ORIGINAL RESEARCH

A Multivariate Statistical Analysis of Equitable and Sustainable Well-Being Over Time

Adriana Monte[1] · Gabriella Schoier[1]

Abstract
The Italian National Institute of Statistics (ISTAT) declined a multidimensional approach to measure equitable and sustainable well-being (Benessere Equo Sostenibile, BES) at a detailed territorial level, that is, at the provincial level (NUTS3) entailing a wide spectrum of indicators grouped into domains related to Health, Education, Work and life balance, Economic well-being, Social relationships, Politics and institutions, Security, Landscape and cultural heritage, Environment, Innovation research and creativity, and Quality of services. These indicators can help in describing the territories because they can spot situations of concern, such as in the South of Italy. The gap between North and South Italy has increased over time, a picture we can describe using each of the indicators, but also by jointly considering them by using a factor analysis and clustering and constructing a composite indicator. Contrary to some researchers who have suggested the employment of time series indicators for the period 2004–2016 looking for a latent factor, that takes into account the joint temporal trend of the indicators, in this paper, the situation among three different moments over time is compared: before the 2008 crisis and a few years after it. To do this, we have chosen a certain number of indicators on the basis of their features.

1 Introduction

In recent years, the study of economic evolution regarding the well-being of people and households has become an important research issue because of the possibility to gather and store a large amount of data collected over time (see, e.g., Mazziotta and Pareto 2017).

The rapid progress made in science and technology has contributed to the evidence of an evolving world, and for this reason, new perspectives in knowledge discovery upon economic

---

✉ Adriana Monte
    adriana.monte@deams.units.it

    Gabriella Schoier
    gabriella.schoier@deams.units.it

[1]   DEAMS, Department of Economic, Business, Mathematic and Statistical Sciences, "Bruno de Finetti", University of Trieste, Trieste, Italy

and social data with a time-oriented perspective are needed. The paradigm of "change mining" (Oliveira and Gama 2012) has arisen as a consequence of this evolution.

Data mining mechanisms that monitor models and patterns over time, compare them, detect changes, and describe these changes have become more and more useful. Having this in mind, some data mining researchers have developed methods and techniques to study the evolution of different phenomena over time (see Aggarwal 2005; Spiliopoulou et al. 2009).

Moreover some macro-economic statistics such as GDP seem to not give a detailed picture of the living conditions of common people (see, e.g., Maggino 2017; Chelli et al. 2016).

The perception of credibility and accountability of public policies has been highlighted in recent years because of the financial and economic crisis.

The Organization for Economic Co-operation and Development (OECD) states, "the OECD Framework for Measuring Well-Being and Progress is built around three distinct domains: material conditions, quality of life and sustainability, each with their relevant dimensions" (OECD 2019).

In this paper, we address the problem of monitoring the evolution of well-being in Italy over time. This can help decision makers of different areas make better economic and political decisions. The indicators, one for each BES domain, have been used to measure well-being. They were selected from the ISTAT database (see ISTAT 2019) and refer to the period 2004–2018.

The aim of this research is to understand how equitable and sustainable well-being in the territories (provinces) has changed; this will be done through a comparison between the pre-crisis year 2008 and subsequent years. In particular, we are interested not so much in identifying a ranking among the provinces but rather in seeing how the various territorial areas have moved in the period and which indicators have a greater weight. For this purpose, the ISTAT database containing the BES territorial indicators was used. This database contains the time series of different indicators, but not all of them are available for the period of interest.

The new idea regards the application of the Multiple Factor Analysis (MFA) to selected indicators for Italian provinces in different periods. MFA is applied to tables in which a set of individuals (one individual = one row) is described by a set of variables (one variable = one column). The advantage of MFA lies in the fact that within the active variables, it can account for a group structure defined by the user. Such data tables are called *individuals × variables organized into groups* (Escofier and Pages 1998).

The idea is different from that proposed by Mazziotta and Pareto (2017) because our objective is not to construct a composite indicator but rather to see how different indicators (one for each domain) move to analyze and then monitor the evolution of well-being in Italy over time.

In the analysis, we do not use a composite indicator for each domain of the BES (see Mazziotta and Pareto 2019); we only use one of the indicators of each domain, choosing it based on objective criteria.

We performed the analysis using FactoMineR (Lê et al. 2008), a package of the R language.

2 Equitable and Sustainable Well-Being at Local Level

Equitable and sustainable well-being (BES) is a multidimensional approach that identifies 12 well-being domains (ISTAT 2013; ISTAT 2018); for each of them, a set of indicators is given (at NUTS2 level). BES is becoming a more and more important tool to evaluate the progress of society from an economic, social, and environmental point of view.

Consequently, the Italian Economic and Financial Document has included some BES selected indicators since 2017.[1] The interest in BES has been growing over time, especially for Italian provinces and cities (NUTS3) (see Taralli 2013), so in the 2018, ISTAT issued for the first time a system of BES indicators at the NUTS3 level.

The BES domains at the local level are the same as those at national level, with an exception made for the Subjective well-being domain because of the lack of subjective indicators at the local level. The 11 domains and the indicators are listed in Table 1, and they belong to the 2019 version of the database (see ISTAT 2019). This version differs from the previous one in some aspects: variable definitions, introduction of new indicators, elimination of some previously available indicators, time span of new indicators, and territorial units (see ISTAT 2019).

In 2019, a set of indicators consisting of 55 measures was published; each domain is not formed by the same number of indicators, and almost half of the indicators do not give values before 2008. Table 1 shows three domains (Social relations, Landscape and cultural heritage, and Innovation research and creativity) that do not have data from before the 2008. Furthermore, the data related to the Social relations domain are still missing up to 2014. Some indicators present values only in well-defined years, that is, Voter turnout in European elections and Voter turnout in regional elections (Politics and institutions domain).

3 The Method: Multiple Factor Analysis

The need to simultaneously introduce quantitative and qualitative variables (known as mixed data) as active elements of one factorial analysis is usual in different statistical analyses (see e.g. Bolasco 1999).

