THE EQUITABLE AND SUSTAINABLE WELL-BEING AT LOCAL LEVEL: A FIRST ATTEMPT OF TIME SERIES AGGREGATION[1]

Mariateresa Ciommi, Chiara Gigliarano, Stefania Taralli, Francesco M. Chelli

1. Introduction

The Istat project aimed at measuring well-being at local level[2] has recently made available time series data on a collection of BES indicators also at NUTS3 level. This recent novelty provides the motivation for studying appropriate methodologies to analyze such complex information. In particular, synthesizing individual indicators into a single index, especially if in a time perspective, by constructing a composite indicator, is an important challenge (Nardo, 2008). Indeed, the choice of the aggregation method is very important and delicate, as the composite indicator is required to be sensitive to territorial disparities, able of capturing transformations over time and, at the same time, robust with respect of the variability and over-dispersion that characterize many BES indicators at NUTS3 level (Taralli *et al*., 2015)

In this paper, we implement a latent variable (factor) model with the aim to construct a well-being composite indicator for the 110 Italian Provinces based on a selection of eight BES indicators over the period of time between year 2004 and year 2015.

Factor models are widely used in the construction of composite indicators. For instance, Chelli et al. (2015) applied a factor analysis to reduce the high number of indicators involved in the definition of the BES at local level. Based on the result of the factor analysis computed among indicators of each domain, the authors then

[1] Disclaimer: This paper is a collective work, realized under the collaboration agreement between Istat and Polytechnic University of Marche for the analysis of the well-being at territorial level. In any case, the opinions expressed do not necessarily reflect those of the Italian National Institute of Statistics.

[2] The data used in this paper were produced by Istat under the project "Equitable and Sustainable Well-being Measures at local level", started as a continuation of the "Provinces' BES" project, that was promoted by CUSPI (Coordination of Statistical Offices of the Italian Provinces) and realized under the Istat's methodological and technical coordination.

constructed a composite indicator as linear combination of the estimated factor scores with weights based on the Gini index of concentration.

The rest of the paper is organized as follow. Section 2 is devoted to the description of the data used in our analysis. Section 3 describes the methodology. Section 4 and Section 5 are devoted to the discussion of the results and conclusions, respectively.

2. Data

This study is based on the data extracted from the 2016 release of Istat's Provinces' BES dataset, which contains time series of all the indicators included in the Provinces' BES project. Following Rijpma (2017) we have chosen one leading indicator for each of the eight well-being domains to be considered in our analysis. Table 1 shows the selected indicators and the related domains.

Table 1 – *The selected measures of equitable and sustainable well-being (Bes indicators).*

Domain	Indicator	Code	Unit
Work	Non-participation rate (15-74)	LAB	Percentage
Education	People with lower secondary education or less (25-64)	EDU	Percentage
Health	Life expectancy at birth	LE	Years
Economic well-being	Household disposable income	INC	Euro
Politics and Institutions	Young people in local government (≤39)	POL	Percentage
Security	Homicide rate	HOM	For 100,000 inhabitants
Quality of services	Separate collection of municipal waste	WAST	Percentage
Environment	Air pollution	PM10	Total days per year

Source: Provinces' BES dataset.

For a detailed description of the indicators involved in our analysis, see the Appendix of this paper.

The dataset contains 110 observations corresponding to the 110 Italian Provinces listed in the NUTS 2013 classification[3] and a complete time series over

[3] The current NUTS 2013 classification is valid from 1 January 2015, and for Italy at the NUTS3 level it includes 110 territorial units, coinciding with the 110 provinces that existed in Italy at the reference date. During 2016, following the reform of the Local Authorities implemented by the Italian Government, some Provinces have become Metropolitan cities, while some others have been suppressed or modified due to regional laws. Since these changes have not yet been transposed into the statistical classification, in this paper, the term Provinces refers to the 110 units accounted in NUTS3, so including the new Metropolitan cities and the Provinces that no longer exist.

the years between 2004 and 2015 for each of the eight indicators. Missing data are about 2.5% of total observations. This percentage is mostly due to the new Provinces that were established in the year 2010.

