CrossMark

Use and Misuse of PCA for Measuring Well-Being

Matteo Mazziotta[1] · Adriano Pareto[2]

Abstract The measurement of well-being of people is very difficult because it is characterized by a multiplicity of aspects or dimensions. Principal Components Analysis (PCA) is probably the most popular multivariate statistical technique for reducing data with many dimensions and, often, well-being indicators are reduced to a single index of well-being by using PCA. However, PCA is implicitly based on a reflective measurement model that is not suitable for all types of indicators. In this paper, we discuss the use and misuse of PCA for measuring well-being, and we show some applications to real data.

Keywords Data reduction · Composite indicator · Measurement model · Well-being

1 Introduction

International interest in well-being research has significantly increased in recent years due to the boost of the "Beyond GDP" initiative and the Stiglitz et al. report (2009). Policy makers and researchers have become more and more aware of the fact that well-being is relevant for countries at all levels of development, and that the GDP (Gross Domestic Product) per capita cannot alone explain this concept (Boarini et al. 2014; OECD 2015). In fact, human well-being is determined by a wide range of factors that are not captured by GDP, such as health, education, environmental quality, meaningful work, leisure time, and so on (Sen 1985). Furthermore, the GDP is positively correlated with some of these factors (e.g. health and education), while in other cases the relationship is weak, if not negative. For example, some indicators of environmental

✉ Adriano Pareto
  pareto@istat.it

  Matteo Mazziotta
  mazziott@istat.it

1 Italian National Institute of Statistics, Viale Oceano Pacifico 171, 00144 Rome, Italy

2 Italian National Institute of Statistics, Via A. Depretis 77, 00184 Rome, Italy

performance (e.g. carbon dioxide [CO_2] emissions) tend to worsen with increased GDP (Nahman et al. 2016).

In well-being research, we often distinguish between objective and subjective well-being. Objective well-being concerns observable factors such as richness, health, and tangible goods. Subjective well-being concerns psychological experiences (Michalos 2014). Hence, the objective approach looks at 'harder' data, such as income per capita or gross enrolment ratios, while the subjective approach considers 'softer' matters, such as an individual's satisfaction with income and his perceived adequacy of educational opportunities (Bleys 2012). As a result, objective well-being can be assessed in terms of indicators of *outcome*; whereas subjective well-being is often measured as 'happiness' or 'life satisfaction' by response scales in questionnaires surveys (Van Beuningen et al. 2014).

Well-being indicators are often analysed by multivariate statistical technique, such as Principal Components Analysis (PCA), in order to summarize the data. The use of PCA is probably due to its computational simplicity (Krishnakumar and Nagar 2008). Ram (1982) applies PCA for constructing composite indices of economic development that capture per capita income, basic needs fulfilment, and other possible indicators of well-being. Slottje (1991) follows the same approach by selecting 20 attributes for 126 countries in computing a composite index of quality of life. Many other composite indices of welfare derived from PCA can be found in literature (see, e.g., Biswas and Caliendo 2002; Lai 2003; McGillivray 2005; Wong 2012; Haq and Zia 2013; Ferrara and Nisticò 2014). Moreover, PCA can be used as extraction technique when performing Factor Analysis (FA).

However, a fundamental distinction must be made between reducing dimensionality and constructing composite indicators.

Reducing dimensionality is a purely mathematical operation that consists in summarizing a set of individual indicators, so that most of the information in the data is preserved. Many techniques have been developed for this purpose, but PCA is one of the oldest and most widely used (Hotelling 1933). Its idea is simple: reduce the dimensionality of a dataset, while preserving as much 'variability' as possible. This translates into finding new variables that are linear functions of the original ones, that successively maximize variance and that are uncorrelated with each other. Finding such new variables reduces to solving an eigenvalue/eigenvector problem, and the results depend on the dataset, rather than being pre-defined basis functions. Because the new variables are defined by the dataset at hand, and not a priori, PCA can be considered an adaptive data analysis tool (Jolliffe and Cadima 2016).

Constructing a composite index (or composite indicator) is a conceptual, as well as mathematical, operation that consists in summarizing (or aggregating as it is termed) a set of individual indicators, on the basis of a well-defined measurement model: formative or reflective (Michalos 2014). Therefore, a composite indicator is formed when individual indicators are compiled into a single index, on the basis of an underlying model of the multi-dimensional concept that is being measured (OECD 2004). Constructing a composite index is a complex task. Its steps involve several alternatives and possibilities that affect the quality and reliability of the results. The main problems, in this approach, concern the choice of theoretical framework, the selection of the more representative indicators and their treatment in order to compare and aggregate them (Salzman 2003; Mazziotta and Pareto 2017).

Obviously, a composite index can be obtained by reducing dimensionality (with an appropriate model of measurement), but not necessarily reducing dimensionality provides a composite index.

In this paper, we discuss the use of PCA for studying well-being indicators and we explain how and why it can be improperly used as a method for constructing composite indices. The paper is organized as follows. Section 2 introduces the difference between formative and reflective measurement models. Then a brief description of PCA is reported and pro and cons of this technique are discussed. In particular, it is shown that PCA rests on a reflective model, even if it is used in a formative approach. An illustrative example is also provided to show the theoretical and empirical limits of PCA when summarizing a set of well-being indicators. Section 3 reports some case studies for measuring well-being in Italy at the regional and provincial level, where PCA is used. Finally, in Sect. 4 conclusions are drawn and some suggestions for measuring well-being by using PCA are given.

2 How and When to Use PCA

According to the "Handbook on Constructing Composite Indicators. Methodology and user guide" by OECD (2008), PCA should be used to study the overall structure of the dataset, assess its suitability, and guide some methodological choices in constructing a composite indicator. In particular, PCA may help to identify groups of individual indicators or groups of units that are statistically 'similar' and to provide an interpretation of the results.

Nevertheless, PCA can also be used for constructing composite indices. For this purpose, it is essential to define the model of measurement in order to describe relationships between the phenomenon to be measured (latent variable) and its measures (individual indicators).

2.1 Formative Versus Reflective Measurement Models

As it is known, a model[1] of measurement can be conceived through two different conceptual approaches: reflective or formative (Jarvis et al. 2003; Diamantopoulos et al. 2008).

The most popular approach is the reflective model, according to which individual indicators denote effects (or manifestations) of an underlying latent variable. Therefore, causality is from the concept to the indicators and a change in the phenomenon causes variation in all its measures. In this model, the concept exists independently of awareness or interpretation by the researcher, even if it is not directly measurable (Borsboom et al. 2003).

Specifically, the latent variable R represents the common cause shared by all indicators X_i reflecting the concept, with each indicator corresponding to a linear function of the underlying variable plus a measurement error:

$$X_i = \lambda_i R + \varepsilon_i \qquad (1)$$

where X_i is the indicator i, λ_i is a coefficient (loading) capturing the effect of R on X_i and ε_i is the measurement error for the indicator i. Measurement errors are assumed to be independent and unrelated to the latent variable.

A fundamental characteristic of reflective models is that individual indicators are interchangeable (the removal of one of the indicators does not change the essential nature of the underlying concept) and correlations between indicators are explained by the measurement model (all indicators must be intercorrelated).

---

[1] For the sake of simplicity, only linear models will be considered.

Another important issue concerns the polarity of the individual indicators. The 'polarity' of a individual indicator is the sign of the relation between the indicator and the concept to be measured. For example, in the case of well-being, "Life expectancy" has positive polarity, whereas "Unemployment rate" has negative polarity. In a reflective model, individual indicators with equal polarities must be positively correlated, whereas individual indicators with opposite polarities must be negatively correlated. Otherwise, the model will produce inconsistent results (for a numerical example, see Sect. 2.4).

A typical example of reflective model is the measurement of the intelligence of a person. In that case, it is the 'intelligence level' that influences the answers to a questionnaire for measuring attitude, and not vice versa. Hence, if the intelligence of a person increased, this would be accompanied by an increase of correct answers to all questions (Simonetto 2012).

The second approach is the formative model, according to which individual indicators are causes of an underlying latent variable, rather than its effects. Therefore, causality is from the indicators to the concept and a change in the phenomenon does not necessarily imply variations in all its measures. In this model, the concept is defined by, or is a function of, the observed variables.

The specification of the formative model is:

$$R = \sum_i \lambda_i X_i + \zeta \tag{2}$$

where λ_i is a coefficient capturing the effect of X_i on R, and ζ is an error term.[2]

In this case, indicators are not interchangeable (omitting an indicator is omitting a part of the underlying concept) and correlations between indicators (r_{ij}, $i \neq j$) are not explained by the measurement model (high correlations between indicators are possible, but not generally expected). So, in a formative model, polarities and correlations are independent and individual indicators can have positive, negative or zero correlations.

It is noteworthy that, because a formative model is not based on the hypothesis that the indicators are correlated, the correlation structure of the data cannot be used to determine the latent variable. Rather, the latent variable is estimated by taking a weighted[3] average of the indicators that comprise the concept (Shwartz et al. 2015).

A typical example of formative model is the measurement of well-being of society. It depends on health, income, occupation, services, environment, etc., and not vice versa. So, if any one of these factors improved, well-being would increase (even if the other factors did not change). However, if well-being increased, this would not necessarily be accompanied by an improvement in all factors.

One of the oldest and most famous formative composite indices is the Human Development Index (HDI) by United Nations Development Programme (UNDP 1990, 2010). It is a composite measure of human development that includes three theoretical dimensions: Health, Education and Income. Any change in one or more of these components is likely to cause a change in a country's HDI score, but there is no reason to expect the components are correlated. The same goes for the Canadian Index of Well-being (CIW), a composite

---

[2] Some authors exclude the error term so that Eq. (2) reduces to a weighted linear combination of the X_i (Diamantopoulos 2006).

