VOLUME LXXV – N. 1 GENNAIO – MARZO 2021

# RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA

**DIRETTORE**
CHIARA GIGLIARANO

**GUEST EDITOR**
LIVIA CELARDO, MARIATERESA CIOMMI

**COMITATO SCIENTIFICO**
GIORGIO ALLEVA, GIAN CARLO BLANGIARDO, LUIGI DI COMITE, MAURO GALLEGATI, GIOVANNI MARIA GIORGI, ALBERTO QUADRIO CURZIO, CLAUDIO QUINTANO, SILVANA SCHIFINI D'ANDREA

**COMITATO DI DIREZIONE**
CHIARA GIGLIARANO, CLAUDIO CECCARELLI, PIERPAOLO D'URSO, SALVATORE STROZZA, ROBERTO ZELLI

**REDAZIONE**
LIVIA CELARDO, MARIATERESA CIOMMI, ANDREA CUTILLO, GIUSEPPE GABRIELLI, ALESSIO GUANDALINI, SIMONA PACE, GIUSEPPE RICCIARDO LAMONICA, ANDREA SPIZZICHINO

*Sede Legale:* C/O Studio Associato Cadoni, Via Ravenna n. 34 – 00161 ROMA.
info@sieds.it, rivista@sieds.it

# SIEDS
# SOCIETÀ ITALIANA
# DI ECONOMIA DEMOGRAFIA E STATISTICA


**CONSIGLIO DIRETTIVO**

*Presidenti Onorari:* LUIGI DI COMITE, GIOVANNI MARIA GIORGI,
FRANCESCO MARIA CHELLI

*Presidente:* SALVATORE STROZZA

*Vice Presidenti:* LEONARDO BECCHETTI, CLAUDIO CECCARELLI,
VENERA TOMASELLI

*Segretario Generale:* MATTEO MAZZIOTTA

*Consiglieri:* MARCO ALFÒ, GIUSEPPE GABRIELLI, CHIARA GIGLIARANO, LUCIANO NIEDDU,
SIMONE POLI, MARIA CRISTINA RECCHIONI, STEFANIA RIMOLDI, SILVANA MARIA ROBONE

*Segretario Amministrativo:* ALESSIO GUANDALINI

*Revisori dei conti:* MICHELE CAMISASCA, FABIO FIORINI, DOMENICO SUMMO

*Revisori dei conti supplenti:* MARGHERITA GEROLIMETTO, GIUSEPPE NOTARSTEFANO



**SEDE LEGALE:**
C/O Studio Associato Cadoni, Via Ravenna n. 34 – 00161 ROMA

info@sieds.it

rivista@sieds.it

---

VOLUME FUORI COMMERCIO – DISTRIBUITO GRATUITAMENTE AI SOCI

## INDICE

# LOCAL INEQUALITY ANALYSIS IN THE US: EVIDENCE FROM SOME METROPOLITAN STATISTICAL AREAS

Margherita Gerolimetto, Stefano Magrini

/

## 1. Introduction

Understanding the impact of policy changes on the distribution of income first requires a good representation of the distribution. There are various ways to do this that range from simple approaches, as the calculus of inequality measures, to more sophisticated approaches, as the kernel estimation of the distribution and, if possible, the observation of its evolution over time. All these methods can be jointly employed to have a clearer view of the concentration of the income, to compare different distributions and to understand the impact of difference policy actions.

In this work, we are interested in the analysis of income inequalities in the US between 2010 and 2018. In particular, we concentrate on a local perspective, by studying the inequality recorded in seven Metropolitan Statistical Areas, which have been chosen in order to cover geographically the US territory. It is important to emphasize that evaluating the degree of inequality at a local level, such a city or a metropolitan statistical area, is as important as at the national level. The connection, for instance, between inequality and crime is as strong within urban areas as it is across countries. Moreover, urban inequality seems as likely to generate political uprisings as inequality across large geographic units (Glaeser *et al*, 2009).

The aim of this work is two-fold. On the one hand, we study inequality, measured via some of the most well-known inequality indices. On the other hand, we complete the analysis via the kernel estimate of the distribution and some distribution dynamics analysis. The idea is that to have a completely informative inequality analysis, different perspectives should be considered.

The structure of the paper is as follows. In section 2, we present an overview of the main inequality measures. In section 3, we recall two tests to compare distributions. In section 4, we present our empirical analysis on 7 Metropolitan Statistical Areas across the USA. In section 5, we present some conclusions.

## 2. Inequality measures: an overview

Measures of inequality are widely used to study income and welfare. They are often a function that assigns a value to a specific distribution of income so that direct and objective comparisons across different distributions are possible: i) dynamic comparison (i.e. comparing inequality measures across time) and ii) comparisons for policy analysis (i.e. comparing the redistributive effects of current tax policy).

To do this, inequality measures should have certain properties and behave in a certain way, given certain events. No single measure satisfies all the properties, so the best approach is to look at more than one measure, trying to capture all the different perspectives into which the phenomenon is observed. In this overview, we focus on indices and ratios.

The Gini index (Gini, 1912) is the most widely cited measure of inequality; it measures the extent to which the income distribution within an economy deviates from a perfectly equal distribution. In its most intuitive definition, the Gini index is calculated as the ratio of the area between the Lorenz curve and the 45-degree line to the area underneath the 45-degree line. The larger is the index the higher the level of inequality. Being scale invariant, the Gini index allows for direct comparison between two populations regardless of their size. Among its limitations, one is that it is not easily decomposable or additive and it does not respond in the same way to income transfers between people in opposite tails of the distribution as it does to transfers between people in the middle of the distribution. Moreover, very different distribution can be characterized by the same value of the Gini index.

The Atkinson's inequality measure (Atkinson, 1970) is known for being a welfare-based measure of inequality. It presents the percentage of total income that a given society would have to forego in order to have more equal shares of income between its citizens. The index depends on the degree of risk aversion to inequality that characterizes a society. An important feature of the index is in that it can be decomposed into within-group and between-group inequality. Furthermore, it can provide welfare implications of alternative policy options, thus allowing the researcher to possibly include some normative content in the analysis.