The first method suggested within this framework is the Classical Canonical Analysis that, in practice, is not useful in the case of creating groups based on a given set of variables.

The problem of variables partition in different subspaces can be solved by using MFA (see e.g. Pages 2014). It relates to a Principle Component Analysis (PCA) that can analyse both quantitative and qualitative data (Escofier and Pages 1998); that is, it can handle multiple data tables that measure sets of variables collected on the same observations and these variables can be of different type.

In particular MFA is applied to tables in which a set of individuals (one individual = one row) is described by a set of variables (one variable = one column). The basic idea of MFA lies in the fact that within the active variables, it can account for a group structure defined by the user. Such data tables are called *individuals × variables organised into groups* (Escofier and Pages 1998).

In order to describe the MFA algorithm, one can see it as a "mixture" between a PCA for quantitative variables and a Multiple Correspondence Analysis (MCA) for the qualitative variables.

Its goal is to analyze several data sets measured on the same observations, to obtain a set of common factor scores and to plot each of the original data sets in a two dimensional space.

---

[1] See https://www.gazzettaufficiale.it/eli/id/2017/11/15/17A07695/sg.

Table 1 Time availability of indicators (ISTAT database from 2004 to 2018) Coloured cells refers to data availability; different colors allow to distinguish domains

Time availability of indicators (ISTAT database from 2004 to 2018)
Coloured cells refers to data availability; different colors allow to distinguish domains

N.	Indicators	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
	Health															
I.1	Life expectancy at birth															
I.2	Infant mortality rate															
I.3	Road accidents mortality rate(15–34 years old)															
I.4	Age-standardised cancer mortality rate															
I.5	Age-standardised mortality rate for dementia and nervous system diseases															
	Education															
I.6	Participation in early childhood education															
I.7	People with at least upper secondary education level (25-64 years old)															
I.8	People having completed tertiary education (30-34 years old)															
I.9	First-time entry rate to university by cohort of upper secondary graduates															
I.10	People not in education, employment, or training (Neet)															
I.11	Participation in long-life learning															
I.12	Level of literacy															
I.13	Level of numeracy															
	Work and life balance															
I.14	Employment rate (20-64 years old)															
I.15	Non-participation rate															
I.16	Incidence rate of fatal occupational injuries or injuries leading to permanent disability															
I.17	Young employment rate (15-29 years old)															
I.18	Non-participation rate (15-29 years old)															
I.19	Yearly paid days (employees)															
	Economic well-being															
I.20	Per capita disposable income															
I.21	Annual average remuneration of employees															
I.22	Per capita yearly pension amount															
I.23	Retired with pension of low amount															
I.24	Per capita assets															
I.25	Rate of bad debts of bank loans to families															
	Social relationships															
I.26	No-profit organizations															
I.27	Accessible schools															
	Politics and institutions															
I.28	Voter turnout in European elections															
I.29	Voter turnout in regional elections															
I.30	Women and political representation at municipality level															
I.31	Youngs (younger 40 years) and political representation at municipality level															
I.32	Percentage of prisoners in penal institutions on the total capacity of penal institutions															
I.33	Municipalities: capacity collection															
I.34	Provinces: capacity collection															
	Security															
I.35	Number of homicide on total population per 100,000 inhabitants															
I.36	Other violent crimes reported on total population per 10,000 inhabitants															
I.37	Number of other crimes reported (theft of any kind and robberies at home) on total population per 10,000 inhabitants															
I.38	Road mortality in extra-urban areas															
	Landscape and cultural heritage															
I.39	Density and importance of museums' heritage															
I.40	Spread of rural tourism facilities															
I.41	Presence of Historic Parks/Gardens and other Urban Parks recognised of significant public interest															
	Environment															
I.42	Water losses in urban supply system															
I.43	Landfill of waste															
I.44	Quality of urban air - PM10															
I.45	Quality of urban air - nitrogen dioxide															
I.46	Urban green															
I.47	Energy from renewable sources															
I.48	Separate collection of municipal waste															
I.49	Waterproofing of the soil from artificial cover															
	Innovation, research and creativity															
I.50	Cultural employment (% of total employment)															
I.51	Brain circulation (italians, 25-39 years old)															
	Quality of services															
I.52	Children who benefited of early childhood services															
I.53	Irregularities in electric power distribution															
I.54	Place-Km of public transport networks															
I.55	Hospital patients emigration to another region															

MFA procedures compute a PCA of each data table and normalize them by dividing all elements by the first singular value obtained. All the normalized data tables are aggregated into a new table that is analyzed via a non-normalized PCA; this new PCA is obtained by decomposing the variance of the "compromise" into a set of new orthogonal variables (i.e., the principal components are also often called dimensions, axes, factors, or even latent variables) ordered by the amount of variance that each component explains. The coordinates of the observations on the components are called *factor scores*; these can be used to plot maps of the observations in which the observations themselves are represented as points such that the distances in the map best reflect the similarities between them. The positions of the observations are called *partial factor scores* and can be represented as points on a map (Abdi et al. 2013).

In other words, the heart of MFA is a PCA in which the weights are assigned to the variables used in the analysis. More precisely, the same weight is associated to each variable of the group of the PCA on the group j ($j = 1,..., J$). The importance of the dimension represented by the principal component is given by its eigenvalue, which indicates how much of the total inertia (i.e., variance) of the data is explained by this component.

This shows that the inertia of *a group* represents the individuals' variability both from the point of view of their deviation from the centre of gravity and from of the between-individuals distances. Thus, the maximum axial inertia of each group of variables is equal to one.

The influence of the groups of variables in the global analysis must be balanced and the structure of each group must also be respected. The weight assigned to each variable presents a simple direct interpretation. It allows to consider MFA as a particular Generalized Canonical Analysis. For each group of variables, MFA analysis associates a set, that is, a "cloud" of individuals and a representation of these clouds.