[3] Experts suggest that weights could be determined a priori, according to the theoretical contribution of the indicators to the concept (Howell et al. 2007). For Cadogan and Lee (2013), if there is no theory suggesting the contrary, individual indicators should have equal weightings.
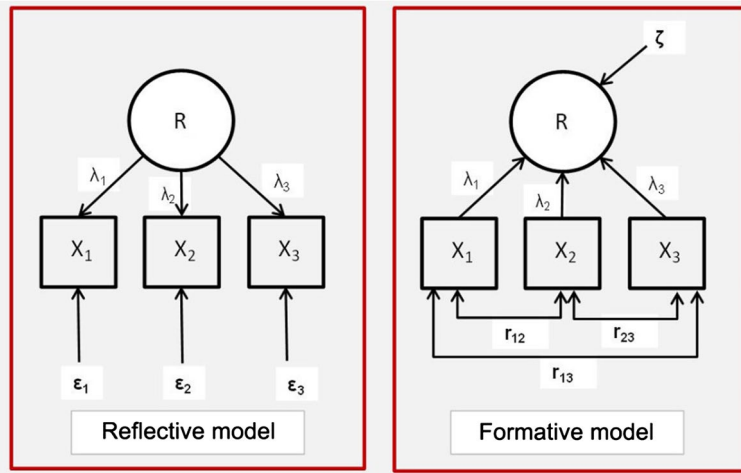
Fig. 1 Alternative measurement models

measure of well-being based on eight domains: Living Standards, Healthy Populations, Community Vitality, Democratic Engagement, Leisure and Culture, Time Use, Education, Environment (Michalos et al. 2011).

Note that (1) is a system of simple regression equations where each individual indicator is the dependent variable and the latent variable is the explanatory variable; whereas (2) represents a multiple regression equation where the latent variable is the dependent variable and the indicators are the explanatory variables.[4] Hence, the correct interpretation of the relationships between indicators and latent variable allows the procedure aimed at aggregating individual indicators to be correctly identified (Maggino 2017).

In Fig. 1, the two different approaches are graphically represented. Traditionally, the reflective model is applied in the development of scaling models for subjective measurement (e.g. attitude or satisfaction scale construction), whereas the formative model is commonly used in the construction of composite indices based on both objective and subjective indicators (Maggino and Zumbo 2012). Although the reflective view dominates the psychological and management sciences, the formative view is common in economics and sociology (Coltman et al. 2008).

2.2 Pros and Cons of PCA

PCA is a multivariate statistical technique that, starting from a large number of quantitative[5] individual indicators, allows to identify a small number of composite indices (principal components or factors) that 'explain' most of the variance observed (Dunteman 1989).

---

[4] Because the formative measurement model is based on a multiple regression, the stability of the coefficients λ_i is affected by the strength of the indicator intercorrelations. Therefore, multicollinearity must be avoided. (Diamantopoulos and Winklhofer 2001).

[5] Individual indicators must have at least an interval level of measurement. For variables measured on nominal or ordinal scale, we recommend the use of Categorical Principal Components Analysis (CATPCA). For a introduction and application of CATPCA, see Linting et al. (2007).

The first factor of PCA is often used as the 'best' composite index (Booysen [2002](); Mishra [2007](), [2008](); Krishnakumar and Nagar [2008]()). Denoting with C_{i1} the score of factor 1 (the first component extracted) for unit i, the composite index is defined as:

$$C_{i1} = \sum_{j=1}^{m} a_{j1} x_{ij}$$

where a_{j1} is the weight for indicator j, as used in creating factor 1, x_{ij} is the value of indicator j for unit i, and m is the number of individual indicators.

PCA has a number of excellent mathematical properties (Kendall and Stuart [1968]()). The most important property is that the index obtained from the first principal component accounts for the largest amount of total variance in the individual indicators. This is obtained by maximizing the sum of the squares of the coefficients of correlation between the composite index and the individual indicators. Thus, the first factor will be correlated with at least some of the individual indicators. Often, it is correlated with many.

However, the first principal component accounts for a limited part of the variance in the data, so we can lose a consistent amount of information.[6] Moreover, the PCA based index is often 'elitist' (Mishra [2008]()), with a strong tendency to represent highly intercorrelated indicators and to neglect the others, irrespective of their possible contextual importance. Consequently, many highly important but poorly intercorrelated indicators may be unrepresented by the composite index. On many occasions, it is found that some very important indicators are roughly dealt with by PCA, simply because those variables exhibited widely distributed scatter or points did not fall within a narrow band around a straight line. In addition, data may have outliers. These outliers can pull down (or up) correlation coefficients of some individual indicators with the others and then affect the index unpredictably. In such a case, the indicators favoured or disfavoured by PCA may obtain entirely unwarranted weights (Mishra [2007]()).

On the other hand, PCA is a blindly empiricist method based on the observed correlations and it ignores the polarity of the individual indicators. Therefore, if the normalized indicators[7] are not all positively intercorrelated, the PCA based index is not correct, as individual indicators are summarized without regard to the proper polarities.

Another aspect to be taken into account in constructing a composite index by PCA is that the meaning of the weights is clear from a mathematical point of view, but it makes little sense in relation to the aim of measuring well-being. So, the weights of the individual indicators often lack socio-economic interpretation (Somarriba and Pena [2009]()). This is because the factors found by PCA are 'empirical' dimensions[8] (based on the variability), and not 'theoretical' dimensions (based on a conceptual framework). Empirical dimensions and theoretical dimension often do not match (even if this would be desirable), which could makes it difficult to assign a clear meaning to the factors.

---

[6] Often, the use of the first principal component as the 'only' composite index is a bad practice that reduces the PCA potentials.

[7] Normalization aims to make individual indicators comparable, as they often have different measurement units and/or different polarities. Normalized indicators are calculated by transforming individual indicators into pure, dimensionless, numbers, with positive polarity (Mazziotta and Pareto [2017]()).

[8] Principal components can be real features of the data, or more or less convenient fictions and summaries. That they are real is a hypothesis for which PCA can provide only a very weak evidence (Shalizi [2009]()).

In addition, PCA does not allow making inter-spatial (for different groups of units) or inter-temporal (for different times) comparisons, as the amount of variance accounted for, and the weights computed by PCA change for each data matrix, and then the results of different analysis are not easily comparable. This can be a big problem, if the composite index must be calculated and assessed over time. The question could be addressed by using STATIS[9] (Structuration des tableaux à trois indices de la statistique) or similar techniques, but the composite index would be recalculated each time new data is available. Note also that PCA cannot be applied to matrices containing values of a set of indicators for different months or years, because correlations must be computed on observations that are independent (e.g. individuals or geographical areas).

Last, but not least, PCA can be little robust and very sensitive to the inclusion or exclusion of an individual indicator.[10] The smaller the correlation of the indicator with the others, the less the robustness of the results.

2.3 Is PCA Formative or Reflective?

To answer to this question it is important to distinguish between PCA and FA,[11] since they are sometimes considered more or less interchangeable (Krishnakumar and Nagar 2008; Osborne 2014).

PCA is a pure data reduction technique that aggregates the observed variables (indicators) in order to reproduce the most amount of variance with fewer variables (principal components or factors). PCA works without an explicit hypothesis on the latent structure of the variables, so that the observed variables are themselves of interest. This makes PCA similar to multiple regression in some ways, in that it seeks to create optimized weighted linear combinations of variables (Osborne 2014).

FA is an explanatory model in which the observed variables (indicators) are assumed to be (linear) functions of a certain (fewer) number of unobserved variables (latent factors). FA hypothesizes an underlying latent structure of the variables and estimates latent factors influencing observed variables.

On the basis of these features, PCA is often views as formative, whereas FA is a reflective measurement model (Edwards and Bagozzi 2000; Zumbo 2007; Markus and Borsboom 2013). However, the question whether PCA is formative or reflective is not trivial. Indeed, although the definition of principal component as weighted sum of individual indicators suggests a formative model, some important issues are involved.

In particular:

1. In a PCA based index (e.g. the first factor), the weights depend on the correlations among indicators. But correlations among individual indicators are not relevant in a formative model and cannot be explained by it. Indeed, in a formative model, the indicators do not necessarily share the same theme and hence have no a preconceived pattern of intercorrelation (Coltman et al. 2008).

---

[9] STATIS is an exploratory technique of multivariate data analysis for handling three-way matrices, where the same units have measures on a set of indicators under a number of conditions (Lavit et al. 1994).

[10] Several robust PCA methods have been introduced in the literature (Filzmoser 1999; Hubert et al. 2005), but they make the analysis resistant to outlying observations.

[11] There are two types of FA: exploratory and confirmatory. In this paper, we consider exploratory factor analysis (Fabrigar and Wegener 2011).

2. Individual indicators aggregated by a PCA based index (e.g. the first factor) are—by construction—highly correlated. But in a multiple regression, such as Eq. 2, individual indicators should have little or no correlation among themselves in order to avoid multicollinearity. Indeed, an excessive collinearity among indicators makes it difficult to separate the distinct influence of the individual indicators on the latent variable (Diamantopoulos and Winklhofer 2001).
3. Under certain conditions, the principal components are equivalent to the factor scores obtained by FA and then they can be considered estimators of latent factors (Krishnakumar and Nagar 2008). But FA is a reflective measurement model, so PCA cannot be considered really formative.

In the light of the above, a composite index based on PCA looks more suited for a reflective approach than a formative one.