The Theil (Theil, 1967) index is a special case of the General Entropy index. It ranges between zero (perfect equality) to one, if normalized. A key feature of these measures is that they are fully decomposable, i.e. inequality can be broken down by population groups or income sources or using other dimensions, which is very useful for policy makers. Another peculiarity of this index is that its mathematical expression depends on the value of a parameter, $\alpha$, that represents a weight to distances between incomes in different parts of the income distribution. For low values of $\alpha$, the index is more sensitive to changes in the lower tail of the distribution while for higher values it becomes more sensitive to changes in the upper tail of the distribution. The most commonly adopted values for α are 0, 1, 2. When α=0 the index is called "Theil's L", when α=1 the index is called "Theil's T", or more simply Theil index, when α=2, the index is called "coefficient of variation". Similarly, to the Gini index, when income redistribution occurs, change in the indices depends on the level of individual incomes involved in the redistribution and the population size.

Finally, we present some ratios that represent a basic inequality measure. In particular, we focus on decile dispersion ratios, which express the income of the richest as a multiple of the income of the poorest. They are simple, direct and easy to understand. At the same time, however, they do not provide as much information as the indexes listed before. A very commonly reported decile ratio is the D9/D1: the ratio of the income of the 10 percent richest part of the population to the income of the 10 per cent poorest. Another frequently adopted ratio is the 20/20 ratio that compares the ratio of the average income of the richest 20 per cent of the population to the average income of the poorest 20 per cent of the population.

## 3. Inference for comparing distributions

In order to focus on the entire distribution, we also resort on the representation of the empirical distribution, via the kernel density estimation (see for example Silverman (1986) for a very good presentation). In particular, we concentrate on the adaptive kernel density estimation based on the nearest neighbors' approach. As typical feature of the adaptive kernel, the smoothing parameter (bandwidth) employed in the estimate is not constant but instead varies according to the degree of clustering of the data. This allows a less biased estimate, while reaching a smoother graphical representation.

Moreover, in order not to confine ourselves to the pure graphical inspection, we consider to two well-known tests to compare distributions: the Kolmogorov-Smirnov test and Kramér-Von Mises test. The Kolmogorov-Smirnov (Kolmogorov, 1933; Smirnov, 1948) test is a nonparametric goodness-of-fit test to assess whether one random sample obtained from a population can be generated by a certain distribution function, that must be specific and known. The Kolmogorov-Smirnoff (KS hereafter) test may also be used to test whether two underlying one-dimensional probability distributions differ, as in the analysis carried out in this paper. The KS test uses the maximal absolute difference between these curves as its test statistic, that for space reasons we do not present here[1]. An attractive feature of this test is that the distribution of the KS test statistic itself does not depend on the underlying cumulative distribution function being tested.

[1] Interested readers may refer to the original papers instead.

Several goodness-of-fit tests, such as Kramér-Von Mises test (Kramér, 1928; von Mises, 1928), KvM hereafter, are refinements of the KS test. As these refined tests are generally considered to be more powerful than the original KS test, many analysts prefer them. In addition, the independence of the KS critical values of the underlying distribution is not as much of an advantage as it first appears. This is due to the fact that the distribution parameters are typically not known *a priori* and have to be estimated from the data. So, in practice the critical values for the KS test have to be determined by simulation just as for the Kramér-Von Mises (and related) tests.

## 4. Empirical analysis

We now present our empirical analysis. Data come from the IPUMS-USA database (Ruggles *et al*., 2020) and consists of per capita income net of transfers, relatively to year 2010 and year 2018 for the following 7 Metropolitan Statistical Areas (MSAs): Chicago-Naperville-Elgin, IL-IN-WI; Dallas-Fort Worth-Arlington, TX; Los Angeles-Long Beach-Anaheim, CA; Miami-Fort Lauderdale-West Palm Beach, FL; Minneapolis-St. Paul-Bloomington, MN-WI; Phoenix-Mesa-Scottsdale, AZ.

The principle behind the concept of Metropolitan Statistical Area (MSA) is that of a core area containing a substantial population nucleus, together with adjacent communities having a high degree of economic and social integration with that core. So, the area defined by the MSA is typically characterized by significant social and economic interaction. As of September 2018, (OMB Bulletin, 2018) there are 392 regions that meet the requirements to be designated as MSA in the U.S. and Puerto Rico (384 in the United States and 8 in Puerto Rico). Currently delineated metropolitan statistical areas are based on application of 2010 standards (which appeared in the Federal Register on June 28, 2010) to 2010 Census and 2011-2015 American Community Survey data, as well as 2018 Population Estimates Program data. The Bureau of Labor Statistics (BLS) uses MSA data to analyze labor market conditions within a geographical area. Within a metropolitan statistical area, workers can presumably change jobs without having to move to a new location, creating a relatively stable labor force. Consequently, MSAs offer a more representative view of the income variable behavior. At the same time, the disadvantage of adopting these geographical units is that they do not correspond to natural political units which make them awkward units for analyzing or discussing public policy.

We begin our empirical analysis by the calculus of the inequality measures we presented in the second section. These results are reported in Table 1. For all 7 MSAs we calculate the 4 inequality indexes and 2 ratios, for 2010 and 2018. We also computed the percentage variation of the indexes. As expected, all indexes and ratios are very well aligned. Apart from a couple of exceptions, all the considered inequality measures increase over the 8-year time span under examination. For two MSAs, in particular, this increase appears to be very severe and is emphasized in bold in the Table: they are Dallas-Fort Worth-Arlington, TX and Miami-Fort Lauderdale-West Palm Beach, FL. A third case captured our attention: it is Minneapolis-St. Paul-Bloomington, MN-WI. Contrary to the previous two mentioned MSAs, this area instead shows an increase in the inequality measures that is not so severe. Interpreting this as a virtuous behavior, we recall that Minneapolis-St. Paul-Bloomington is an area characterized by a special form of local governance, where there is a high degree of overlap between the administrative delimitation of the area and its economic delimitation.