Table 2 – *Summary statistics (2004-2015)*

Indicator	Mean	Standard Deviation	Min	Max
Per capita GDP	24339	6419	12800	51200
Non-participation rate (15-74)	17.89	10.83	3.70	48.45
People with lower secondary education or less (25-64)	47.00	7.73	27.30	69.34
Life expectancy at birth	81.58	0.92	78.40	84.10
Household disposable income	39016	6909	25132	59609
Young people in local government (≤39)	30.85	5.28	5.00	58.00
Homicide rate	0.99	2.56	0.00	83.14
Separate collection of municipal waste	34.18	20.14	0.00	87.40
Air pollution	58.05	45.93	-1	321.00

Source: Our elaboration on Eurostat database and Provinces' BES dataset.

A further indicator was included in the analysis, which is the Per capita Gross domestic product (PCGDP) at current market prices, measured in Purchasing Power Standard (PPS), provided by the Eurostat's Regional economic accounts[4].

Table 2 reports summary statistics of the indicators considered over 2004-2015 period, while Figure 1 shows their trends.

Figure 1 clearly shows that the series have very different trends, which reflect the divergences among the territorial units, and the different nature of the phenomena.

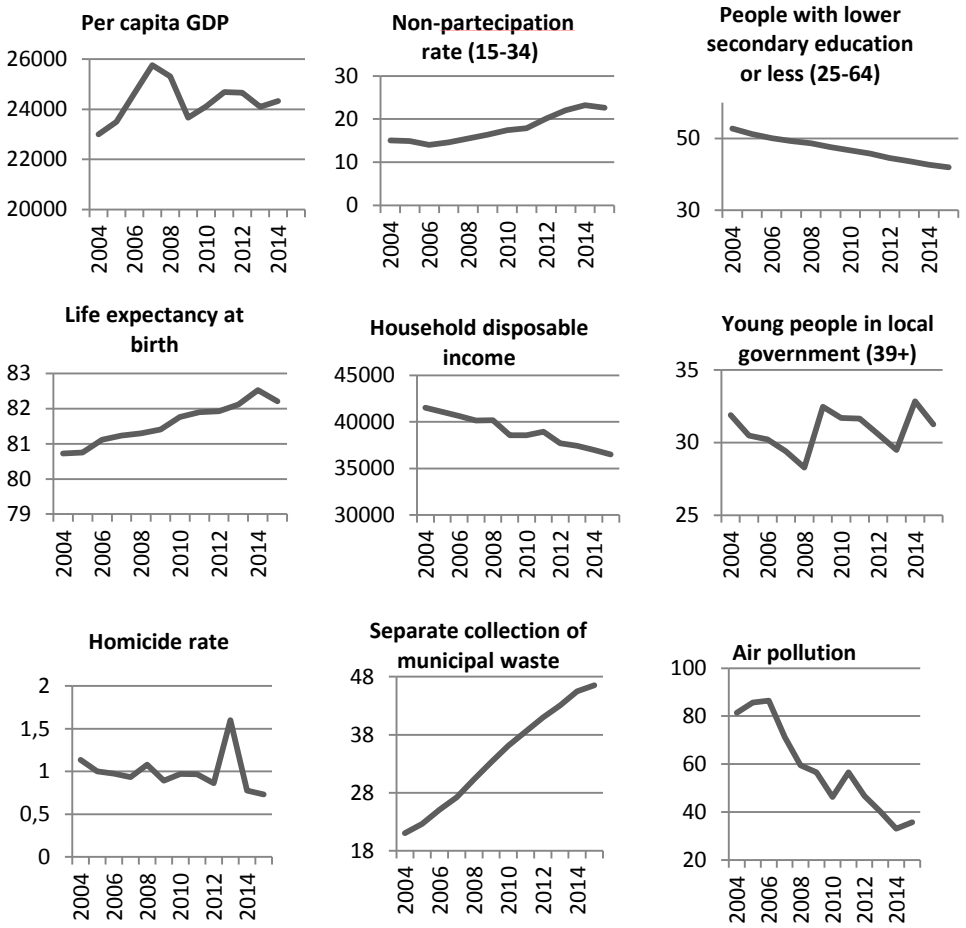
The indicator with the highest temporal variability, computed through the coefficient of variation (data not shown), is Homicide rate (CV=2.6), followed by Air pollution (CV=0.8), Separate collection of municipal waste (CV=0.6), and Non-participation rate (CV=0.6), while the remaining indicators show a CV below the threshold of 0.2.