In fact, PCA is commonly used for the evaluation of reflective measurement models (Götz et al. 2010) and it is considered an appropriate method for examining the indicators' underlying factor structure in order to check the content validity (Bohrnstedt 1970; Vinzi et al. 2003).

2.4 A Numerical Example

In this section, we consider a numerical example where a formative composite index is requested. A simple arithmetic mean and the first factor of PCA are compared as composite indices, but the PCA based index fails because the conditions required for a reflective model (individual indicators with opposite polarities must be negatively correlated) are not met.

Let us imagine that we want to construct a composite index of well-being in the work dimension, for several countries or regions, based on the following individual indicators:

X_1 = Employment rate;
X_2 = Incidence rate of occupational injuries.

Indicator X_1 has positive polarity (it is positively correlated with well-being), whereas indicator X_2 has negative polarity (it is negatively correlated with well-being).

Suppose also that X_1 and X_2 are positively correlated, i.e. $r(X_1, X_2) > 0$, so that high employment rates tend to be associated with higher rates of occupational injuries.

In a formative approach, such as Eq. (2), we can create a composite index by arithmetic mean. However, the first factor of PCA could be the best solution, since it accounts for as much variance as possible.

In Table 1 is reported an example where five countries are considered. The table also provides the normalized indicators[12] Z_1 and Z_2, the ranks R_1 and R_2, the arithmetic mean of the normalized values M_1, and the first factor[13] scores PC_1. Note that $r(X_1, X_2) = 0.45$, whereas $r(Z_1, Z_2) = -0.45$, because the polarity of X_2 has been inverted in order to construct the composite index.[14]

---

[12] Individual indicators were normalized as *z*-scores. The signs were reversed if the polarity is negative.

[13] The first factor of PCA accounts for 72.4% of the variance in the data.

[14] Note that, for constructing a composite index, all the normalized indicators must have positive polarity, so that an increase in each of them corresponds to an increase in the composite index (Mazziotta and Pareto 2013).

Table 1 Comparing arithmetic mean and first component score as composite indices. *Source*: Mazziotta and Pareto (2016b)

Country	Original values		Normalized values		Ranks		M_1		PC_1	
	X_1	X_2	Z_1	Z_2	R_1	R_2	Value	Rank	Value	Rank
1	80.0	1.9	1.58	−0.71	1	4	0.44	2	1.20	1
2	50.0	2.0	0.00	−1.41	3	5	−0.71	5	0.74	2
3	50.0	1.8	0.00	0.00	3	3	0.00	3	0.00	3
4	50.0	1.6	0.00	1.41	3	1	0.71	1	−0.74	4
5	20.0	1.7	−1.58	0.71	5	2	−0.44	4	−1.20	5
Mean	50.0	1.8	0.00	0.00						
SD	19.0	0.1	1.00	1.00						
Average rank correlation with normalized indicators							0.520		0.050	

Table 2 Influence analysis. Absolute difference of rank when excluding an indicator

Country	M_1				PC_1			
	X_1	X_2	Mean	SD	X_1	X_2	Mean	SD
1	2.0	1.0	1.5	0.5	3.0	0.0	1.5	1.5
2	0.0	2.0	1.0	1.0	3.0	1.0	2.0	1.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	2.0	1.0	1.0	3.0	1.0	2.0	1.0
5	2.0	1.0	1.5	0.5	3.0	0.0	1.5	1.5
Mean	0.80	1.20	1.00	0.20	2.40	0.40	1.40	1.00

As we can see, units 2, 3, and 4 have the same employment rate ($X_1 = 50.0$) and decreasing values of the rate of occupational injuries. Nevertheless, unit 2 ranks 5th according to M_1 and ranks 2nd according to PC_1, whereas unit 4 ranks 1st according to M_1 and ranks 4th according to PC_1. So, the average Spearman rank correlation coefficient between the composite index and the individual indicators is 0.52 for M_1 and 0.05 for PC_1. This is due to the fact that PCA ignores the polarity of the individual indicators and normalized indicators are not positively correlated. Therefore, the use of PC_1 for aggregating X_1 and X_2 results in an inconsistent composite index and an unrealistic ranking of units, because PC_1 is concordant with both X_1 and X_2, whereas it should be concordant with X_1 and discordant with X_2 (as is the case for M_1), according to the polarities.

Finally, an Influence Analysis[15] is performed to assess the robustness of the composite indices when excluding an individual indicator. In particular, for each method (M_1, PC_1) and for each country (1, 2, …, 5), the composite index is computed, by excluding each time a different indicator (X_1, X_2). The absolute differences of rank (shifts) between the new rank and the original one are reported in Table 2. The table

---

[15] Influence Analysis is a particular case of Uncertainty Analysis that aims to empirically quantify the 'weight' of each individual indicator in the calculation of the composite index (Mazziotta and Pareto 2017).

Table 3 Well-being indicators for Italian regions. *Source*: Istat (2015a)

Label	Well-being indicator	Year
HEA	Composite index of health	2013
EDU	Composite index of education and training	2014
QOW	Composite index of quality of work	2014
EMP	Normalized employment rate	2014
INC	Composite index of income and inequality	2014
HAR	Composite index of economic hardship	2014
REL	Composite index of social relationships	2014
SAF	Composite index of safety	2014
HOM	Normalized homicide rate	2014
LSI	Life satisfaction index	2014
LAN	Composite index for landscape and cultural heritage	2011
ENV	Composite index of environment	2012

provides also the main characteristics of these distributions, such as mean and standard deviation (SD).

There are a number of points of interest in Table 2. For example, considering M_1 for country 2, we have no shifts when X_1 is removed, and 2 shifts when X_2 is removed. On the contrary, considering PC_1, we have 3 shifts when X_1 is removed, and only 1 shift when X_2 is removed. Overall, when X_1 is removed we have a mean shift of 0.8 for M_1 versus a mean shift of 2.4 for PC_1, whereas when X_2 is removed we have a mean shift of 1.2 for M_1 versus a mean shift of 0.4 for PC_1. Hence, on average, excluding an indicator, we have a greater shift with PC_1 (1.4 versus 1.0). Note also that M_1 has a low SD of the mean shift (0.20), whereas PC_1 has a high SD (1.00). This means that PC_1 is less robust and more sensitive to the inclusion or exclusion of an individual indicator compared to M_1.

3 Use of PCA for Studying Well-Being Indicators

In this section, some applications of PCA to well-being indicators are presented.

As we have seen, the measurement model for measuring well-being is formative. For this reason, it does not make sense to summarize correlated indicators for constructing a composite index, as they are not functions of a conceptual (latent) variable. Nevertheless, correlated indicators can be summarized for reducing data dimensionality in order to simplify graphical representation or to detect clusters of similar units. Correlations between principal components and original indicators can also help to identify groups of indicators that provide the same information and to find redundant indicators.

In the first case study, a set of composite indices of well-being for Italian regions were summarized by PCA. In the second case, a set of composite indices of well-being (one for each dimension and a global index) for Italian provinces were calculated and relationships between the global index of well-being and the principal components were investigated. It is noteworthy that, in both cases, PCA allows to 'quantify' the amount of information on well-being that cannot be derived from GDP per capita.

3.1 A Case Study for Italian Regions

The well-being indicators used in this work are selected from BES 2015 report (Istat 2015a). In particular, we considered the composite indices of 9 dimensions of the BES (Health, Education and training, Work and life balance, Economic well-being, Social relationships, Security, Subjective well-being, Landscape and cultural heritage, Environment) and some complementary indicators such as employment rate, homicide rate, and life satisfaction index. All the indicators are calculated at the Italian regional level.

In Table 3, the list of the indicators, with label and year of reference, is reported. For a detailed description of the indicators, see Istat (2015a).

Table 4 shows the correlation matrix of the 12 well-being indicators and the correlation between each of them and the GDP per capita (GDP) for 2014. As we can see, the majority of indicators are positively correlated among them (HEA, EDU, QOW, EMP, INC, HAR, REL, LSI e LAN), and the values are very high ($r \geq 0.700$). Even the composite index of environment (ENV) and the homicide rate (HOM) are positively correlated with this set of indicators, but the strength of the relationship is more moderate for ENV ($0.450 \leq r \leq 0.700$) and it is weak for HOM ($0.200 \leq r \leq 0.450$).

The composite index of safety (SAF), instead, shows a slight negative correlation with the other well-being indicators ($-0.250 \leq r \leq 0.200$).

Regarding the correlations of the 12 well-being indicators with the GDP, the highest correlation is observed with the employment rate (EMP), followed by the composite index of income and inequality (INC), the composite index of quality of work (QOW) and the composite index of health (HEA). The indicators less concordant with the GDP are the homicide rate (HOM), with $r = 0.554$, and the composite index of environment (ENV), with $r = 0.577$; whereas the composite index of safety (SAF) is the most discordant, as it shows a negative correlation with GDP ($r = -0.221$).

These results confirm that if, on the one hand, the main well-being indicators can be 'explained' by the GDP, some of them (e.g. those related to security and environment) are almost completely independent from this measure.

In order to study the overall structure of the dataset, an exploratory PCA was performed, as suggested in OECD (2008). As seen above, principal components are linear combinations of the starting indicators, they have decreasing importance and they are linearly uncorrelated themselves. This allows to describe the statistical units with a lower number of new dimensions, maximizing the proportion of variance accounted for.

