Sole 24 Ore" has been publishing for 25 years a ranking of Quality of Life (QoL) for Italian provinces according to the NUTS3 classification. This survey produces every year a great media resonance, recalling a long-debated issue about the so-called "Questione Meridionale"; economic and social territorial disparities are highlighted by a strong polarization among the provinces of North Italy, with higher levels of QoL than the provinces of the South (Felice, 2013;

[1] Gianna Agrò, DSEAS, University of Palermo; email: gianna.agro@unipa.it
Mariantonietta Ruggieri, DSEAS, University of Palermo; email: mariantonietta.ruggieri@unipa.it
Erasmo Vassallo, DSEAS, University of Palermo; email: erasmo.vassallo@unipa.it

IlSole24Ore, 2017). Actually, this result is strongly influenced by the choice of the basic dimensions, the kind of data and indicators, the standardization and aggregation techniques to obtain a Composite Indicator (CI). There is no doubt that the economic and the social context presents a greater level of discomfort in the South of Italy (ISTAT, 2017a), but this can not immediately be related to lower levels of QoL. It is very difficult to provide a univocal definition of QoL, as this concept invests personal and/or community aspects, depending on subjective and/or objective well-being and happiness (Cummins, 1998; Nussbaum and Sen 2003). Certainly, the concept of QoL is much more than "standard of living" one, that is basically connected to income levels (ISTAT, 2017b). The absence of a clear and univocal definition makes the concept more difficult to translate into data, indicators and measures that are free of criticism. Nevertheless, the aim of this paper is not to discuss about the different definitions of QoL used in the theoretical and applied literature, but to focus on the statistical techniques used to standardize and synthesize data, influencing the final ranking of the Italian provinces. Although we highlight many perplexities on the redundancy of some data and their ability to correctly represent some dimensions, here we accept the choice of "Il Sole 24 Ore" about pillars and indicators used in its QoL dossier, well-known by journalists and politicians. In particular, we compare four procedures, described in the Section 2, including the BoD-DEA method, which provides endogenous weights emphasizing the strength of each province (Giambona and Vassallo, 2014). The results and comments are reported in Section 3, while in Section 4 we draw some conclusions.

2 "Il Sole 24 Ore" quality of life and BOD approach

The 2017 "Il Sole 24 Ore" dossier uses 42 indicators divided into 6 dimensions (income and consumption, labor and innovation, environment and services, demography and society, justice and safety, culture and leisure), each consisting in 7 indicators to measure QoL among the 110 Italian provinces, according to the NUTS3 classification of the territory. In this paper, we do not discuss the choice of indicators and dimensions of the analysis that, moreover, is not constant over time, but the procedure used to aggregate the data. In particular, the dossier proposes to standardize each of the 42 indicators for the *i*-th province with $std.value_i=(indicator\ value_i\ /\ max(indicator\ values))\cdot 1000$ if the indicator polarity is positive (higher values correspond to a better condition) and $std.value_i=(min(indicator\ values)/indicator\ value_i\ )\cdot 1000$ if the polarity is negative (lower values correspond to a better condition). In this way, the best province (benchmark) always assumes a maximum score of 1000 and the other provinces assume decreasing values up to a minimum; then a simple arithmetic mean among indicators/dimensions provides an overall score for each province. This method has the advantage of being very simple and intuitive, but it ignores some statistical features of data, for example the different variability exhibited by the different indicators. Besides, the simple arithmetic mean is a fully compensatory aggregation procedure; this implies that the indicators and dimensions have the same importance and they can also be considered perfectly complementary. A better way to standardize the indicators is to use the min-max procedure, i.e. *std.value*$_i$ = *(indicator value*$_i$ *- min(indicator values))/(max(indicator values)-min(indicator values))* if the polarity is positive and *std.value*$_i$ = *(max(indicator values) - indicator value*$_i$*)/(max(indicator values)-min(indicator values))* if the indicator polarity is negative; in this way, the standardized values always range between 0 and 1 for all indicators (OECD, 2008). Obviously, the compensatory nature of the mean is not resolved.

Among other proposals in literature (Munda and Nardo, 2009), the AMPI index is widely used by some ISTAT reports; it presents a simple solution to the full compensation of the arithmetic mean through the use of a penalty based on a function of variability (Mazziotta and Pareto, 2017). So, the indicators are standardized via min-max rescaled between 70 and 130 and the corresponding mean is reduced if the province has different values (higher variability) among the dimensions, i.e. $AMPI_i = M_i - S_i \cdot cv_i$, where M_i, S_i and cv_i are the mean, the standard deviation and the coefficient of variation, respectively, of the indicators for the i-th province.