This representation can be obtained in different ways: as a projection of a cloud of points, as a canonical variable or using, another idea, such as that proposed by Pages (2014). According to this last proposal the structure of the variables in the *J* groups ($j = 1,...,J$) and the use of a weighting of MFA given by the reciprocal of the first eigenvalue are taken into account. This prescaling entails that when a PCA is performed on the merged prescaled data sets, the resulting components will reflect a structure common to the data set.

Given the transition formula of the space of variables into the space of individuals, as written by Pages and Husson (2005), and taking into account the structure of variables in *J* groups and the weighting of MFA ($\overline{\lambda_1^j}$ if x_k belongs to group *j*), the $F_s(i)$, that is, the score of the individual *i* on the axis (of rank) *s* is given by:

$$F_s(i) = \frac{1}{\sqrt{\lambda_s}} \sum_{j=1}^{J} \frac{1}{\lambda_1^j} \sum_{k=1}^{K_j} x_{ik} G_s(k)$$

where:

K_j is the number of variables in group *j*,
$G_s(k)$ is the score of the individual *i* on the axis (of rank) *s*,
λ_s is the *s* eigenvalue associated to axis *s*,
λ_1^j is the first eigenvalue of group *j*,
x_{ik} is the general term of the data table (row *i*, column *k*).

This relationship is very important for interpreting the position of individuals with respect to the variables. We must note that on the graphical displays derived from MFA, each individual appears as a centroid of its partial representations. (see Pages and Husson 2005).

In the next paragraph, practical and theoretical notions referring to the object of this study are considered.

4 The Application

This section describes the data used and the results of applying MFA. As previously anticipated, we use the ISTAT database *Misure del benessere dei territori. Tavole di dati* (2019).[2] Because our aim is to describe how equitable and sustainable well-being has changed, we compare the situation before the 2008 crisis, after the 2011 crisis, and to the current one. Based on the considerations presented in the second paragraph, we choose the database performances of Italian provinces in 2007, 2012, and 2017.

We apply one indicator for each domain. This is because of the following considerations:

– the different number of indicators by domain would lead to an initial distortion, resulting in different weights for each domain;
– the use of a composite indicator for each domain to apply the MFA (see Mazziotta and Pareto 2019) is not appropriate here because in some domains, the indicators are not available for all the years considered;
– there are cases in which the choice of only one comparable indicator over time is the only possible, because of the presence of numerous missing data in the table in relation to some domains;
– a similar approach is used by Ciommi et al. (2017), in which the domains of the territorial BES are described by a single indicator given the limited availability of homogeneous time series.

4.1 The Data

The choice of indicators used in the analysis is based on different criteria: availability over time, discriminant capacity, variability, and the correlation among indicators inside each domain.

The evaluation of the discriminant capacity of indicators is based on Analysis of Variance to test for the differences among group means; we grouped provinces by macro socioeconomic region (North, Center, and South Italy). Because some indicators have not shown a significant difference in the mean values among macro-regions, we decided not to use them in the application. We report the selected indicators in Table 2, which shows the domains with one indicator (Economic well-being, Spread of rural tourism facilities, Innovation, research, and creativity) and others with a different number of indicators (from 2 to 5). To choose the better indicator in each domain, we considered the correlations within each domain (Figs. 1, 2, 3, 4, 5, and 6) and the variability indexes (Table 2) (i.e., coefficient

---

[2] The data are downloaded from https://www.istat.it/it/archivio/230627. In a previous presentation in abstracts SIEDS 2019, we used an ISTAT database published in 2018. In June 2019, a new database was issued by ISTAT, and the indicators are not always the same. The changes refer to variable definitions, introduction of new indicators, elimination of some previous ones, and time availability. See https://www.istat.it/it/benessere-e-sostenibilita/la-misurazione-del-benessere-(bes)/il-bes-dei-territori.

Table 2 Means, coefficients of variation and standardized Q_3-Q_1 (2007, 2012. 2017) for indicators selected by availability and discriminant capacity