In Fig. 2 the scree-plot (a) and the correlation circle (b) of PCA are shown. From the scree-plot examination, an elbow is evident at the second factor. This means that most of the variability of Italian regions (80.77%) can be explained by the first two factors. The third factor accounts for 7.62% of the remaining variance, but having an eigenvalue less than 1 ($\lambda = 0.914$) may be non-significant, according to the Kaiser's criterion (Guttman 1954; Kaiser 1961). By projecting the original indicators in the plane of the first two principal components, the circle of correlations is obtained, where each well-being indicator is represented by a point with coordinates equal to the two coefficients of correlation with the first and second factor. Note that the first factor is strongly correlated with 9 indicators of 12 (HEA, EDU, QOW, EMP, INC, HAR, REL, LSI and LAN), whereas the second one represents only the composite index of safety (SAF). Finally, the normalized homicide rate (HOM) and the composite index of environment (ENV) are to be placed in an intermediate position between the two axes, as they are partially correlated with both factors.

Figure 3 displays the graphical representation of the relationships between GDP per capita and the first two factors of PCA. The correlation between GDP per capita and the

Table 4 Correlations among well-being indicators and GDP

| Well-being indicator	HEA	EDU	QOW	EMP	INC	HAR	REL	SAF	HOM	LSI	LAN	ENV
HEA	1.000	0.842	0.911	0.917	0.906	0.876	0.902	-0.232	0.457	0.871	0.803	0.559
EDU	0.842	1.000	0.807	0.850	0.841	0.832	0.829	-0.119	0.190	0.826	0.763	0.540
QOW	0.911	0.807	1.000	0.963	0.963	0.913	0.842	-0.201	0.452	0.791	0.784	0.538
EMP	0.917	0.850	0.963	1.000	0.969	0.908	0.884	-0.229	0.413	0.821	0.809	0.494
INC	0.906	0.841	0.963	0.969	1.000	0.916	0.887	-0.172	0.452	0.858	0.799	0.555
HAR	0.876	0.832	0.913	0.908	0.916	1.000	0.845	-0.185	0.404	0.785	0.700	0.478
REL	0.902	0.829	0.842	0.884	0.887	0.845	1.000	-0.084	0.427	0.927	0.865	0.639
SAF	-0.232	-0.119	-0.201	-0.229	-0.172	-0.185	-0.084	1.000	-0.048	0.020	-0.129	0.178
HOM	0.457	0.190	0.452	0.413	0.452	0.404	0.427	-0.048	1.000	0.418	0.228	0.428
LSI	0.871	0.826	0.791	0.821	0.858	0.785	0.927	0.020	0.418	1.000	0.775	0.696
LAN	0.803	0.763	0.784	0.809	0.799	0.700	0.865	-0.129	0.228	0.775	1.000	0.532
ENV	0.559	0.540	0.538	0.494	0.555	0.478	0.639	0.178	0.428	0.696	0.532	1.000
GDP	0.889	0.748	0.889	0.928	0.899	0.834	0.873	-0.221	0.554	0.847	0.733	0.577
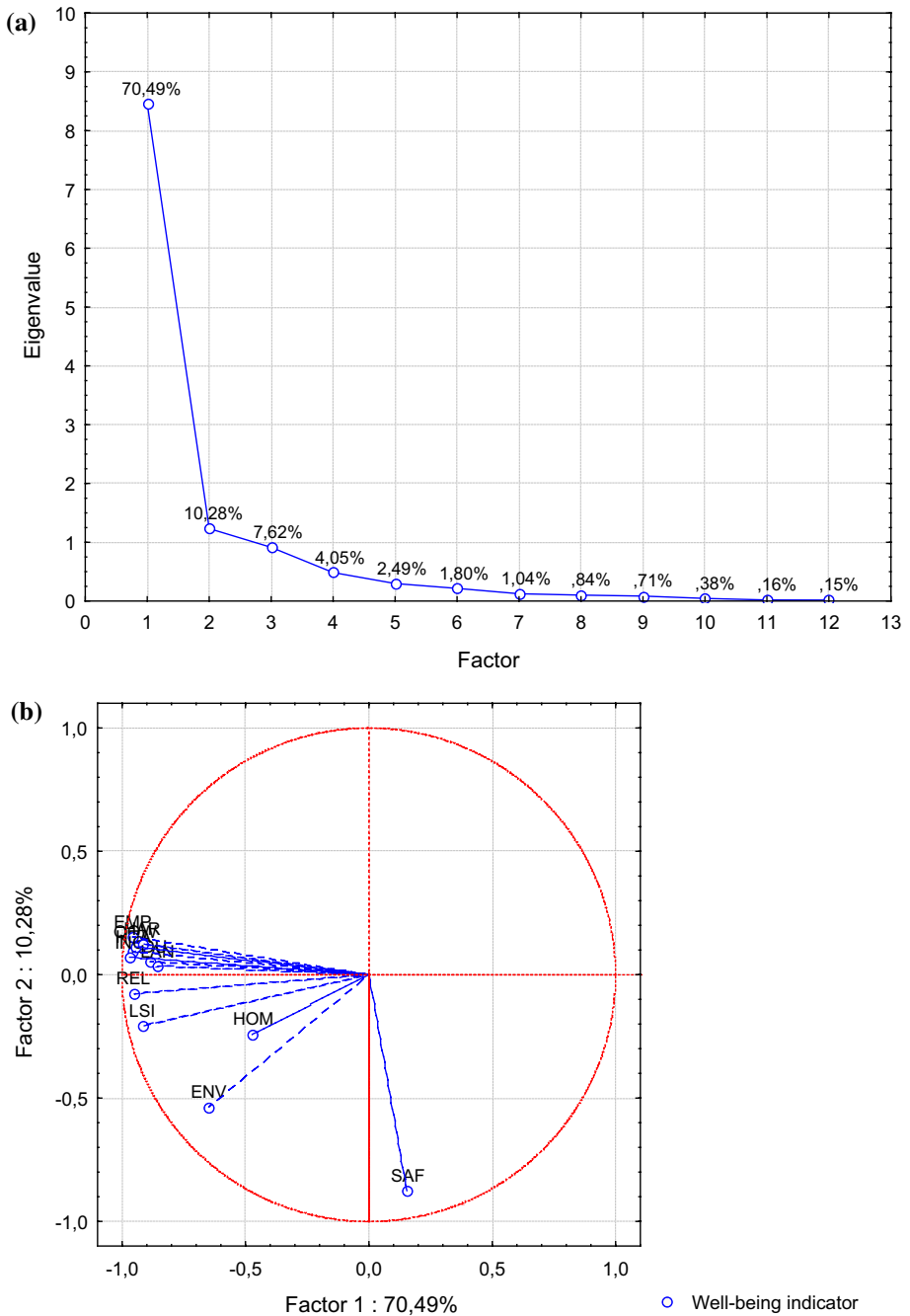
Fig. 2 Scree-plot and correlation circle of PCA
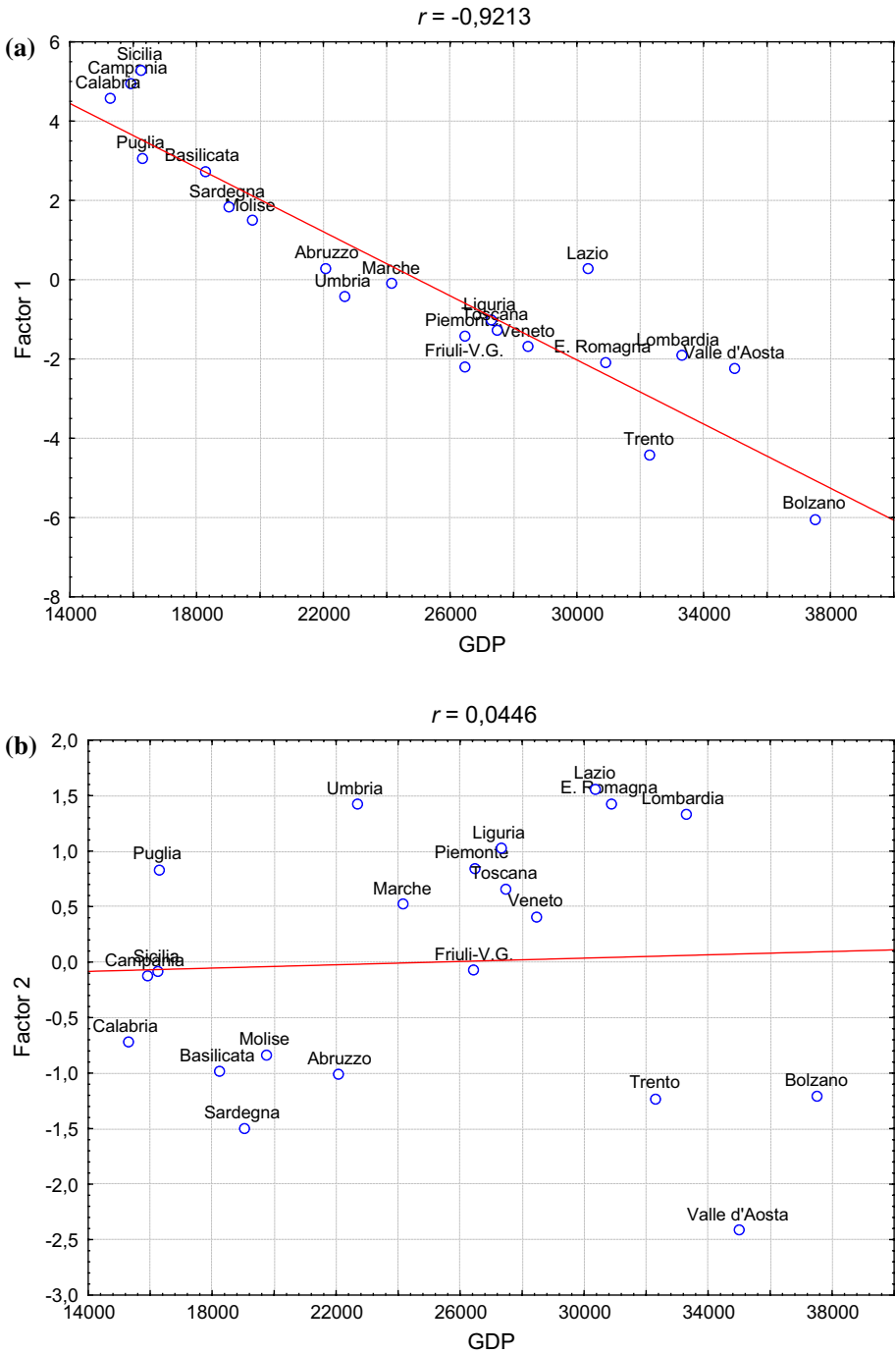
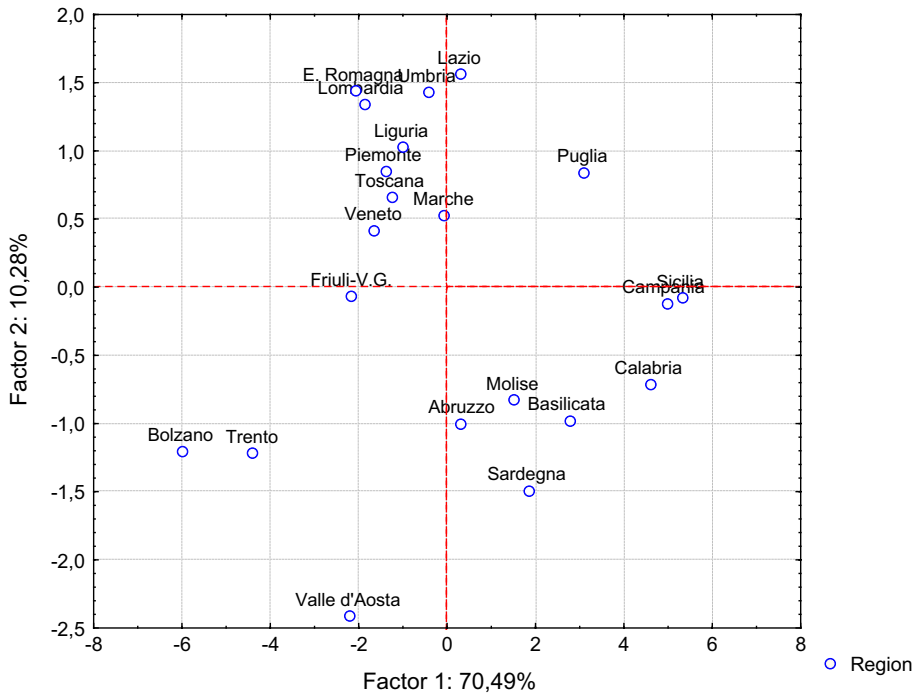Fig. 3 Relationships between GDP per capita and the first two factors of PCA