**Table 1 –** *Inequality indices and ratios in 2010, 2018 and percentage variation (in italic)*

| MSA name | Year | Gini | Theil L | Theil T | Atkinson | S80S20 | P90P10 |
|---|---|---|---|---|---|---|---|
| | 2010 | 0.483 | 0.491 | 0.432 | 0.388 | 16.450 | 12.014 |
| Chicago | 2018 | 0.497 | 0.516 | 0.467 | 0.403 | 17.236 | 13.368 |
| | Δ% | *2.903* | *5.115* | *8.069* | *3.913* | *4.778* | *11.268* |
| | 2010 | 0.475 | 0.466 | 0.414 | 0.373 | 15.262 | 10.341 |
| Dallas | 2018 | 0.501 | 0.517 | 0.473 | 0.404 | 16.175 | 11.818 |
| | Δ% | **5.327** | **10.946** | **14.120** | **8.376** | **5.980** | **14.285** |
| | 2010 | 0.502 | 0.512 | 0.464 | 0.401 | 16.565 | 11.889 |
| Los Angeles | 2018 | 0.518 | 0.539 | 0.510 | 0.417 | 18.090 | 13.000 |
| | Δ% | *3.188* | *5.342* | *9.977* | *4.036* | *9.202* | *9.346* |
| | 2010 | 0.490 | 0.473 | 0.454 | 0.377 | 14.462 | 10.021 |
| Miami | 2018 | 0.517 | 0.521 | 0.523 | 0.406 | 16.566 | 11.538 |
| | Δ% | **5.636** | **10.186** | **15.258** | **7.777** | **14.552** | **15.142** |
| | 2010 | 0.450 | 0.437 | 0.374 | 0.354 | 14.156 | 10.407 |
| Minneapolis | 2018 | 0.460 | 0.449 | 0.402 | 0.362 | 14.043 | 10.909 |
| | Δ% | **2.202** | **2.753** | **7.588** | **2.182** | **-0.799** | **4.823** |
| | 2010 | 0.507 | 0.528 | 0.489 | 0.410 | 18.194 | 12.900 |
| New York | 2018 | 0.516 | 0.547 | 0.508 | 0.421 | 19.466 | 12.939 |
| | Δ% | *1.728* | *3.513* | *3.914* | *2.643* | *6.989* | *0.304* |
| | 2010 | 0.457 | 0.433 | 0.380 | 0.351 | 13.593 | 10.323 |
| Phoenix | 2018 | 0.466 | 0.444 | 0.404 | 0.358 | 13.014 | 10.934 |
| | Δ% | *1.925* | *2.617* | *6.240* | *2.080* | *-4.255* | *5.927* |

The graphical representation of the distributions of the logarithm of income and the results of the KS and KvM tests are reported in Figures 1-7. The estimates of the distributions are obtained with a nearest neighbors Gaussian kernel, where the percentage of neighbors is set equal to 25%. Although the 2010 and 2018 distribution representations seem to overlap, in line with the increase in inequality indexes documented in Table 1 both tests lead always to a rejection of the null hypothesis that the distribution is the same. Consistently with those results, we note that the

MSA for which the rejection is less strong is Minneapolis-St. Paul-Bloomington, for which, as previously seen, the increase in inequality measures is generally smaller.

## 5. Concluding remarks

In this paper, we study the evolution of per capita personal income inequalities within selected urban areas of the USA between 2010 and 2018. In particular, we adopt the Metropolitan Statistical Area as the basic spatial unit of analysis as this is an urban region characterized by a significant degree of social and economic interaction.

We first calculate several well-known income inequality indexes for 7 large MSA distributed around the US territory and document a significant increase in income disparities over the 2010-2018 period. Then, for each MSA we produce kernel density estimates of the distributions and performed Kolmogorov-Smirnoff and Kramér-Von Mises tests to evaluate whether the distributions are the same. In all cases, we reject the null hypothesis and conclude that the 2010-2018 period is characterized by a significant increase in income inequalities.