N.	Name	Mean			C.V. %			Standardized Q_3-Q_1		
		2007	2012	2017	2007	2012	2017	2007	2012	2017
	Health									
I.1	Life expectancy at birth	81.24	81.93	82.72	0.90	0.89	0.91	1.54	1.51	1.46
I.2	Infant mortality rate	3.26	2.81	2.69	43.07	44.46	47.64	1.35	1.52	1.25
I.5	Age-standardised mortality rate for dementia and nervous system disease	23.83	29.87	31.04	18.42	17.06	18.64	1.37	1.28	1.49
	Education									
I.7	People with at least upper secondary education level (25-64 years old)	51.18	55.85	60.08	12.56	13.15	12.73	1.37	1.38	1.41
I.8	People having completed tertiary education (30-34 years old)	16.97	20.37	24.83	22.44	23.00	22.89	1.31	1.38	1.44
I.10	People not in education. employment. or training (Neet)	16.81	22.68	22.98	49.03	34.64	37.37	1.42	1.45	1.50
	Work and life balance									
I.14	Employment rate (20-64 years old)	63.19	60.99	62.31	14.61	16.41	17.47	1.59	1.63	1.71
I.15	Non-partecipation rate	14.58	20.12	20.73	67.99	51.92	53.17	1.48	1.72	1.59
I.16	Incidence rate of fatal occupational injuries or injuries leading to permanent disability	17.81	15.55	13.57	31.81	34.90	38.95	1.48	1.27	1.16
I.17	Young employment rate (15-29 years old)	41.78	33.57	31.35	25.75	27.30	26.85	1.67	1.61	1.70
I.18	Young non-partecipation rate (15-29 years old)	25.83	37.28	37.67	61.06	41.03	42.77	1.53	1.79	1.75
	Economic well-being									
I.25	Rate of bad debts of bank loans to families	0.95	1.32	1.34	34.91	28.71	30.40	1.05	1.59	1.48
	Social relationships									
I.26	No-profit organizations	-	-	62.48	-	-	25.85	-	-	1.20
I.27	Accessible schools	-	-	32.03	-	-	26.48	-	-	1.46
	Politics and institutions									
I.30	Women and political representation at municipality level	17.05	20.34	31.98	29.61	27.55	13.69	1.59	1.59	1.58
I.31	Youngs (younger 40 years) and political representation at municipality level	29.55	30.64	30.81	16.61	16.57	16.97	1.33	1.27	1.26
I.33	Municipalities: capacity collection	66.15	70.79	78.93	11.30	10.07	6.64	1.20	1.40	1.14
I.34	Provinces: capacity collection	68.85	70.17	77.17	18.13	15.80	14.56	1.52	1.44	1.47
	Security									
I.35	Number of homicide on total population per 100.000 inhabitants	0.90	0.85	0.61	108.79	112.02	104.09	0.61	0.87	0.83
I.37	Number of other crimes reported (theft of any kind and robberies at home) on total population per 10,000 inhabitants	194.70	210.91	170.65	40.04	41.86	44.16	1.18	1.05	1.18
I.38	Road mortality in extra-urban areas	6.02	5.33	4.86	42.81	49.20	50.22	1.29	1.23	1.14
	Landscape and cultural heritage									
I.40	Spread of rural tourism facilities	-	6.82	8.00	-	91.67	85.91	-	0.90	0.97
	Environment									
I.43	Landfill of waste	62.52	56.28	47.38	71.10	101.56	163.64	1.19	1.03	0.56
I.48	Separate collection of municipal waste	27.27	41.00	55.45	68.68	45.31	31.08	1.66	1.74	1.43
	Innovation, research and creativity									
I.50	Cultural employment (% of total employment)	1.21	1.17	1.11	37.19	35.58	33.42	0.66	0.48	0.81
	Quality of services									
I.52	Children who benefited of early childhood services	12.46	12.63	12.80	60.82	56.35	57.40	1.36	1.39	1.39
I.53	Irregularities in electric power distribution	2.10	2.41	1.830	63.50	60.40	46.75	1.13	1.46	1.31
I.55	Hospital patients emigration to another region	8.17	7.85	8.65	62.89	66.51	65.29	1.42	1.27	1.08
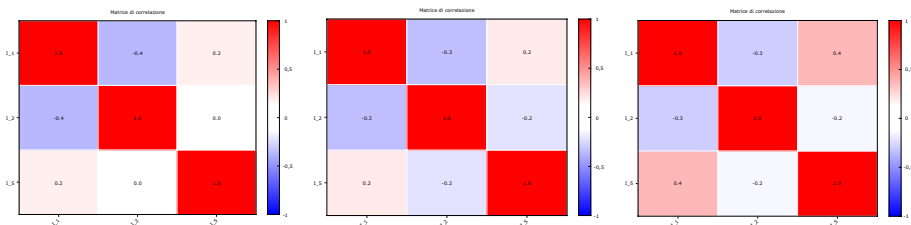
Fig. 1 Correlation matrices for Health indicators, 2007, 2012, 2017
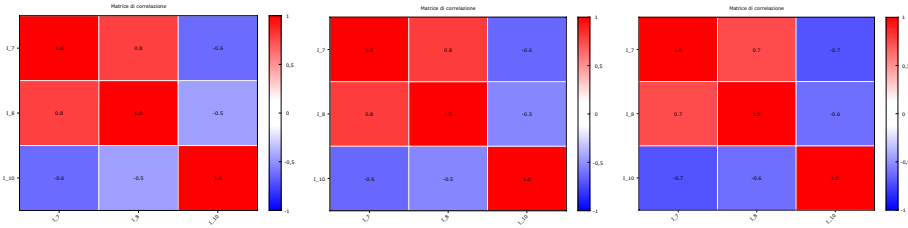
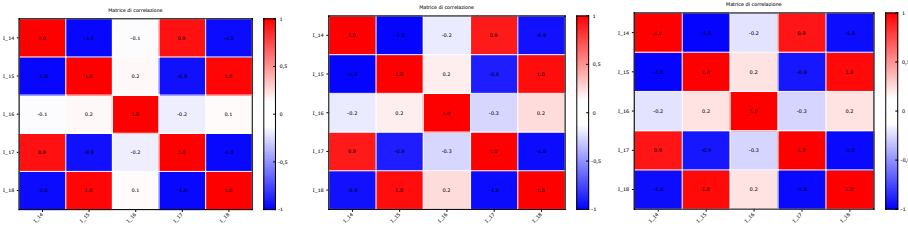**Fig. 2** Correlation matrices for Education indicators, 2007, 2012, 2017

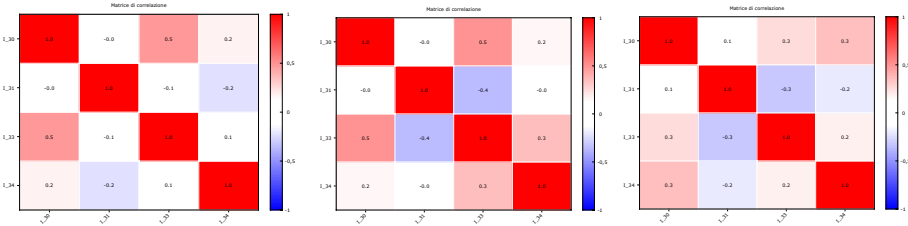**Fig. 3** Correlation matrices for Work and life balance, 2007, 2012, 2017