Fig. 4 First plane of PCA

first factor (in absolute value) is very high ($r=-0.9213$) confirming that a meaningful amount part of information on the well-being of the regions can be derived from GDP. On the other hand, it is noteworthy that the first factor accounts for about 70% of the total variance. Therefore, GDP does not 'capture' the remaining 30% of the information. In fact, the second factor of PCA, that represents security (SAF) and, in part, the environment (ENV), is totally uncorrelated with GDP per capita ($r=0.0446$).

In Fig. 4 the projections of Italian regions on the first plane of PCA are shown. The scatterplot highlights the usual polarization between northern regions (to the left along the x-axis) and southern regions (to the right along the x-axis). The higher the value of the first factor, the lower the GDP per capita of the region. The second factor, by contrast, represents most of the safety information.

Note that the first factor cannot be used as a composite index of well-being at least for two reasons. Firstly, it summarize a set of indicators only because they are correlated among themselves, but not because they are functions of a common latent variable. Secondly, it ignores some important indicators, such as SAF. In fact, it accounts for only 70% of the information about well-being.

3.2 A Case Study for Italian Provinces

The BES project has been extended for measuring well-being not only at the Italian regional level but also at the provincial level (Istat 2015b). From this point of view, the

analysis is even more interesting than the regional one as the number of statistical units is greater (110 provinces versus 21 regions).

In this case, we computed 11 composite indices for Italian provinces[16] with the aim of representing different dimensions or 'pillars'[17] of well-being (Health, Education and training, Work and life balance, Economic well-being, Social relationship, Politics and institutions, Security, Landscape and cultural heritage, Environment, Research and innovation, Quality of service). The correlations among composite indices and GDP per capita were evaluated and a PCA was carried out in order to reduce data complexity.

Table 5 reports the list of individual indicators used for constructing each composite index (Chelli et al. 2017). The polarity of each indicator is also provided.

Composite indices were created with a formative model by applying the same method as used in 2015 BES Report for Italian regions, namely the Adjusted Mazziotta-Pareto Index (AMPI). Specifically, for each pillar P_i ($i = 1, \ldots, 11$), a composite index was computed, under the hypothesis of non-substitutability of the components, and the formula of the AMPI with negative penalty was used (Mazziotta and Pareto 2016a). Similarly, a global well-being index was obtained, by aggregating the 11 composite indices with AMPI. In this way, we obtained both a ranking of Italian provinces for each dimension of well-being and a general ranking ('one number' for each province). The individual indicators used try to emulate the theoretical framework of the national BES even if, in some cases, it is impossible have exactly the same measure since many sample surveys estimate parameters only at the regional level (Istat 2015b).

Correlations among the 11 composite indices and GDP per capita are shown in Table 6. The most of the composite indices (P_1–P_6, P_8, P_{10} and P_{11}) are positively intercorrelated ($0.244 \leq r \leq 0.810$), excepted for P_7 (Security) and P_9 (Environment) that are negatively correlated with some of them. This means that the dimensions of well-being concerning Health, Education and training, Work and life balance, Economic well-being, Social relationship, Politics and institutions, Landscape and cultural heritage, Research and innovation, Quality of service are, with different intensity, concordant among themselves. Only Security and Environment are, in some cases, discordant from the others dimensions. P_7 and P_9 are also negatively correlated with the GDP per capita; whereas the other composite indices are all positively correlated with it ($0.302 \leq r \leq 0.848$).

Figure 5 displays the scree-plot (a) and the correlation circle (b) of PCA for this analysis. From the two graphs, we see that the first factor of PCA for Italian provinces accounts for 47.22% of the total variance and it is negatively correlated with P_1-P_6, P_8, P_{10} and P_{11}. By contrast, the second factor accounts for 16.30% of the total variance and it is negatively correlated, above all, with P_7 and P_9. So, the first plane of PCA accounts for about 63.5% of the variability of Italian provinces.

The scatterplots of the first two factors versus the GDP per capita are given in Fig. 6. Similarly to the case of Italian regions, the first factor is strongly correlated (in absolute values) with the GDP per capita ($r = -0.8133$), despite the presence of two outliers, such as Rome (RM) and Milan (MI). On the contrary, the second factor is weakly correlated with it ($r = 0.2646$). However, the amount of total variance 'explained' from GDP per capita seems very lower for Italian provinces, as the variance accounted for by the first factor is less than 50%.

---

[16] Note that only individual indicators are released by Istat at the provincial level.

[17] A pillar describes a particular aspect—not directly observable—of the latent phenomenon by a set of individual indicators which are assumed to be related to it.

Table 5 Well-being indicators for Italian provinces. *Source*: Istat (2015b)

Label	Composite index	Individual indicator	Polarity
P_1	Health	Life expectancy at birth (M)	+
		Life expectancy at birth (F)	+
		Avoidable mortality	−
P_2	Education and training	Young people leaving school early	−
		People of working age with no higher education	−
		Competence level alphabetic students	+
		Level of digital competence of students	+
		People of working age in lifelong learning	+
P_3	Work and life balance	Rate of non-attendance at work (15–74 years)	−
		Gender difference in the rate of non-participation (F–M)	−
		Employment rate (20–64)	+
		Gender differences in the employment rate (M–F)	−
		Youth employment rate (15–29 years)	+
		Rate risk for serious accidents at work	−
P_4	Economic well-being	Estimated gross disposable income per household	+
		Average amount of family assets	+
		Gender differences in the average wage employees (M–F)	−
		Differences of generation in the average wage employees	−
P_5	Social relationship	Dissemination of non-profit institutions	+
		Volunteers for 100 residents aged 14 and over	+
P_6	Politics and institutions	Turnout in the European elections	+
		Turnout in provincial elections	+
		Percentage of women in municipalities	+
		Percentage of young people (< 40 years old) in municipalities	+
P_7	Security	Violent crimes reported	−
P_8	Landscape and cultural heritage	Consistency of the historic urban fabric in good condition	+
		Density of urban parks and green of historical interest	+
		Museums accessible	+
P_9	Environment	Availability of urban green	+
		Overruns limits air pollution—PM10 (Maximum)	−
		Energy produced from renewable sources	+
		Municipal waste landfilled	−
P_{10}	Research and innovation	Propensity to patent (applications)	+
		Flows of new graduates in S & T residents (total)	+
		Specialization in knowledge-intensive sectors	+
P_{11}	Quality of service	Electricity outages without notice	−
		Children 0–2 years old receiving services for children	+
		Separate collection of municipal waste	+
		Index of overcrowding of prisons	−
		Emigration hospital in another region	−
		Density of urban networks of local transport	+