**Figure 1** – *Kernel Representation of Income Distributions*

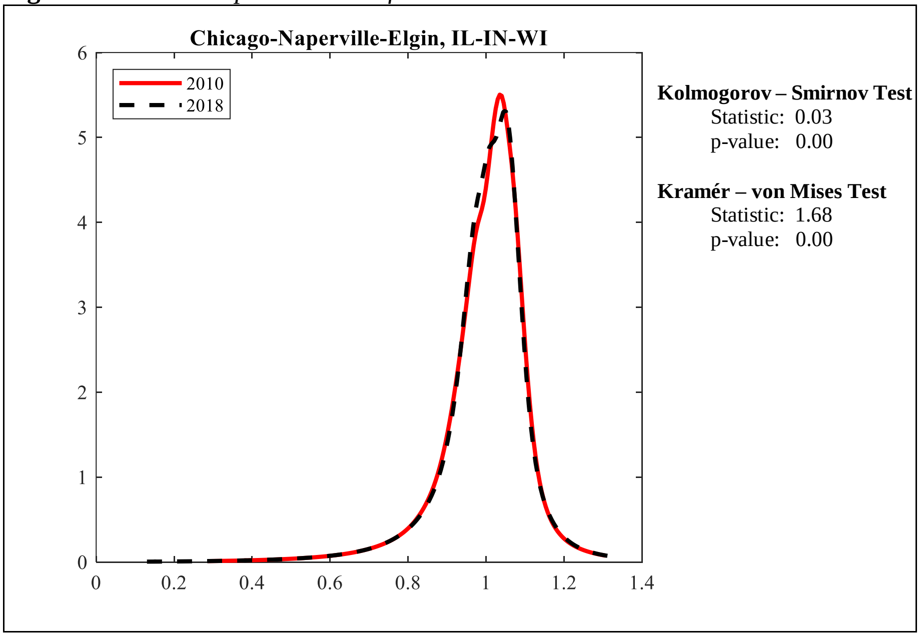


**Figure 2** – *Kernel Representation of Income Distributions*

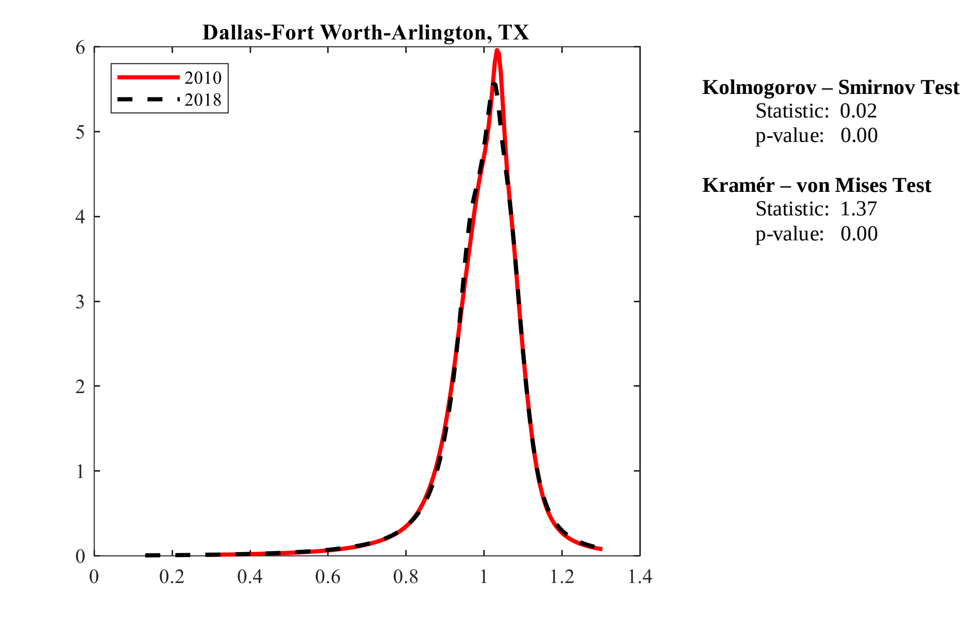



**Figure 3** – *Kernel Representation of Income Distributions*

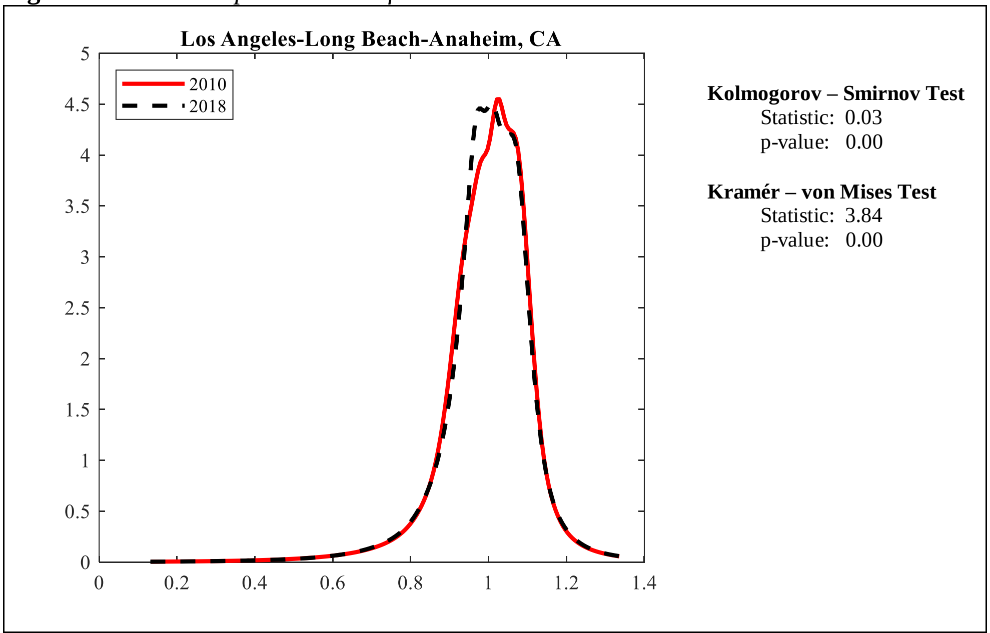
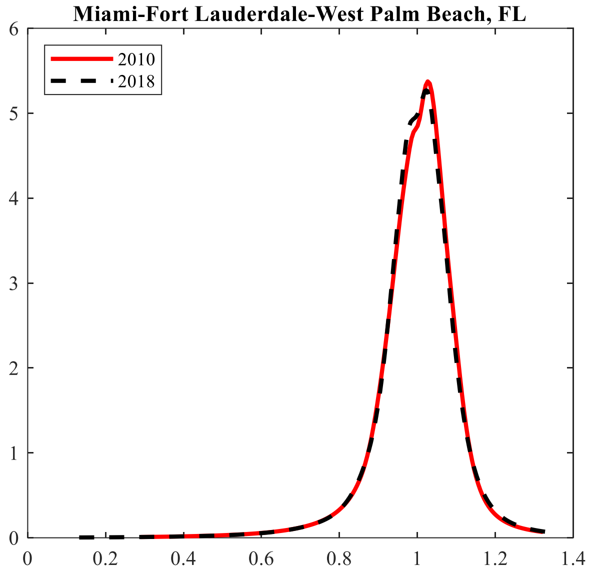


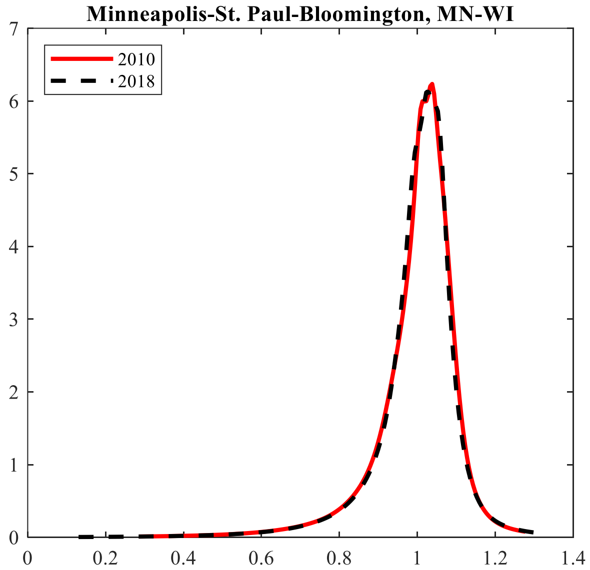
**Figure 4** – *Kernel Representation of Income Distributions*




**Kolmogorov – Smirnov Test**
Statistic: 0.03
p-value: 0.00

**Kramér – von Mises Test**
Statistic: 1.67
p-value: 0.00

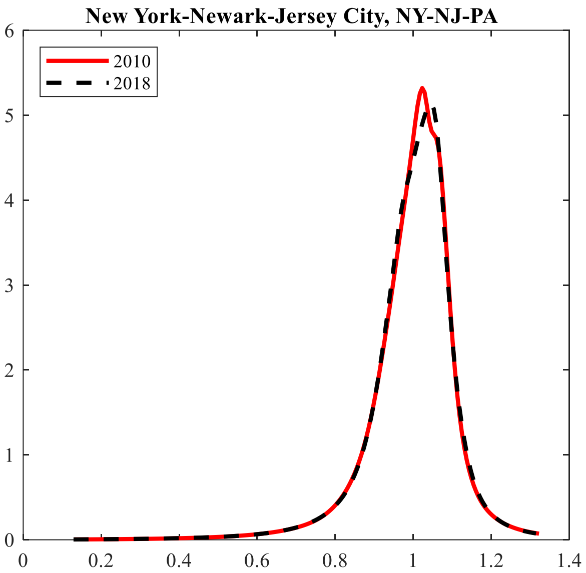
**Figure 5** – *Kernel Representation of Income Distributions*