Fig. 4 Correlation matrices for Politics and institutions, 2007, 2012, 2017
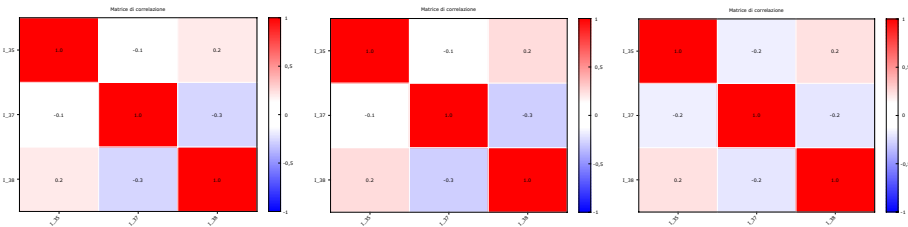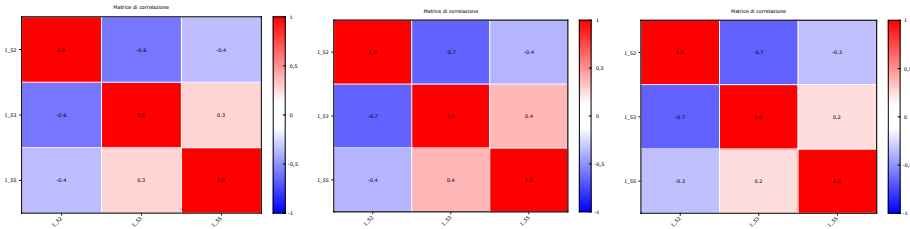

Fig. 5 Correlation matrices for Security, 2007, 2012, 2017

Fig. 6 Correlation matrices for Quality of services, 2007, 2012, 2017

Table 3 Domain, name and description of indicators used in MFA

Domain	N	Name	Description
Health	I_1	Life_Exp	Life expectancy at birth
Education	I_10	Neet	People not in education, employment, or training (Neet)
Work and life balance	I_15	Unempl	Non-partecipation rate
Economic well-being	I_25	Loans	Rate of bad debts of bank loans to families
Social relationships	I_27	Acc_Sc	Accessible schools
Politics and institutions	I_30	Women	Women and political representation at municipality level
Security	I_37	Crimes	Number of other crimes reported (theft of any kind and robberies at home) on total population per 10,000 inhabitants
Landscape and cultural heritage	I_40	Rural	Spread of rural tourism facilities
Environment	I_48	Waste	Separate collection of municipal waste
Innovation, research and creativity	I_50	Cult_Emp	Cultural employment (% of total employment)
Quality of services	I_53	Elect	Irregularities in electric power distribution

of variation and the quartile difference for standardized data). Figures 1, 2, 3, 4, 5 and 6 show the correlation matrices for each domain. The red colors indicate a positive correlation, while the blue colors show a negative one. The color intensity increases when the correlation rises. Within each domain, we selected those indicators that show both the largest correlation and the highest variability.

We perform our analysis by making use of different sets of indicators, and we decided to report those providing the best outcome. Table 3 lists the indicators used in the analysis; their outputs are described in Sect. 4.2.

4.2 The results of the MFA

In the current paper, we use the MFA methodology with three indexes: $i = 1,..,110$ for the provinces, $j = 1, 2, 3$ for the years, and $k = 1,…11$ for indicators. We run the MFA using the variables described in the previous paragraph. In the application of the MFA, we use indicators regardless of the polarity because we are not interested in a synthetic indicator obtained by a PCA (Mazziotta and Pareto 2019), but only in detecting the unit position over time regarding the selected indicators.

Because in some cases the choice of the indicator in each domain is not straightforward, we decided to repeat the analysis with a different set of indicators. This is the case of two domains and two different variables (Neet and Graduates for Education, Child and Elect for Quality of services). Below we report the application that provides the best results in terms of the explained variance, that is, using the indicators Neet and Elect.

The eigenvalues shown in Table 4 and in Fig. 7 suggest choosing the first two dimensions that explain the 60.54% of the total variance.

The partial analysis (see Table 5) for each of the three years reveals a decrease over time in the variability, which is explained by the first component, highlighting a process of convergence toward more similar values. The second component shows a different behavior.

Figure 8 presents the results of the analysis on the two-dimensional graph. If a variable is well represented (in the sense that its variability is well explained in the factorial dimension, i.e., that much of the variability is expressed in that factor), then its image on the factorial plane approaches the circumference, and the colors visually reinforce this fact. The more a variable forms a small angle with the factorial dimension, the more it is correlated with the factor and determines the interpretation of the axis. The variables that are well

Table 4 Eigenvalues of the first five factors (dimensions)

	Eigenvalue	Percent	Cumulative percent
Dimension 1	2.932	48.02	48.02
Dimension 2	0.765	12.52	60.54
Dimension 3	0.425	6.95	67.49
Dimension 4	0.374	6.13	73.62
Dimension 5	0.320	5.25	78.87
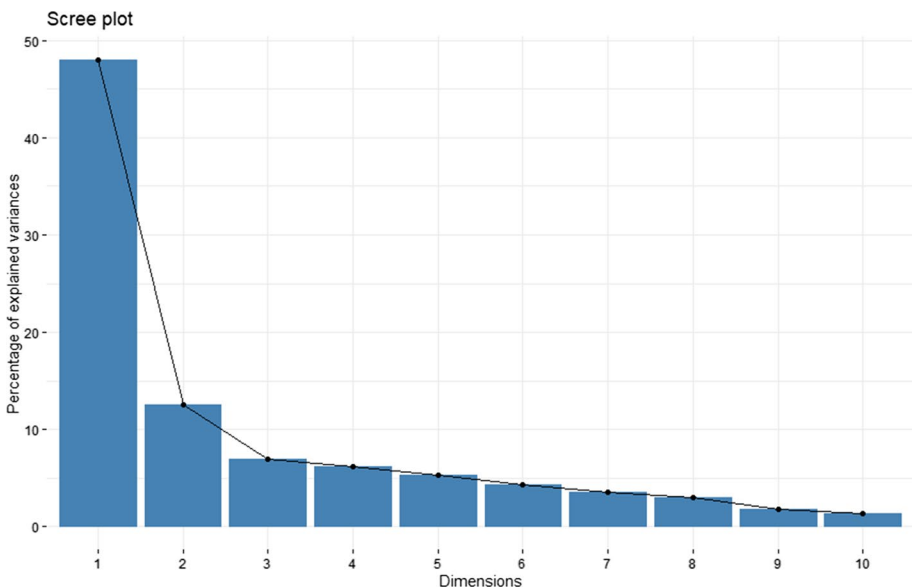
Fig. 7 Screeplot