Table 6 Correlation among well-being composite indices and GDP

| Composite index	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
P1	1.000	0.499	0.619	0.436	0.448	0.496	0.162	0.547	−0.060	0.585	0.594
P2	0.499	1.000	0.639	0.300	0.387	0.410	0.079	0.480	−0.149	0.542	0.439
P3	0.619	0.639	1.000	0.492	0.668	0.547	0.058	0.722	−0.139	0.725	0.810
P4	0.436	0.300	0.492	1.000	0.520	0.403	0.153	0.530	0.113	0.249	0.445
P5	0.448	0.387	0.668	0.520	1.000	0.244	0.317	0.694	0.343	0.365	0.584
P6	0.496	0.410	0.547	0.403	0.244	1.000	−0.035	0.443	−0.195	0.511	0.426
P7	0.162	0.079	0.058	0.153	0.317	−0.035	1.000	0.090	0.332	0.045	−0.030
P8	0.547	0.480	0.722	0.530	0.694	0.443	0.090	1.000	0.149	0.404	0.598
P9	−0.060	−0.149	−0.139	0.113	0.343	−0.195	0.332	0.149	1.000	−0.322	−0.213
P10	0.585	0.542	0.725	0.249	0.365	0.511	0.045	0.404	−0.322	1.000	0.676
P11	0.594	0.439	0.810	0.445	0.584	0.426	−0.030	0.598	−0.213	0.676	1.000
GDP	0.630	0.632	0.848	0.302	0.550	0.472	−0.115	0.549	−0.257	0.738	0.748

(a)

(b)
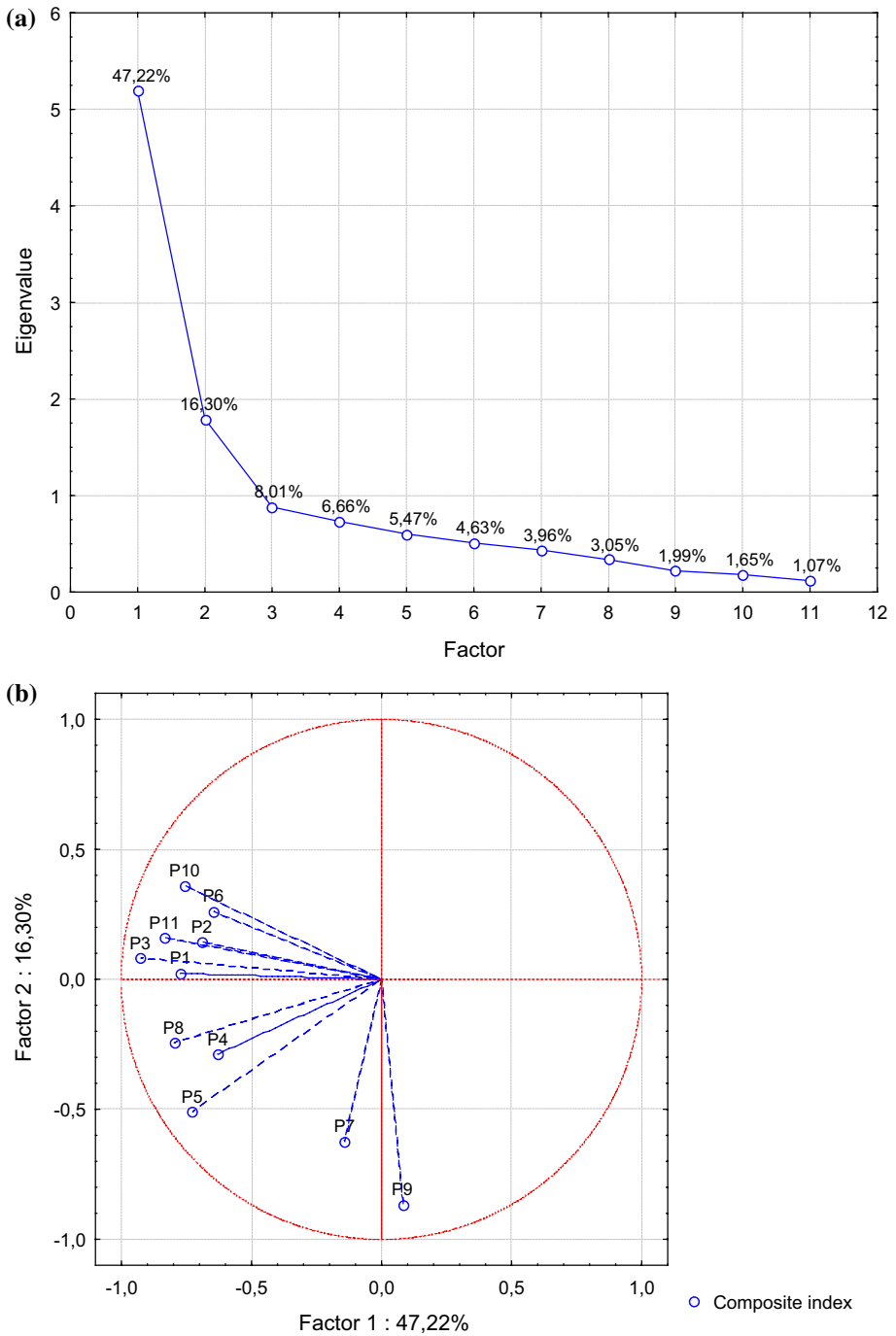
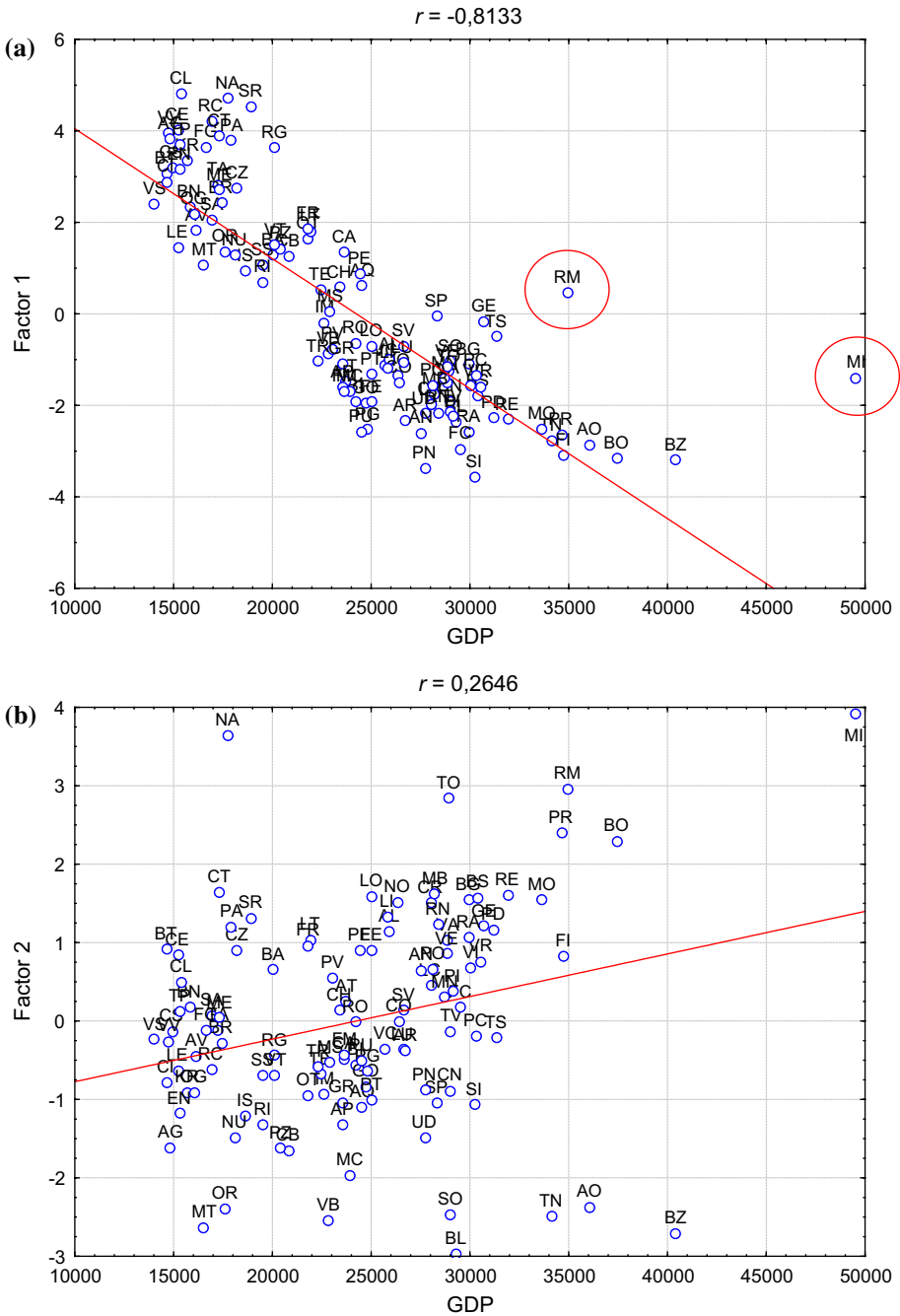**Fig. 5** Scree-plot and correlation circle of PCA