**Kolmogorov – Smirnov Test**
Statistic: 0.03
p-value: 0.00

**Kramér – von Mises Test**
Statistic: 0.84
p-value: 0.01

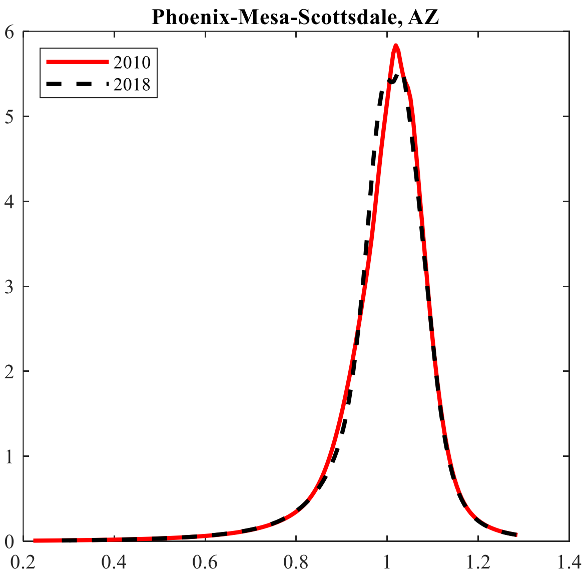
**Figure 6** – *Kernel Representation of Income Distributions*




**Kolmogorov – Smirnov Test**
Statistic: 0.02
p-value: 0.00

**Kramér – von Mises Test**
Statistic: 2.95
p-value: 0.00

**Figure 7** – *Kernel Representation of Income Distributions*




**Kolmogorov – Smirnov Test**
Statistic: 0.03
p-value: 0.00

**Kramér – von Mises Test**
Statistic: 1.31
p-value: 0.00

## References

ATKINSON A.B. 1970. On the measurement of economic inequality, *Journal of Economic Theory*, Vol. 2, pp. 244–263.

GINI C. 1912. Variabilità e Mutabilità. *Studi Economico-giuridici*, *Università di Cagliari*, Anno III, Parte 2°.

GLAESER E.L., RESSENGER M., TOBIO K. 2009. Inequality in cities, *Journal of Regional Science*, Vol. 49, pp. 617-646.

KOLMOGOROV A. 1933. Sulla determinazione empirica di una legge di distribuzione, *Giornale dell'Istituto Italiano degli Attuari*, Vol. 4, pp. 83–91.

KRAMÉR A. 1928. On the composition of elementary errors II, *Scandinavian Actuarial Journal*, Vol. 1, pp. 13–74.

von MISES R.E. 1928. *Wahrscheinlichkeit, Statistik und Wahrheit*. Berlin: Julius Springer.

RUGGLES S., FLOOD S., GOEKEN R., GROVER J., MEYER E., PACAS J., SOBEK M. *IPUMS USA: Version 10.0 [dataset]*. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0

SILVERMAN B.W. 1986. *Density estimation for statistics and data analysis*. London: Chapman and Hall.

SMIRNOV N. 1948. Table for estimating the goodness of fit of empirical distributions, *Annals of Mathematical Statistics*, Vol. 19, pp. 279–281.

THEIL H. 1967. *Economics and Information Theory*. Amsterdam: North-Holland Publishing.

## SUMMARY

# Local inequality analysis in the US: Evidence from some Metropolitan Statistical Areas

In this paper we present an analysis of per capita personal income inequality within seven Metropolitan Statistical Areas of the USA between 2010 and 2018. The analysis is conducted by calculating several well-known income inequality indexes and their variations over the selected period. Then, for each MSA we produce kernel density estimates of the distributions in 2010 and 2018 and perform Kolmogorov-Smirnoff and Kramér-Von Mises tests to evaluate whether they are the same. All results unequivocally portray a picture of significant increases in per capita personal income inequalities.

Margherita GEROLIMETTO, Università Ca' Foscari Venezia, Dipartimento di Economia. E-mail: margherita.gerolimetto@unive.it

Stefano MAGRINI, Università Ca' Foscari Venezia, Dipartimento di Economia. E-mail: stefano.magrini@unive.it

# USO COMBINATO DEGLI INDICI DI MORAN E DI THEIL: UN'APPLICAZIONE SULLA VULNERABILITA' SOCIALE E MATERIALE A LIVELLO TERRITORIALE[2]

Simona Ballabio, Alberto Vitalini

## 1. Introduzione

Molte discipline scientifiche studiano la distribuzione delle caratteristiche socio-economiche all'interno delle unità amministrative o di altre unità territoriali e le differenze tra di esse. Ma spesso si concentrano sulle differenze tra unità amministrative - solitamente macro-aree o regioni - oppure sulla distribuzione spaziale senza tener conto dei vari confini amministrativi.

La variabilità delle caratteristiche della popolazione e la loro distribuzione a livello geografico sono argomenti chiave per lo studio dei fenomeni socio-economici a diversi livelli territoriali. Sono molte le tecniche presenti nel panorama scientifico che indagano questo aspetto, tuttavia non esiste un quadro universale per considerare congiuntamente le suddivisioni amministrative e spaziali nell'analisi dei dati areali.

Il modo più semplice per misurare la variabilità geografica sono le misure statistiche standard di variabilità. Tuttavia, queste misure sono insensibili ai pattern geografici (Fotheringham *et al.*, 2000) e trascurano l'autocorrelazione spaziale portando agli stessi risultati sia quando le unità con alti valori di concentrazione sono adiacenti (alta autocorrelazione) sia quando le unità sono situate in parti anche molto lontane dell'area di studio (bassa autocorrelazione).