Recent developments aim at removing any kind of subjectivity in the choice of right weights for aggregating indicators and/or dimensions. In this framework, the Benefit of Doubt (BoD) based methods have received many consents. These methods exploit the Data Envelopment Analysis (DEA), a frontier technique that has been usually used for measuring the efficiency in production. Several variants of BoD models have been proposed in literature (Rogge and Van Nijverseel, 2018). In this work, we use an appropriate geometric mean of the indicators, in which the weights are endogenously defined, depending on the characteristics of the data according to the principle of BoD. A Composite Indicator $CI_i = \prod_{j=1}^{n} I_{ij}^{w_j}$ is defined, where I_{ij} is the *j*-th indicator of QoL ($j = 1,...,n$) for the *i*-th province ($i = 1,...,m$) with weight w_j determined endogenously by an algorithm based on a multiplicative optimization model, that solves the following problem: $SI_i = \max_w \prod_{j=1}^{n} I_{ij}^{w_j}$ with constraints $\prod_{j=1}^{n} I_{ij}^{w_j} \leq e$ and $w_j \geq 0$, where e is the Napier's constant (Zhou et al., 2007; Giambona and Vassallo, 2016). In this way, the CI is obtained by multiplying the basic indicators of QoL with weights calculated in the best possible conditions, i.e. increasing as much as possible the composite score for a given province. In short, a low value of the CI for the *i*-th province is due to low values of the indicators that compose it and not to specific weights, calculated to obtain the best, i.e. the maximum possible, result for the *i*-th unit compared to the benchmark province. At the end, we obtain scores between 1 and 2.71 (the Napier's constant), attributable to the most intuitive interval between 0 and 1 by applying the antilogarithm. However, the optimization problem could determine zero weight to some indicators and attribute too much weight to other indicators, and this is not desirable if all the dimensions are relevant. So, we add specific constraints on the weights; in particular, we add proportion constraints to the model: $\left(\prod_{j=1}^{n} I_{ij}^{w_j}\right)^{L} \leq I_{ij}^{w_j} \leq \left(\prod_{j=1}^{n} I_{ij}^{w_j}\right)^{U}$, where U and L ranging between 0 and 1 to represent the lower and the upper bound (in percentage terms) for the contribution of the j-th indicator. In fact, without constraints on the weights, the model could ignore the contribution of the under-performing indicators or dimensions to maximize the best solution, and this is not admissible. In our case L=10% (U is defined accordingly); we note that bounds slightly lower or slightly higher lead to similar results, so an intermediate value has been chosen among the possible alternatives, taking into account that L (to guarantee sufficient flexibility to the method) must be relatively low.

3 Results and comments

This Section presents the results obtained by means of the methods described in Section 3, that is: 1) the original "IlSole24Ore" with min or max standardization on 1000 and use of arithmetic mean (CI.SOLE); 2) the "range" with min-max standardization rescaled between 1 and 10 and arithmetic average (CI.RANGE); 3) the "ampi" with min-max standardization between 70 and 130 and use of a penalized arithmetic mean (CI.AMPI); 4) the "BoD" with min-max standardization, rescaled for calculation purposes between 2-10, and geometric mean with weights endogenously defined (CI.BOD).
In Table 1 the correlation matrix among the considered CIs is reported, showing a very high correlation between the series.

Table 1: Correlation matrix among CIs

	CI.SOLE	**CI.RANGE**	**CI.AMPI**	**CI.BOD**
CI.SOLE	1.00	0.94	0.93	0.82
CI.RANGE	0.94	1.00	0.99	0.93
CI.AMPI	0.93	0.99	1.00	0.94
CI.BOD	0.82	0.93	0.94	1.00

Figure 1 reports the kernel density estimates of the four CIs, where the BoD kernel density estimates uses the Silverman reflection method to avoid bias and inconsistency near the boundaries (Silverman, 1986; Scott, 1992). We note some “twin peaks” distributions, substantially in correspondence with the provinces of the southern and northern Italy. A similar polarization can be seen in BoD, although this distribution suggests a more detailed interpretative framework. In fact, the scores of the different methods are consistent each other, but the rankings are sometimes quite different. In particular, the BoD ranking shows a less intense polarization and the distance North-South appears less severe. In this regard, it is interesting to note the strong change of position of some provinces compared to the original classification of the "IlSole24Ore"; for example, some northern provinces lose many positions:

Monza e Brianza loses 55 positions (from 29 to 84), Verbano-Cusio-Ossola loses 49 positions (from 7 to 56), Sondrio loses 41 positions (from 3 to 44), and so on. On the contrary, other provinces acquire positions, such as Brescia (from 46 to 14), Padova (from 42 to 12), Venezia (from 43 to 15), etc.