Looking at the trends, we note a relevant improvement over time in the values of Separate collection of municipal waste, which passes from 21.0% in 2004 to 46.5% in the year 2015 (an overall increase of 121.4%). Time-variation is smaller, but still substantial for Air pollution, which shows an overall decrease of 55.9% and for Non-participation rate, with an increase over time of 49.7%. Generational replacement contributes to an overall reduction of 20.5 per cent points for the indicator People in working-age with lower secondary education or less. In the

[4]Available at: http://ec.europa.eu/eurostat/web/regions/data/database

same period the Household disposable income decreases by 12.1%, while the Per capita GDP increases by 9.1%.

Figure 1 – *Trends of the eight BES indicators and of the Per capita GDP - average of NUTS3 data (2004-2015).*



Source: Our elaboration on Eurostat database and Provinces' BES dataset

Notice the outlier value of the Homicide rate in correspondence to the year 2013, due to the large number of migrants who died in a shipwreck near the Lampedusa Island and were considered as victims of a crime of massacre. Overall, the global Homicide rate marks a steady and consistent decrease (by 35.4%) over the period 2004-2015, while Life expectancy at birth has a low rate steady growth,

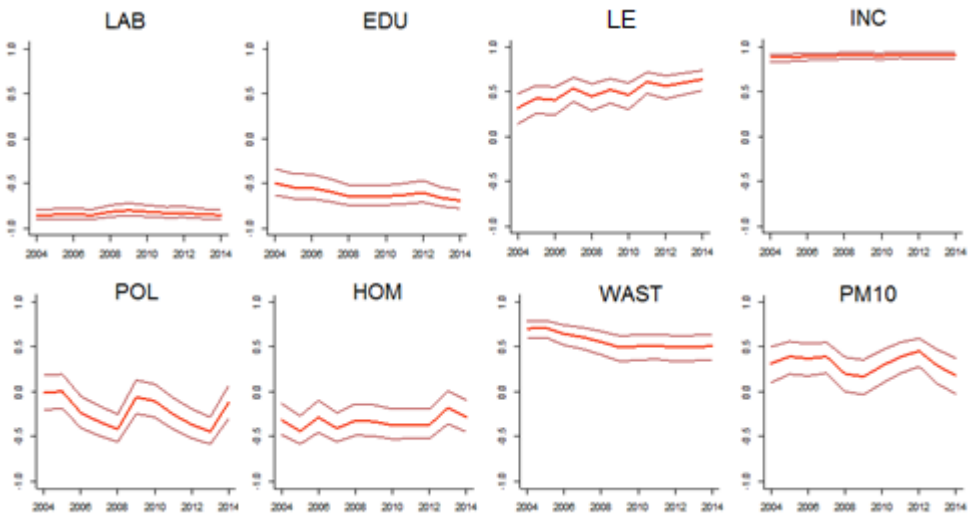
except for the 2015 breakdown due to exceptional number of deaths in that year in Italy. The percentage of young people in local government shows a cyclical trend, with high peaks at the years of municipal elections. The underlying trend is moderate growth (from 32% in 2004 to 33% in 2015).

Table 3 – *Correlations among BES indicators (year 2015)*

20015	LAB	EDU	LE	INC	POL	HOM	WAST	PM10
LAB	1							
EDU	0.61	1						
LE	-0.58	-0.45	1					
INC	-0.70	-0.51	0.63	1				
POL	0.27	0.15	0.12	-0.12	1			
HOM	0.32	0.14	-0.35	-0.32	0.05	1		
WAST	-0.56	-0.27	0.32	0.45	-0.20	-0.13	1	
PM10	0.10	0.04	-0.13	-0.05	-0.04	0.08	0.24	1

Source: Our elaboration on Eurostat database and Provinces' BES dataset

Figure 2 – *Cross-sectional correlation coefficients and 95 per cent confidence intervals between BES indicators and PCGDP.*



Source: Our elaboration on Eurostat database and Provinces' BES dataset

Looking now at the correlations among the eight BES indicators (see Table 3) we note that the strongest positive correlations are between the domains Work and

Education (with correlation coefficient equal to 0.49 in 2014 and to 0.61 in 2015), while the domains Work and Quality of services show the strongest negative correlation coefficient both in 2004 (-0.75) and in 2015 (-0.56).