Questa situazione però può essere in parte superata utilizzando un approccio che analizza congiuntamente la variabilità e l'autocorrelazione spaziale globale (Netrdová e Nosek, 2017). Tale approccio permette di considerare gli effetti di prossimità nella misurazione della variabilità e di identificare se esiste un modello spaziale significativo. Gli indicatori dell'autocorrelazione spaziale globale misurano l'estensione del raggruppamento spaziale di valori "simili". Ci sono molti modi per misurare l'autocorrelazione spaziale globale a seconda della natura e delle proprietà dei dati (Anselin, 1988), tra questi è frequentemente utilizzato l'indice di Moran, che

[2] L'articolo è frutto del lavoro comune degli autori. Il paragrafo 2 è attribuito ad Alberto Vitalini, il paragrafo 3 è attribuito a Simona Ballabio.

si basa sulla covarianza e presenta molte analogie con il coefficiente di correlazione di Pearson.

In linea con quanto proposto da Netrdová e Nosek (2017) nel presente lavoro utilizziamo congiuntamente due tecniche che combinano tra loro sia l'approccio che tiene conto dei confini amministrativi sia l'approccio spaziale nell'analisi della variabilità. In particolare, per il primo approccio utilizziamo l'indice di Theil, (T di Theil), un indice di scomposizione della varianza che quantifica l'effetto dei confini amministrativi sul fenomeno indagato. Mentre per il secondo approccio utilizziamo l'indice di Moran (I di Moran), misura di autocorrelazione spaziale in grado di quantificare il livello di clustering spaziale nell'area di studio.

Il fenomeno indagato è la vulnerabilità sociale e materiale a livello familiare. In particolare, con l'uso combinato dei due approcci, l'obiettivo è quello di delineare dei pattern territoriali di vulnerabilità e di capire quanto questi pattern siano confinati nei limiti amministrativi provinciali all'interno delle singole regioni italiane. Il vantaggio dell'analisi è anche quello di considerare congiuntamente più confini amministrativi, quello provinciale e quello comunale. La comprensione della distribuzione della vulnerabilità a livello territoriale ha evidentemente valenza non solo in termini di mappatura ma anche in termini di implementazione di efficaci politiche locali.

## 2. Dati e metodi

La base dati utilizzata è di tipo censuario e si riferisce al Censimento della Popolazione e delle Abitazioni del 2011 condotto dall'Istituto Nazionale di Statistica Italiano (Istat). Nello specifico, i dati sono stati estratti da 8milaCensus (Istat, 2015), un sistema di diffusione dei dati censuari sintetizzati attraverso una selezione di 99 indicatori che consentono di delineare un profilo del territorio italiano a diversi livelli territoriali.

La variabile considerata per le analisi è l'indice di vulnerabilità sociale e materiale, testato e validato dall'Istat. Si tratta di un indice che cerca di rilevare l'esposizione di alcune fasce di popolazione a situazioni di rischio, di incertezza della propria situazione sociale ed economica che però non necessariamente si traducono in effettive situazioni di disagio materiale e sociale. L'indice è costruito attraverso la sintesi di sette indicatori[3] e tiene conto della multidimensionalità del fenomeno

[2] Gli indicatori considerati sono: 1. incidenza percentuale della popolazione di 25-64 anni analfabeta e alfabeta senza titolo di studio; 2. incidenza percentuale delle famiglie con potenziale disagio economico, ossia la quota di famiglie giovani o adulte con figli nei quali nessuno è occupato o pensionato; 3. incidenza percentuale delle famiglie con potenziale disagio assistenziale, ossia la quota di famiglie composte solo da anziani ultra 65-enni con almeno un componente ultraottantenne; 4.

indagato cercando di sintetizzare in un unico valore sia la dimensione materiale sia quella sociale della vulnerabilità,

La metodologia utilizzata per la costruzione dell'indice è basata sull'ipotesi di non "sostituibilità" delle diverse componenti (Mazziotta e Pareto, 2014 e 2015) e consente di produrre un indice sintetico non compensativo confrontabile in termini "assoluti" (Adjusted Mazziotta-Pareto Index – AMPI+/-), in grado di facilitare l'individuazione di potenziali aree di criticità attraverso la comparazione tanto spaziale quanto temporale.

Le tecniche di misurazione dell'autocorrelazione spaziale (approccio spaziale) e della variabilità tra unità amministrative (approccio che tiene conto dei confini amministrativi provinciali) utilizzate sono, nel primo caso, l'indice di Moran (I) e, nel secondo caso, l'indice di Theil (T) e la sua scomposizione. Le formule per l'I di Moran (Equazione 1) e l'indice di Theil T (Equazione 2) possono essere scritte come:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - y)(y_j - y)}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\sum_{i=1}^{n}(y_i - y)^2} \tag{1}$$

dove: n = numero di comuni, i e j = indici per due diversi comuni, y = media della variabile in analisi, $w_{ij}$= valore del peso spaziale tra due comuni i, j con $w_{ii}$= 0 (Anselin, 1995);

$$T = (\sum_{j=1}^{k} \frac{n_j}{n}\frac{y_j}{y} ln\frac{y_j}{y}) + \left(\sum_{j=1}^{k} \frac{n_j}{n}\frac{y_j}{y}\sum_{i=1}^{n_j}\frac{y_{ij}}{y_j} ln\frac{y_{ij}}{y_j}\right) = T_B + T_W \tag{2}$$

dove: n= numero di comuni, i = indice per i singoli comuni, j = indice per le province, k = numero di province, y = media della variabile in analisi, per T (indice di Theil complessivo), TB = componente interprovinciale dell'indice Theil, TW = componente intra-provinciale dell'indice Theil (Elbers *et al.*, 2008).