Fig. 6 Relationships between GDP per capita and the first two factors of PCA
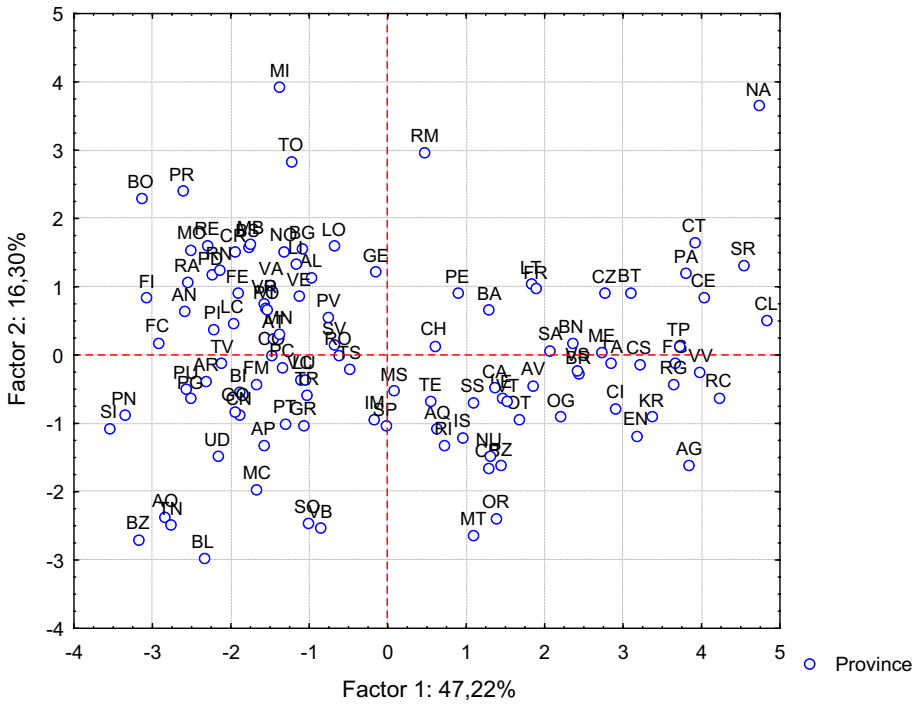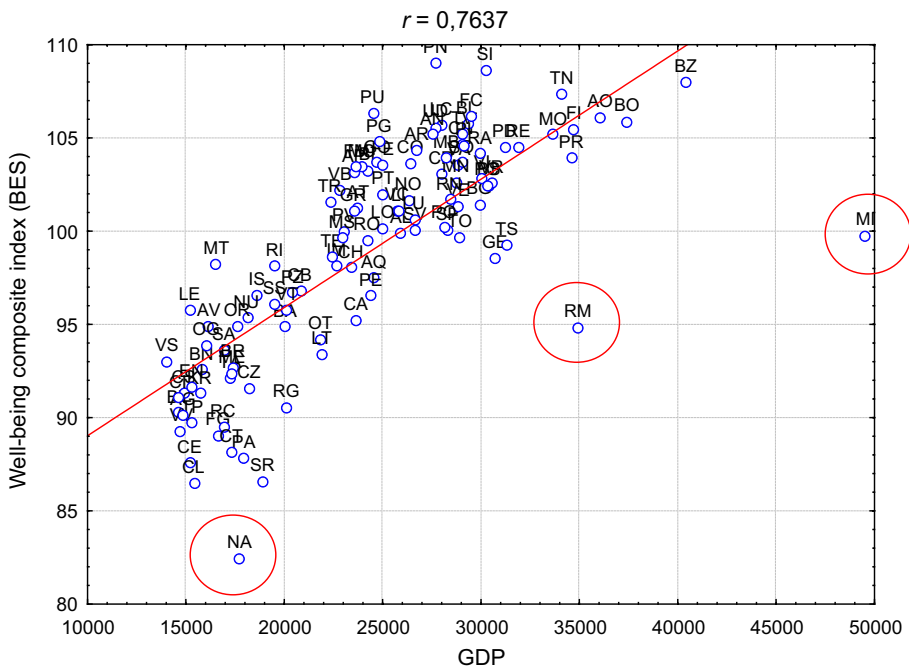
Fig. 7 First plane of PCA

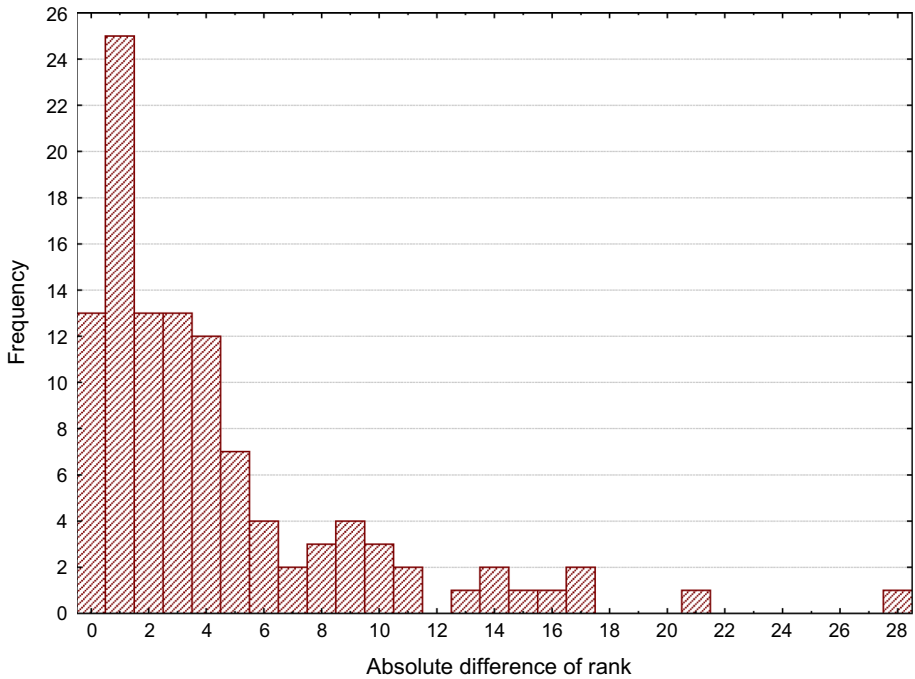**Fig. 8** Relationship between GDP per capita and Well-being composite index

Fig. 9 Comparison of rankings based on first factor of PCA and BES index

Figure 7 shows the projection of the provinces on the first plane of PCA, where the polarization between northern provinces (to the left along the *x*-axis) and southern provinces (to the right along the *x*-axis) is reproduced. The higher the value of the first factor, the lower the GDP per capita of the province. Note that three big provinces such as Rome (RM), Milan (MI) and Naples (NA) are placed at the top of the map, away from the rest of the group.

After calculating the global well-being composite index (BES), it was correlated with the GDP per capita ($r = -0.7637$). The relationship between this two measures is shown in Fig. 8 and it is very similar to the relationship between GDP per capita and first factor of PCA (Fig. 6a).

However, in Fig. 8, also Naples (NA) can be considered an outlier, although it has different characteristics from Rome (RM) and Milan (MI). This means that the BES index is able to 'capture' some aspects of well-being that the first factor of PCA ignores. In fact, Naples has a GDP per capita greater than Caltanissetta (CL), but a very lower level of well-being.

Comparing the two rankings based on the first factor of PCA and the BES index, we obtain a mean absolute difference of rank of 4.3 (i.e. the rank of each province changes, on average, by 4.3 positions), with a maximum of 28 positions. Figure 9 shows the distribution of absolute differences of rank. As can be seen from the histogram, only 12 percent of the provinces occupy the same place in the two rankings, whereas 42 percent of them move at least 4 ranking positions, because the first factor does adequately

consider P_7 and P_9. Indeed, the weights of the individual indicators on the first factor are based on the correlations among indicators and not on their real importance.

In this case too, PCA can be an useful tool for understanding the phenomenon, analysing correlations and visualizing data, but a composite index of well-being, such as the BES index, must be created following a formative approach.

4 Final Remarks

The construction of composite indices for measuring multidimensional phenomena, such as the human well-being, is a central issue in data analysis. Researcher cannot solve this question simply by using PCA or related methods, such as Factor Analysis, since they are typically used for a reflective approach and they ignore the polarities, namely the meaning of the individual indicators. Furthermore, a PCA based index accounts for a limited part of the total variance, it does not include all the non-redundant information of the individual indicators and it does not allow making inter-spatial and inter-temporal comparisons.

Reducing dimensionality and constructing composite indicators are two separate issues that are repeatedly confused. Both the procedures aims to summarize a set of variables or individual indicators, but reducing dimensionality focuses on extracting the most important information from the data, whereas constructing composite indicators focuses on the use of a measurement model that can be reflective or formative.

Extracting the most important information from the data translates in summarizing correlated indicators, but correlations can indicate causal, non-causal (spurious) and coincidental relationships, making the principal components meaningless or difficult to interpret. On the contrary, defining a measurement model means assuming a specific direction of causality between the measures (individual indicators) and the latent variable (phenomenon to be measured).

Measuring well-being requires a formative approach, where the index to be constructed does not exist as an independent entity, but it is a composite measure directly determined by a set of non-interchangeable individual indicators or pillars (e.g. the HDI or the CIW).

Therefore, in order to obtain a valid and reliable measure, it is absolutely essential to define the theoretical framework with an appropriate measurement model. This paradigm should always be considered when the objective of the research is to measure a multidimensional phenomenon through composite indices. And this is even more valid if the phenomenon to be measured is human well-being, as this latent factor depends on a set of individual indicators that influence it and not vice versa.

In such a context, PCA is recommended for different reasons. Firstly, PCA is a powerful tool for reducing complexity and visualizing data, so that the researcher can identify clusters of units (regions, provinces or countries) that have the same characteristics. Secondly, it allows for comparing empirical dimensions (factors) with theoretical dimensions (pillars), in order to evaluate any differences and to detect possible dimensions that had not previously been taken into account. Lastly, PCA makes it easy to study correlations among many individual indicators in order to find redundant and non-redundant indicators and to assess linkages with other relevant measures, such as GDP.