Table 5 Partial analysis: eigenvalues of the first five component

	Eigenvalue	Percent	Cumulative Percent
2007			
Component 1	4.551	50.57	50.57
Component 2	1.334	14.82	65.39
Component 3	0.959	10.66	76.05
Component 4	0.697	7.75	83.80
Component 5	0.516	5.73	84.03
2012			
Component 1	4.930	49.30	49.30
Component 2	1.289	12.89	62.19
Component 3	0.913	9.13	71.32
Component 4	0.859	8.59	79.91
Component 5	0.557	5.57	85.48
2017			
Component 1	5.238	47.62	47.62
Component 2	1.105	13.13	60.75
Component 3	0.785	10.04	70.79
Component 4	0.630	7.14	77.93
Component 5	0.435	5.73	83.66

represented all over time refer to the following BES aspects: Health (Life_exp), Education (Neet), Work and life balance (Unempl), Security (Crimes), Environment (Waste), Innovation research and creativity (Cult_Emp). Regarding Social relations (Acc_sc), there is a good representation, but only for the last year, which is also the only available one. Politics (Women) has a good representation in the first two moments although it becomes worse in 2017. Landscape and cultural heritage (Rural) is poorly represented both in 2012 and 2017, but its data are not available in 2007. Economic well-being (Loans) has the worst representation, but it improves over time. The last one is Quality of services (Elect), which is not represented very well.

By focusing on the horizontal axis (dimension 1) we notice on the right hand side, the variables positively correlated with dimension 1 (Life_exp, Women, Rural, Waste) and on the left one those negatively correlated (Unempl, Elect, Loans, Neet). The first dimension explains the *socio-economic and environmental* aspects: increasing the values of this dimension would relate to an improvement in Health, Education, Work and life balance, Economic well-being, Social relationships, Politics and institutions, Landscape and cultural heritage, Environment, and Quality of services. This first dimension explains 48% of the variance.
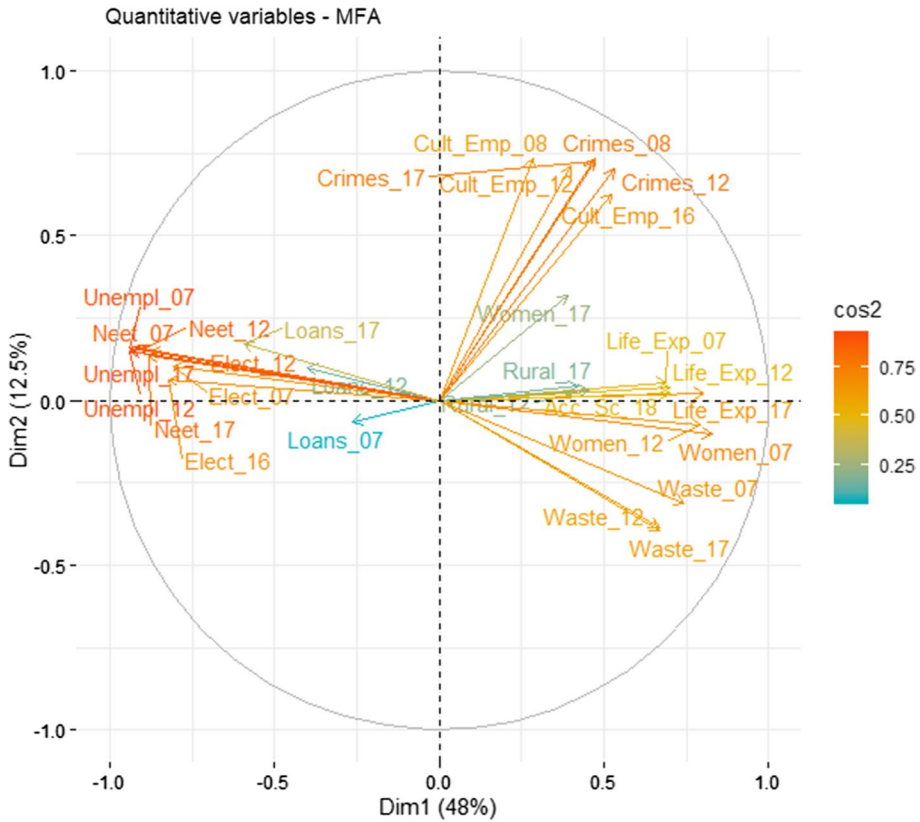
Fig. 8 Correlations between quantitative variables and dimensions. Quality of representation (cos2)

On the other axis (the vertical one), we observe the variables positively correlated with dimension 2 (Crimes, Cult_Emp), which explain 12.5% of the variance, and we can interpret it as a residual one. It is not easy to interpret this dimension: it increases with other crimes reported (Crimes) and with a rise of employers in cultural enterprises.