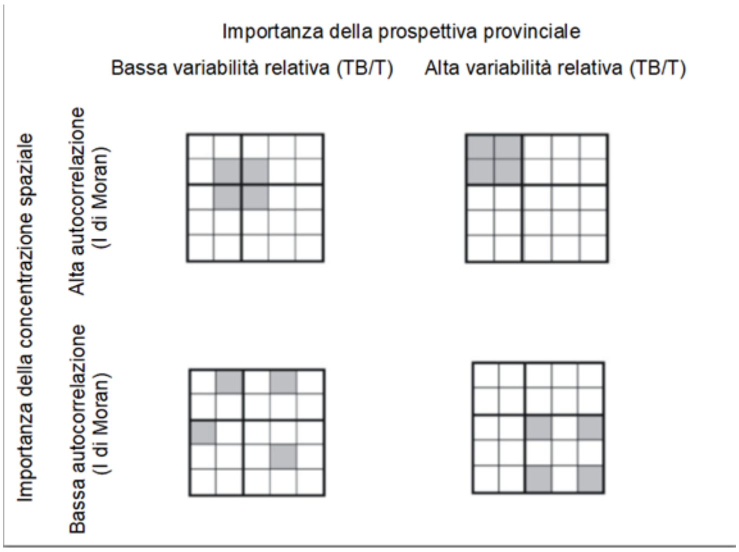
Mentre la TB è una misura della variabilità provinciale, la quota della variabilità provinciale rispetto alla variabilità complessiva (TB/T) misura la variabilità provinciale relativa. Tutti i calcoli riguardanti l'indice di Theil e la sua scomposizione sono stati eseguiti in Stata e tutti i calcoli riguardanti l'I di Moran sono stati eseguiti in GeoDa 1.4.0 (Anselin, 2003; Anselin *et al.*, 2004). Come mostra

incidenza percentuale della popolazione in affollamento grave, ossia il rapporto percentuale tra la popolazione residente in abitazioni con superficie inferiore a 40 mq e più di 4 occupanti o in 40-59 mq e più di 5 occupanti o in 60-79 mq e più di 6 occupanti, e il totale della popolazione residente in abitazioni occupate; 5. incidenza percentuale delle famiglie con 6 e più componenti; 6. incidenza percentuale delle famiglie monogenitoriali giovani (età del genitore inferiore ai 35 anni) o adulte (età del genitore compresa fra 35 e 64 anni) sul totale delle famiglie; 7. incidenza percentuale di giovani di 15-29 anni fuori dal mercato del lavoro e dalla formazione scolastica.

la formula (Equazione 1), per misurare l'autocorrelazione spaziale è necessaria una matrice di peso spaziale ($w_{ij}$) che renda operativa la posizione e la vicinanza delle unità geografiche. In questo lavoro coincide con una matrice di adiacenza, formata da tutte le possibili combinazioni di coppie di comuni, che assume valore "1" (se i due comuni condividono un tratto di confine) o valore "0" (se due comuni non condividono alcun tratto di confine). Per convenzione, ciascun comune non è contiguo a se stesso, pertanto la diagonale principale della matrice è formata da tutti valori pari a zero.

A livello teorico, combinando tra loro le due misure dicotomizzate, I di Moran e misura di variabilità relativa (TB/T) si ricava una tipologia in cui si delineano quattro tipi di risultati (Figura 1).

**Figura 1** – *Tipologia dei dati areali basata sulle prospettive spaziali e provinciali.*




*Note: Ognuno dei quattro quadrati corrisponde ad un tipo di regione. All'interno di ogni tipo di regione i venticinque quadratini piccoli corrispondono ai comuni mentre i quattro quadrati più grandi alle province.*
*Fonte: rielaborazione personale diagramma da Netrdová e Nosek (2017)*

Se sia la variabilità provinciale relativa sia l'autocorrelazione spaziale sono elevate (riquadro in alto a destra), il fenomeno in analisi può essere considerato dipendente dallo spazio e delimitato nelle province. Il fenomeno è concentrato e lo è all'interno dei confini amministrativi predefiniti, le province.

Se il fenomeno si concentra in comuni spazialmente contigui che attraversano i confini provinciali (riquadro in alto a sinistra) la dipendenza spaziale risulta scarsamente correlata con i confini provinciali.

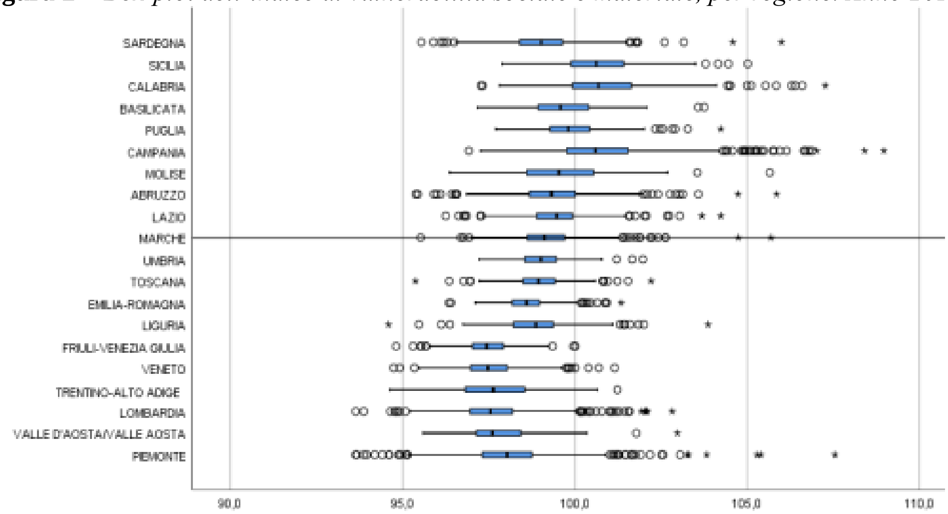
Se il fenomeno non forma concentrazioni spaziali né si concentra nelle province, è indipendente sia dal punto di vista spaziale che provinciale (riquadro in basso a sinistra).

È infine teoricamente possibile che non si osservi alcuna concentrazione spaziale, sebbene il fenomeno si concentri a livello provinciale (riquadro in basso a destra). In quest'ultimo caso il fenomeno sarebbe indipendente dal punto di vista spaziale ma delimitato nelle province.

## 3. Risultati e discussione

I rischi di vulnerabilità sociale e materiale, intesi in termini di incertezza della situazione sociale ed economica familiare che però non necessariamente si traduce in effettive situazioni di disagio, non si distribuiscono in modo omogeneo sul territorio italiano.