The correlation coefficients between each of the BES indicators and the Per capita GDP tend to remain constant over time, with 95% confidence intervals that are rather tight. The strongest correlations are between PCGDP and Household disposable income, as expected (it ranges from 0.85 in 2004 to 0.90 in 2013) (Figure 2).

3. Methodology

In the literature, several methods have been proposed for constructing composite indicators; for a review we refer, among others, to Nardo et al (2008).

Here we adopt a data-driven approach that implies that the aggregation step is essentially based on objective, rather than subjective, criteria.

To the best of our knowledge, this is the first attempt to construct a well-being composite indicator using data for Italian Provinces in a time series perspective (2004-2015). Here we estimate a composite indicator through a latent variable (factor) model (hereafter, LVM), following the methodology introduced in Rijpma (2016).

The main advantage of this model is that it allows for imputation of missing data. This is an important issue, especially when working with time series data and even more at territorial level. In fact, on one hand, with historical data missing values are very frequent, and in particular at a high territorial detail time-series can be affected by administrative change, but, on the other hand, the way in which imputation is done influences the results. According to LMV, data are imputed using the covariance among indicators.

Another advantage of LVM refers to the weighing procedure that is a necessary step to construct a composite indicator, based on multiple series. The LVM assigns higher weights to highly correlated indicators and, as stressed by Foster et al. (2013), this procedure increases the robustness of the rankings.

However, the method is not without disadvantages. Firstly, the know that the correlation between variables does not necessarily capture the right trade-offs between the indicators. In addition, the model assumes that the underlying well-being concept is reflected in the correlation structure, and not that each indicator measures a distinct and unique part of well-being. Concerning this, it has been argued that applying a reflective measurement model to Bes is a misspecification of the measurement model itself as the Bes is a formative construct (Istat, 2015). However, in the Bes the multicollinearity, which is a highly undesirable property in

formative models, is not fully solved, as several indicators show substantial correlations. Finally, the model assumes that the observed and latent variables are continuous and this is a further limitation.

We consider a factor model for the observed data y_{ij}, where *i* is the province and *j* the indicator. The model is defined as

$$y_{ij} \sim N\left(\beta_{0j} + \beta_{1j}\chi_i, \omega_j^2\right) \quad \text{and} \quad \chi_i \sim N(0,1)$$

where *N* denotes the normal distribution, χ_i is the unobserved and latent variable for the province i, β_{1j} is a parameter reflecting how well the observed indicator *j* differentiates between units (Provinces), β_{0j} is an intercept and ω_j^2 represents the variance of indicator *j*. The latent variable (which corresponds to the well-being composite indicator) is assumed to be standardized, with mean equal to zero and standard deviation equal to one. This is not a problem, since the composite indicator has no natural unit of measurement.

The model is estimated in a Bayesian framework. This approach presents several advantages, since it works very well with the multilevel structure of our data and, more important, it may account for similarities among provinces in the same region. Bayesian multilevel models take this information into account through the priors. Rather than considering the same standard normal distribution for all provinces ($\chi_i \sim N(0,1)$), we may follow Høyland et al. (2012) and allow as a prior a distribution of the composite indicator that varies across the *m* macroareas (or macro-regions); formally

$$\chi_i \sim N\left(\alpha_{m[i]}, 1\right) \quad \text{and} \quad \alpha_{m[i]} \sim N(0,1).$$

The use of priors might introduce a degree of arbitrariness and subjectivity. In fact, prior distributions usually reflect the researcher's beliefs about the model, before looking at the distribution of the real data. Thus, to reduce this subjective belief we use uninformative priors, assuming that: (i) the loadings follow a normal distribution, $N(0, 10^{-7})$, (ii) the variance is modelled through a Gamma distribution, $G(0.01, 0.01)$, and (iii) the group-specific error term follows a normal distribution, $N(0, \sigma^2)$ with variance uniformly, $\sigma^2 \sim U(0,100)$.