Nevertheless, the use of PCA for constructing formative composite indices can give very misleading information about the latent variable of interest, because it is exclusively based on the covariance structure between the individual indicators (Fayers and Hand 2002).

References

Biswas, B., & Caliendo, F. (2002). A multivariate analysis of the human development index. *Indian Economic Journal, 49,* 96–100.

Bleys, B. (2012). Beyond GDP: Classifying alternative measures for progress. *Social Indicators Research, 109,* 355–376.

Boarini, R., Kolev, A., & McGregor, A. (2014). *Measuring well-being and progress in countries at different stages of development: Towards a more universal conceptual framework*. Working Paper No. 325. OECD Development Centre.

Bohrnstedt, G. W. (1970). Reliability and validity assessment in attitude measurement. In G. F. Summers (Ed.), *Attitude measurement* (pp. 80–99). London: Rand McNally.

Booysen, F. (2002). An overview and evaluation of composite indices of development. *Social Indicators Research, 59,* 115–151.

Borsboom, D., Mellenbergh, G. J., & Heerden, J. V. (2003). The theoretical status of latent variables. *Psychological Review, 110,* 203–219.

Cadogan, J. W., & Lee, N. (2013). Improper use of endogenous formative variables. *Journal of Business Research, 66,* 233–241.

Chelli, F., Ciommi, M., Emili, A., Gigliarano, C., & Taralli, S. (2017). A new class of composite indicators for measuring well-being at the local level: An application to the Equitable and Sustainable Well-being (BES) of the Italian Provinces. *Ecological Indicators, 76,* 281–296.

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,* 1250–1262.

Diamantopoulos, A. (2006). The error term in formative measurement models: Interpretation and modeling implications. *Journal of Modeling in Management, 1,* 7–17.

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

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

Dunteman, G. H. (1989). *Principal components analysis*. Newbury Park: Sage.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5,* 155–174.

Fabrigar, L. F., & Wegener, D. T. (2011). *Exploratory factor analysis*. New York: Oxford University Press.

Fayers, P. M., & Hand, D. J. (2002). Causal variables, indicator variables and measurement scales: An example from quality of life. *Journal of the Royal Statistical Society, Series A, 165,* 233–261.

Ferrara, A. R., & Nisticò, R. (2014). *Measuring well-being in a multidimensional perspective: A multivariate statistical application to Italian regions*. Working Paper, 6. Dipartimento di Economia, Statistica e Finanza, Università della Calabria.

Filzmoser, P. (1999). Robust principal components and factor analysis in the geostatistical treatment of environmental data. *Environmetrics, 10,* 363–375.

Götz, O., Liehr-Gobbers, K., & Krafft, M. (2010). Evaluation of structural equation models using the partial least squares (PLS) approach. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods, and applications* (pp. 691–711). Berlin: Springer.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika, 19,* 149–161.

Haq, R., & Zia, U. (2013). Multidimensional wellbeing: An index of quality of life in a developing economy. *Social Indicators Research, 114,* 997–1012.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology, 24,* 417–441.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods, 12,* 205–218.

Hubert, M., Rousseeuw, P. J., & Vanden Branden, K. (2005). Robpca: A new approach to robust principal component analysis. *Technometrics, 47,* 64–79.

Istat (2015a). *BES 2015. Il benessere equo e sostenibile in Italia*. http://www.istat.it/it/files/2015/12/Rapporto_BES_2015.pdf. Accessed 21 May 2018.

Istat (2015b). *Il benessere equo e sostenibile delle province*. http://www.besdelleprovince.it/fileadmin/grpmnt/1225/pubblicazione_nazionale.pdf. Accessed 21 May 2018.

Jarvis, C. B., Mackenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30,* 199–218.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society, A, 374,* 20150202. https://doi.org/10.1098/rsta.2015.0202.

Kaiser, H. F. (1961). A note on Guttman's lower bound for the number of common factors. *British Journal of Mathematical ans Statistical Psychology, 14,* 1–2.

Kendall, M. G., & Stuart, A. (1968). *The advanced theory of statistics* (Vol. 3). London: Charles Griffin & Co.

Krishnakumar, J., & Nagar, A. L. (2008). On exact statistical properties of multidimensional indices based on principal components, factor analysis, MIMIC and structural equation models. *Social Indicators Research, 86,* 481–496.

Lai, D. (2003). Principal component analysis on human development indicators of China. *Social Indicators Research, 61,* 319–330.

Lavit, C., Escoufier, Y., Sabatier, R., & Traissac, P. (1994). The ACT (STATIS method). *Computational Statistics & Data Analysis, 18,* 97–119.

Linting, M., Meulman, J. J., Groenen, P. J. F., & Van der Kooij, A. J. (2007). Nonlinear principal components analysis: Introduction and application. *Psychological Methods, 12,* 336–358.

Maggino, F. (2017). Developing indicators and managing the complexity. In F. Maggino (Ed.), *Complexity in society: From indicators construction to their synthesis* (Vol. 70, pp. 87–114)., Social indicators research series Cham: Springer.

Maggino, F., & Zumbo, B. D. (2012). Measuring the quality of life and the construction of social indicators. In K. C. Land, A. C. Michalos, & M. J. Sirgy (Eds.), *Handbook of social indicators and quality-of-life research* (pp. 201–238). Dordrecht: Springer.

Markus, K. A., & Borsboom, D. (2013). *Frontiers of test validity theory. Measurement, causation, and meaning*. New York: Routledge.

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.

Mazziotta, M., & Pareto, A. (2016a). On a generalized non-compensatory composite index for measuring socio-economic phenomena. *Social Indicators Research, 127,* 983–1003.

Mazziotta, M., & Pareto, A. (2016b). On the construction of composite indices by principal components analysis. *Rivista Italiana di Economia Demografia e Statistica, LXX*(1), 103–109.

Mazziotta, M., & Pareto, A. (2017). Synthesis of indicators: The composite indicators approach. In F. Maggino (Ed.), *Complexity in society: From indicators construction to their synthesis* (Vol. 70, pp. 159–191)., Social indicators research series Cham: Springer.

McGillivray, M. (2005). Measuring non-economic well-being achievement. *Review of Income and Wealth, 51,* 337–364.

Michalos, A. C. (2014). *Encyclopedia of quality of life and well-being research*. Dordrecht: Springer.

Michalos, A. C., Smale, B., Labonté, R., Muharjarine, N., Scott, K., Moore, K., et al. (2011). *The Canadian Index of wellbeing*. Technical report 1.0. Waterloo, ON: Canadian Index of Wellbeing and University of Waterloo.

Mishra, S. K. (2007). *A comparative study of various inclusive indices and the index constructed by the principal components analysis*. MPRA Paper, No. 3377. MPRA. http://mpra.ub.uni-muenchen.de/3377. Accessed 21 May 2018.

Mishra, S. K. (2008). On Construction of Robust Composite Indices by Linear Aggregation. SSRN. http://ssrn.com/abstract=1147964. Accessed 21 May 2018.

Nahman, A., Mahumani, B. K., & De Lange, W. J. (2016). Beyond GDP: Towards a green economy index. *Development Southern Africa*. https://doi.org/10.1080/0376835X.2015.1120649.

OECD. (2004). *The OECD-JRC handbook on practices for developing composite indicators*. Paper presented at the OECD Committee on Statistics, 7–8 June 2004, OECD, Paris.

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

OECD. (2015). *How's life? 2015: Measuring well-being*. Paris: OECD Publishing.

Osborne, J. W. (2014). *Best practices in exploratory factor analysis*. Newbury Park: Jason W. Osborne.

Ram, R. (1982). Composite indices of physical quality of life, basic needs fulfilment, and income: A principal component representation. *Journal of Development Economics, 11,* 227–247.

Salzman, J. (2003). *Methodological choices encountered in the construction of composite indices of economic and social well-Being*. Technical Report. Center for the Study of Living Standards, Ottawa.

Sen, A. K. (1985). *Commodities and capabilities*. Amsterdam: North Holland Publishing Company.

Shalizi C. R. (2009). *The truth about principal components and factor analysis*. http://www.stat.cmu.edu/~cshalizi/350/lectures/13/lecture-13.pdf.

Shwartz, M., Restuccia, J. D., & Rosen, A. K. (2015). Composite measures of health care provider performance: A description of approaches. *The Milbank Quarterly, 93,* 788–825.

Simonetto, A. (2012). Formative and reflective models: State of the art. *Electronic Journal of Applied Statistical Analysis, 5,* 452–457.

Slottje, D. J. (1991). Measuring the quality of life across countries. *The Review of Economics and Statistics, 73,* 684–693.

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

Stiglitz, J., Sen, A. K., & Fitoussi, J. P. (2009). Report of the commission on the measurement of economic performance and social progress. Paris. Available online from the Commission on the Measurement of Economic Performance and Social Progress: http://www.stiglitz-sen-fitoussi.fr/en/index.htm. Accessed 21 May 2018.

UNDP. (1990). *Human development report 1990*. New York: Oxford University Press.

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

Van Beuningen, J., Van der Houwen, K., & Moonen, L. (2014). *Measuring well-being. An analysis of different response scales*. Discussion Paper, 3. Statistics Netherlands.

Vinzi, V. E., Lauro, C., & Tenenhaus, M. (2003). *PLS path modeling*. Working paper. DMS – University of Naples, HEC – School of Management, Jouy-en-Josas.

Wong, K. M. (2012). Well-being and economic development: A principal components analysis. *International Journal of Happiness and Development, 1,* 131–141.

Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26, pp. 45–79)., Psychometrics Boston: Elsevier.