The graph in Fig. 9 represents the structure of the unit points (provinces) and their positions in 2007, 2012, and 2017. The black point is the barycenter of three colored points: the red one represents 2007, the green one 2012, and the blue one 2017. The dash is the distance between each year and the barycenter. We observe that the movements seem to be prevalent in the horizontal direction. For example, Roma (point 67) shows a considerable change in dimension 2 but not so much in dimension 1 (*socio-economic and environmental* aspects), while Milano (point 17) improves in dimension 1 a little from 2007 to 2012 and more so from 2012 to 2017. Dimension 1 decreases a little in the province of Rimini (point 47) from 2007 to 2012 but more so from 2012 to 2017. Napoli (point 72) improves
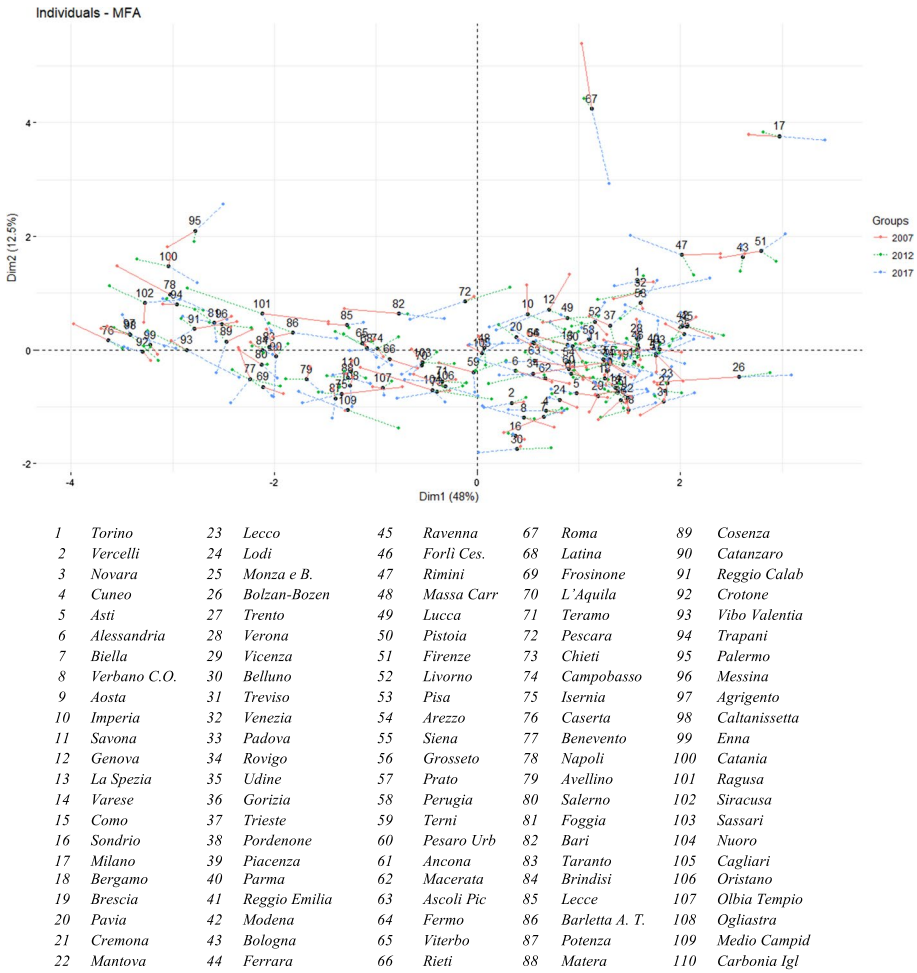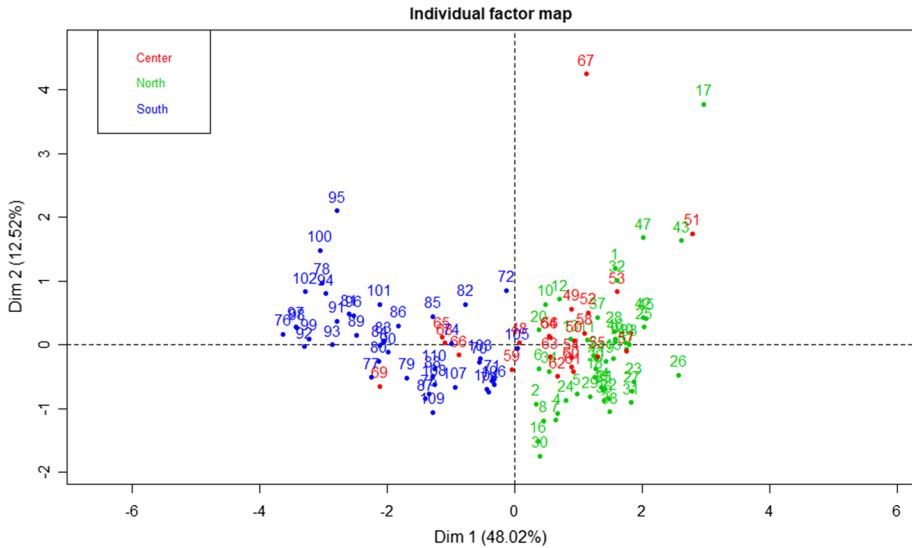
Fig. 9 Individual factor map

1	Torino	23	Lecco	45	Ravenna	67	Roma	89	Cosenza
2	Vercelli	24	Lodi	46	Forlì Ces.	68	Latina	90	Catanzaro
3	Novara	25	Monza e B.	47	Rimini	69	Frosinone	91	Reggio Calab
4	Cuneo	26	Bolzan-Bozen	48	Massa Carr	70	L'Aquila	92	Crotone
5	Asti	27	Trento	49	Lucca	71	Teramo	93	Vibo Valentia
6	Alessandria	28	Verona	50	Pistoia	72	Pescara	94	Trapani
7	Biella	29	Vicenza	51	Firenze	73	Chieti	95	Palermo
8	Verbano C.O.	30	Belluno	52	Livorno	74	Campobasso	96	Messina
9	Aosta	31	Treviso	53	Pisa	75	Isernia	97	Agrigento
10	Imperia	32	Venezia	54	Arezzo	76	Caserta	98	Caltanissetta
11	Savona	33	Padova	55	Siena	77	Benevento	99	Enna
12	Genova	34	Rovigo	56	Grosseto	78	Napoli	100	Catania
13	La Spezia	35	Udine	57	Prato	79	Avellino	101	Ragusa
14	Varese	36	Gorizia	58	Perugia	80	Salerno	102	Siracusa
15	Como	37	Trieste	59	Terni	81	Foggia	103	Sassari
16	Sondrio	38	Pordenone	60	Pesaro Urb	82	Bari	104	Nuoro
17	Milano	39	Piacenza	61	Ancona	83	Taranto	105	Cagliari
18	Bergamo	40	Parma	62	Macerata	84	Brindisi	106	Oristano
19	Brescia	41	Reggio Emilia	63	Ascoli Pic	85	Lecce	107	Olbia Tempio
20	Pavia	42	Modena	64	Fermo	86	Barletta A. T.	108	Ogliastra
21	Cremona	43	Bologna	65	Viterbo	87	Potenza	109	Medio Campid
22	Mantova	44	Ferrara	66	Rieti	88	Matera	110	Carbonia Igl

in dimension 1 from 2007 to 2012, while it worsens from 2012 to 2017. Moreover, the situation in 2017 is worse than that in 2007.