The model is estimated through Gibbs sampling and implemented in JAGS[5].

The Bayesian multilevel latent variable model has two main advantages: first of all, the weights for the aggregation procedure, which are based on correlations, are chosen so that they differentiate between provinces as best as possible; moreover,

[5] JAGS is a program for analysis of Bayesian hierarchical models using Markov Chain Monte Carlo (MCMC) simulation. See Plummer (2003) for more details. We are grateful to Rijpma who kindly provided us with the code.

the model accounts for sources of uncertainty, such as measurement error in a composite indicator or missing data.

4. Results

As a first result, we are interested in analysing the distributions of the factor loadings, synthesized in Table 4, which show how much each BES indicator contributes to the composite indicator.

Table 4– *Summary of the factor loadings' distribution for each BES indicator*

Indicator	Mean	5-th percentile	50-th percentile	95-th percentile
Non-participation rate (15-74)	-0.967	-0.981	-0.967	-0.952
People with lower secondary education or less (25-64)	0.829	0.801	0.829	0.856
Life expectancy at birth	0.282	0.235	0.282	0.327
Household disposable income	-0.445	-0.488	-0.445	-0.402
Young people in local government (39+)	0.331	0.276	0.332	0.385
Homicide rate	-0.286	-0.331	-0.286	-0.240
Separate collection of municipal waste	-0.225	-0.272	-0.225	-0.179
Air pollution	0.470	0.430	0.470	0.509

Source: Our elaboration on Provinces' BES dataset

The first two indicators listed in the table are the main contributors to the composite indicator. In particular, the Non-participation rate and the People with lower secondary education or less are the indicators that affect more the well-being composite indicator, the former with a strongly negative effect and the latter with a strong positive impact. Third, in decreasing order of importance, is the Air pollution, with a mean loading of 0.470, followed by the negative loading of the Household disposable income (-0.445 on average).

Looking at the temporal trend of the composite indicators by macro-area (Figure 3) Italy seems to be divided into two areas of well-being: on the one hand, the South and the Islands, with consistently negative composite indicator's values (that is below the Italian average), indicating a structural and persistent disadvantage; on the other hand, the North and the Centre, always taking positive values (above the Italian average).

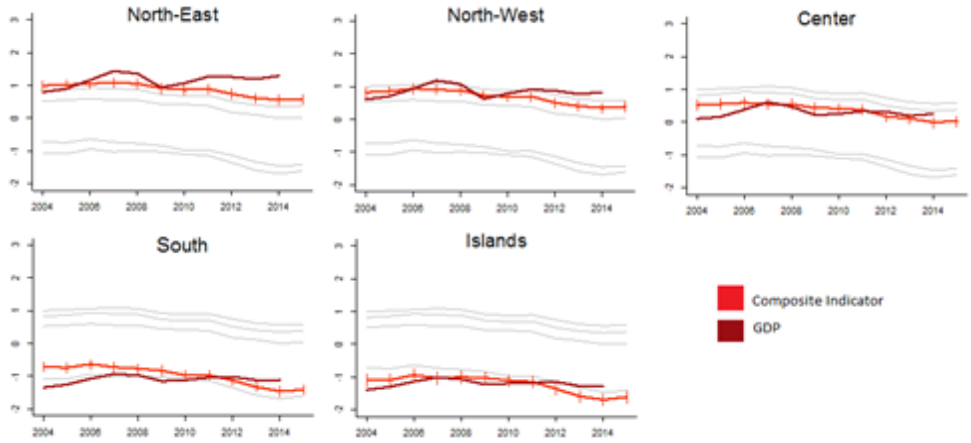
The highest levels of the composite index (red line) are registered, in particular, in the North-East of Italy (NE). The average value of the composite indicator of NE, indeed, remains above the ones of all the other Italian macro-areas (grey lines) during the whole period. In particular, until year 2008 the value of the composite indicator for the NE stays around an average of 1.0-1.1. Starting from 2008 a

slightly decreasing trend leads the BES indicator for NE to levels between 0.4 and 0.6 in the years 2014-15 (with an overall reduction of 44%).