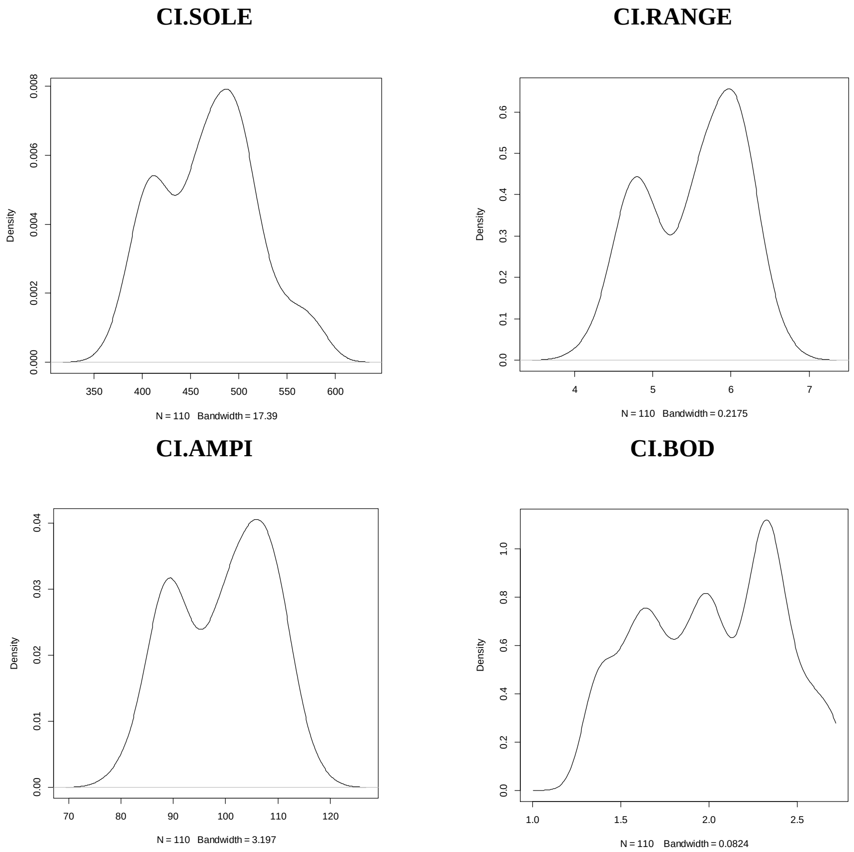

Figure 1: Kernel density estimates of the considered CIs

4 Conclusions

In this paper four different procedures, able to supply a unique composite indicator measuring the QoL, are applied to data reported on the well-known "IlSole24Ore" QoL dossier. We do not discuss the choice of indicators and dimensions, even if there are many critical aspects, but we focus on the possible consequences in applying different aggregative procedures. The original “IlSole24Ore” procedure uses a simple standardization and a fully compensatory arithmetic mean, similarly the "range" procedure that, instead, uses a better min-max standardization. The "ampi" technique uses again an arithmetic mean, but penalizes the provinces with greater variability among indicators/dimensions. Finally, the BoD method applies different weights to dimensions and provinces according to the specific characteristics of the data; besides, it uses a geometric mean to avoid the problem of a full compensation among indicators or dimensions. The BoD method is more specific and advanced from a technical point of view and addresses some critical aspects compared to simpler methods; nevertheless, it is not possible to assert that an endogenous choice of the weights is better compared to an exogenous one or to a choice of identical weights for units and dimensions. Certainly, the techniques of standardization and aggregation strongly influence the rankings, and this generates many doubts about the utility of these classifications. Finally, this work highlights that all the rankings maintain a strong North-South distinction, but this depends on strongly heterogeneous indicators between northern and southern provinces: with less heterogeneous variables, the ranking differences would be even more evident.

References

1. Cummins R.A. (1998). *Quality of Life Definition and Terminology: A Discussion*. Document from the International Society for Quality-of-Life Studies. International Society for Quality-of-Life Studies. *(mimeo).*
2. Felice E. (2013). *Perché il Sud è rimasto indietro*, Il Mulino, Bologna.
3. Giambona F. and Vassallo E. (2014). Composite Indicator of Social Inclusion for European countries. *Social Indicators Research*, 116, 1, 269-293.
4. Giambona F. and Vassallo E. (2016). *Composite Indicator of Social Inclusion for the EU Countries*. In: Alleva G. and Giommi A. (eds.) "Topics in Theoretical and Applied Statistics", Springer, Berlin.
5. IlSole24Ore (2017). *Dossier Qualità di vita 2017*, Il Sole 24 Ore, Milano.
6. ISTAT (2017a). *Rapporto Annuale*, Istat, Roma.
7. ISTAT (2017b). *Rapporto BES*, Istat, Roma.
8. Mazziotta M. and Pareto A. (2017). Measuring Well-Being Over Time: The Adjusted Mazziotta–Pareto Index Versus Other Non-compensatory Indices. *Social Indicators Research*, first online, doi: 10.1007/s11205-017-1577-5.
9. Munda G. and Nardo M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Applied Economics*, 41, 12, 1513-1523.
10. Nussbaum M. and Sen A. (2003). *The Quality of Life*, Oxford Press, Oxford.
11. OECD (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Oecd, Paris.
12. Rogge N. and Van Nijverseel I. (2018). Quality of Life in the European Union: A Multidimensional Analysis. *Social Indicators Research*, first online, doi: 10.1007/s11205-018-1854-y.
13. Scott D. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley, New York.
14. Silverman B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman $ Hall, London.
15. Zhou P., Ang B.W. and Poh K.L. (2007). A mathematical programming approach to constructing composite indicators. *Ecological Economics*, 62, 291–297.