Figure 10 shows one relevant aspect. The points representing the provinces are reported in three different colors depending on the relevant macro region. The graph shows that contiguous provinces are not close in the factorial depiction (see, e.g., the provinces of Lazio indicated by the numbers 65, 66, 67, 68, and 69), and we also see a net separation on the factorial axis between the provinces in the North and those in the South, with the latter showing worse values for dimension 1.

Fig. 10 Individual factor map. Distribution by macro region (North, Center and South Italy)

5 Conclusions

The aim of the research is to understand how and equitable and sustainable well-being in the territories (provinces) has changed, by comparing them between the precrisis 2008 and the subsequent years.

The principal interest is how the various territorial areas have moved over the period and which are the indicators that have a greater weight. For this purpose, the ISTAT database containing the BES territorial indicators is used. We have chosen an indicator for each domain because for some domains, it was the only possible choice, and in other cases, it was not possible to build a composite indicator for each domain for each of the three years considered (2007, 2012, and 2017) because of missing indicators; moreover, the different number of indicators by domain would lead to an initial distortion, resulting in different weights for each domain.

The need to simultaneously introduce quantitative and qualitative variables (known as mixed data) as active elements of one factorial analysis is typical in the problem of variables being partitioned into different subspaces; this can be solved by using MFA. It relates to a PCA, which can analyze both quantitative and qualitative data (i.e., time).

The hypothesis of independence between the variables at different times is satisfied when considering sufficiently distant times (5 years).

The MFA analysis also allows for the use of indicators available in a few years only; this happens for the Social relations domain. The variables considered in the analysis may also be different in the groups identified by the qualitative variable (time). A further characteristic of the method that makes it useful for studying evolution over time is that it is suitable for considering variables that are not available on an annual basis, too.

Our results identify a principal dimension that describes about 48% of the total variability across provinces over time.

By making use of the MFA we have identified the direction in the change and its magnitude with respect to the group of variables that are described with the various dimensions across the considered Italian territorial units.

The new idea regards the application of MFA to the selected indicators for Italian provinces in different periods. MFA is applied to tables in which a set of individuals (one individual = one row) is described by a set of variables (one variable = one column). The advantage of MFA lies in the fact that within the active variables, it can account for a group structure defined by the user. Such data tables are called *individuals × variables organised into groups*.

A different idea could be construct a composite indicator for each domain and then apply the MFA to monitor the evolution of BES in Italy over time.

References

Abdi, H., Williams, L. J., & Valentin, D. (2013). Multiple factor analysis: Principal component analysis for multitable and multiblock data sets. *WIREs Computer Statistics, 2013*(5), 149–179. https://doi.org/10.1002/wics.1246.

Aggarwal, C. C. (2005). On change diagnosis in evolving data streams. *IEEE Transactions on Knowledge and Data Engineering, 17*, 587–600.

Bolasco, S. (1999). *Analisi Multidimensionale dei Dati. Metodi, strategie e criteri d'interpretazione*. Roma: Carrocci editore.

Ciommi, M., Gigliarano, C., Taralli, S., & Chelli, F. M. (2017). The Equitable end Sustainable Well-Being at local level: a first attempt of time series aggregation. *Rivista Italiana di Economia Demografia e Statistica*, *71*(4), 51–60.

Chelli, F. M., Ciommi, M., Emili, A., Gigliarano, C., & Taralli, S. (2016). Assessing the equitable and sustainable well-being of the Italian provinces. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *24*(Suppl. 1), 39–62.

Escofier, B., & Pages, J. (1998). *Analyses Factorielles Simples et Multiples: Objectifs, Methodes et Interpretation* (3rd ed.). Paris: Dunod.

ISTAT. (2013). BES-2013, Il Benessere Equo e Sostenibile in Italia. Retrieved October, 2019, from https://www.istat.it/it/files//2013/03/bes_2013.pdf

ISTAT. (2018). BES-2018, Il Benessere Equo e Sostenibile in Italia. Retrieved October, 2019, from https://www.istat.it/en/archivio/225140.

ISTAT. (2019). Misure del benessere dei territori. Anno 2018. Nota metodologica. Retrieved October, 2019, from https://www.istat.it/it/files//2019/05/Nota-metodologica.pdf

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software., 25*, 1–18.

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.

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.

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.

OECD. (2019). Measuring well-being and progress: Well-being research. Retrieved October, 2019, from https://www.oecd.org/statistics/measuring-well-being-and-progress.htm

Oliveira, M., & Gama, J. (2012). A framework to monitor clusters' evolution applied to economy and finance problems. *Intelligent Data Analysis, 16*, 93–111.

Pages, J., & Husson, F. (2005). Multiple factor analysis with confidence ellipses: a methodology to study the relationships between sensory and instrumental data. *Journal of Chemometrics., 19*, 138–144.

Pages, J. (2014). *Multiple Factor Analysis by Example Using R*. London: Chapman and Hall/CRC.

Spiliopoulou, M., Ntoutsi, I., & Theodoridis, Y. (2009). Tracing cluster transitions for different cluster types. *Control and Cybernetics, 38*, 239–259.

Taralli, S. (2013). Indicatori di Benessere Equo e Sostenibile delle Province: informazioni statistiche a supporto del policy-cycle e della valutazione a livello locale. *RIV Rivista Italiana di Valutazione.*, *55*(2013), 171–187.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.