The composite index for the North-West (NW) shows for each year values that are slightly lower than NE but with very similar trends: it remains stable around average values of 0.8-0.9 until 2008, followed by a gradual decrease, more pronounced than NE, from year 2008 to 2015 (with an overall decrease of 56.2%).

The composite indicator for the Center of Italy is fairly stable around an average value of 0.56 until the year 2008. Starting from 2008 it sharply declines, reaching the value of 0.03 in 2015, which corresponds to an overall decrease of 95% compared to 2008. It seems, therefore, that the 2008 economic crisis has affected the well-being of this area much harder than in the North of Italy. The loss for the Center, in relative terms, is similar to the one suffered by the Islands (-61.5%), which were already in a more disadvantageous condition. However, South is the most penalized and deprived macro-area in Italy: in 2014-2015 the well-being level in the South of Italy lies between -1.25 and -1.62; this corresponds to a negative variation, compared to year 2008, equal to -85%.

Figure 3 – *Standardized Per capita GDP, estimates and 90% Confidence Interval of the BES composite indicator, by macro-area (years 2004-2015).*

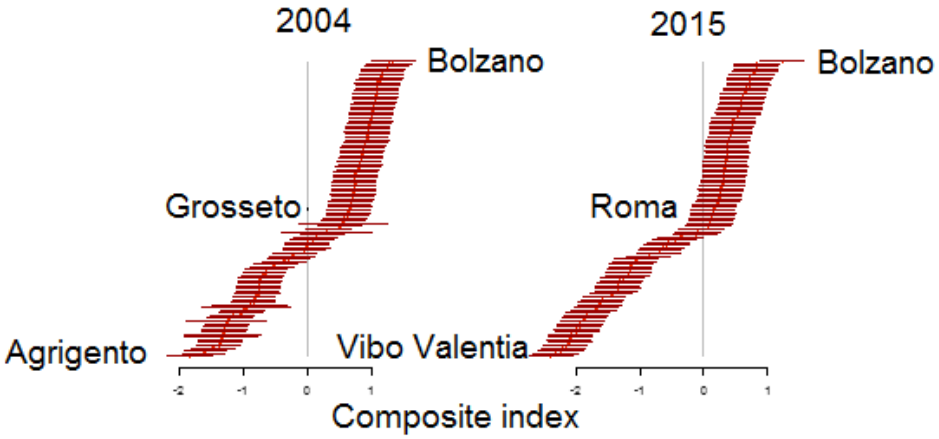


Source: Our elaboration on Eurostat database and Provinces' BES dataset

Figure 3 also compares the trends of the BES composite indicator (red line) and of the Per capita GDP (dark red line) for each macro-area. It can be noted that the two series are generally rather close. In NE, starting from the year 2009, there is a clear-cut separation between PCGDP, which tends to grow, and BES, which shows a decreasing trend. At different points in the time-line, but still after 2008, this reversal of the relationship between BES and PCGDP occurs in all areas of Italy.

Figure 4 depicts the estimates of the composite indicator for all Italian Province in 2004 and 2015. For each Province, the graphic displays the mean value of the composite indicator and its confidence interval at 90%. Looking at the value of the composite index we can evaluate, for each period, the trend and position of the BES levels for each province, with respect to the general mean (grey vertical bar, equal to 0). The form of the right-hand curve compared to the left-hand shows that over the last few years the distances between best and worst performers deepened both in general and compared to the average. It emerges that territorial disparities in 2005 affect mainly the lower part of the distribution, which corresponds to those provinces that were already disadvantaged in the years before the crisis. Bolzano shows the best performance, whereas, although with some exceptions, the South and the Island Provinces are ranked in the last positions, both in 2004 and 2005.

Figure 4 – *Composite indicator and its (90%) confidence interval for 2004 and 2015 for all Provinces.*



Source: Our elaboration on Eurostat database and Provinces' BES dataset

5. Concluding Remarks

A latent variable model has been adopted in order to construct a composite indicator aimed at representing the trend of well-being for the Italian Provinces over the years 2004-2015. The composite indicator proposed may be considered a useful tool for summarizing the well-being trends, since it allows enhancing differences among macro-areas and their trends also during the years of the recent financial crisis.

Future research will consider (i) including additional well-being indicators, in order to cover all the well-being dimensions and (ii) aggregating the analysis also at regional level (and not only at macro-regional level) to provide a more in-depth analysis of territorial disparities, inequalities and divergences across the Italian territories.

Appendix

Here we provide a detailed definition of the BES indicators selected for our analysis.

- Non-participation rate (15-74) is defined as the ratio: (*Unemployed persons aged 15-74 years plus inactive persons - who have not looked for a job in the past 4 weeks but willing to work) / (Labour force aged 15-74 years plus inactive persons).*
- People with lower secondary education or less (25-64): *Percentage of people aged 25-64 years who have completed at most lower secondary education (ISCED level not above 3a, 3b or 3c).*
- Life expectancy at birth: *Average number of years that a child born in a given calendar year can expect to live if exposed during his whole life to the risks of death observed in the same year at different ages.*
- Household disposable income: *Ratio of the annual household disposable income over the total number of household members.*
- Young people in local government (≤39): *Percentage of young people (aged 39 years or less) elected or in charge in municipal councils over the total number of people elected or in charge.*
- Homicide rate: *Number of homicides per 100 thousand inhabitants.*
- Separate collection of municipal waste: *Percentage of municipal waste collected separately over the total collection of municipal waste.*
- Air pollution: *Number of days in a year during which the level of PM10 exceeds the limit of 50g/m3.*

References

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*, Vol. *24*(1), pp. 39-62.

FOSTER J.E., MCGILLIVRAY M., SETH S. 2013. Composite Indices: Rank Robustness, Statistical Association, and Redundancy, *Econometric Reviews*, Vol. 32, pp. 35-56.

HØYLAND, B., MOENE K., WILLUMSEN F. 2012, The Tyranny of International Index Rankings, *Journal of Development Economics*, Vol. 97, pp 1-14.

ISTAT (2015). Il Benessere Equo e sostenibile in Italia, Roma, Istat.

NARDO M., SAISANA M., SALTELLI A., TARANTOLA S. (2008).Handbook on constructing composite indicators: Methodology and user guide. OECD Publishing.

PLUMMER, M., 2003, "JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling," DSC, (Disability Services Commissioner) 2003 Working Papers, www.r-project.org/nosvn/ conferences/DSC-2003/Drafts/Plummer.pdf.

RIJPMA A. 2016. What can't money buy? Wellbeing and GDP since 1820. *Working paper* No. 0078. http://www.cgeh.nl/working-paper-series/

TARALLI S., CAPOGROSSI C., PERRI, G. 2015. Measuring Equitable and Sustainable Well-being (Bes) for policy-making at local level (NUTS3), RIEDS, vol. LXIX n. 3, pp. 97-106.

SUMMARY

The Equitable and Sustainable Well-Being at local level: a first attempt of time series aggregation

The on-going economic crisis has reinforced the feeling that macroeconomic indicators, namely economic indicators at country level, do not provide a correct picture of the living conditions in a territory. In fact, individual and local characteristics also influence the well-being of individuals and, within the same country, territories can vary at a large extent. Thus, the analysis of well-being at local level is crucial.

Here, we analyse the time series of a selection of indicators that constitute the Equitable and Sustainable Well-being (BES) at local level (NUTS3). After providing an overview of the temporal trends in the selected well-being indicators, we construct a composite index of well-being for groups of regions (NUTS1) applying a latent variable model estimated in a Bayesian framework.

Mariateresa CIOMMI, Università Politecnica delle Marche, m.ciommi@univpm.it
Chiara GIGLIARANO, Università dell'Insubria, chiara.gigliarano@uninsubria.it
Stefania TARALLI, Istat, taralli@istat.it
Francesco M. CHELLI, Università Politecnica delle Marche, f.chelli@univpm.it