**Figura 2 –** *Box-plot dell'indice di vulnerabilità sociale e materiale, per regione. Anno 2011*

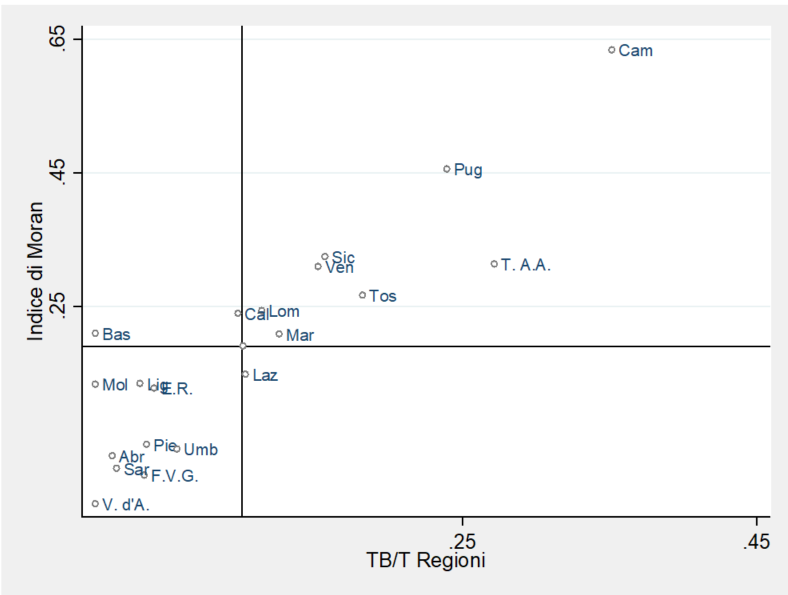



*Fonte: nostre elaborazioni su dati 8milacensus.*

Come mostra la Figura 2, le condizioni di maggior rischio di vulnerabilità sociale e materiale riguardano le regioni del Sud Italia, in particolare Campania, Calabria e Sicilia. In questi casi, oltre ad un valore mediano di incidenza del fenomeno più elevato rispetto alle altre regioni, si osserva anche una maggiore variabilità, quindi una maggiore dispersione dei valori osservati per i singoli comuni. Di converso, confermando anche in questo caso il noto divario Nord-Sud, i minori rischi di vulnerabilità si rilevano nelle regioni settentrionali, in particolare in Friuli-Venezia Giulia, Veneto, Lombardia, Valle d'Aosta, Trentino-Alto Adige e Piemonte. Per il Tentino Alto-Adige si osserva però una variabilità del fenomeno, entro i confini regionali, più marcata rispetto alle altre regioni del Nord.

Per analizzare la diffusione dei fenomeni sia sociali sia economici, compresa la vulnerabilità, l'attenzione è spesso posta sulle differenze tra macro-aree o tra regioni. Tuttavia, è anche interessante capire cosa accade all'interno delle unità territoriali regionali. In particolare, come visto nel paragrafo precedente, l'utilizzo congiunto dell'I di Moran e della misura di variabilità relativa (TB/T) consente, per ciascuna regione italiana, di sintetizzare la distribuzione spaziale della vulnerabilità, rilevata a livello comunale, evidenziando sia il peso dei pattern territoriali sia il loro confinamento entro i limiti amministrativi provinciali.

**Figura 3** – *Diagramma a dispersione dell'I di Moran e dell'indice TB/T (T di Theil) per regione. Anno 2011*




*Fonte: nostre elaborazioni su dati 8milacensus.*

La figura 3 mostra i risultati. Sull'asse verticale si trova l'I di Moran (I), mentre sull'asse orizzontale è presente la misura di variabilità relativa (TB/T) ricavata dall'indice di Theil (T). Come abbiamo visto, la prima fornisce un'indicazione sintetica dell'autocorrelazione spaziale – quindi dell'esistenza di pattern geografici- entro le singole regioni, mentre la seconda misura fornisce un'indicazione sintetica della variabilità provinciale o, detto in altro modo, di quanta parte della variabilità regionale è "spiegata" dalle differenze tra le province.

Dal grafico emerge chiaramente una forte relazione positiva tra le due misure. In particolare, richiamando lo schema teorico presentato nel secondo paragrafo e utilizzando i valori medi a livello nazionale delle due misure per la ripartizione nei quattro quadranti del grafico (linea orizzontale per I e linea verticale per TB/T), si delineano due gruppi di regioni. Nel quadrante in basso a sinistra le regioni in cui il fenomeno della vulnerabilità non forma concentrazioni spaziali né si concentra nelle province, pertanto sembra essere indipendente sia dal punto di vista spaziale sia dal punto di vista provinciale. Nel quadrante in alto a destra invece troviamo le regioni in cui il fenomeno è concentrato spazialmente e lo è all'interno dei confini amministrativi provinciali. Le tre regioni con i valori più bassi di entrambe le misure, tralasciando la Valle d'Aosta che ha un'unica provincia, sono Abruzzo, Sardegna e Friuli-Venezia Giulia. Di converso, le tre regioni con i valori più alti sono Campania, Puglia e Trentino-Alto Adige.

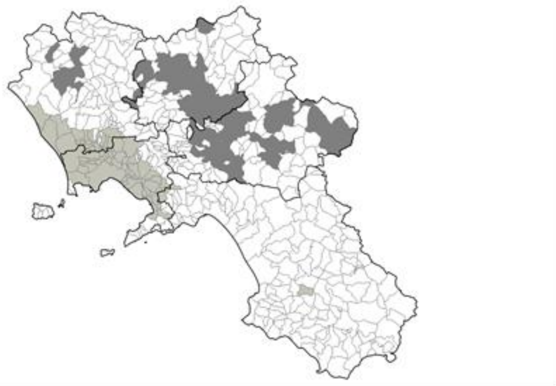
Per una corretta interpretazione e comprensione dei risultati empirici, è fondamentale considerare anche il livello comunale. Sebbene entrambi i metodi considerati (spaziale e provinciale) forniscano importanti informazioni sulla variabilità geografica della vulnerabilità rimangono ancora aspetti da approfondire, ad esempio: come si configurano i cluster spaziali?

Per rispondere a questa domanda ricorriamo alle *Cluster Maps* LISA soffermandoci unicamente sulle tre regioni con i valori più alti delle due misure considerate. Le *Cluster Maps* LISA sono cartografie tematiche che mostrano solo i comuni con valori statisticamente significativi della misura LISA (Anselin, 1995), cioè di una misura di autocorrelazione spaziale locale che rileva la somiglianza tra il valore di una variabile misurata in un'unità di analisi areale (p.es. comune) e i valori della stessa variabile nelle unità vicine. Nel caso specifico, il valore di LISA, calcolato per ogni singolo comune, evidenzia se il comune presenta un valore dell'indice di vulnerabilità tendenzialmente simile o differente rispetto a quello dei comuni vicini, cioè quelli che condividono con il comune un tratto di confine. Valori negativi di LISA indicano diversità: il comune presenta un indice di vulnerabilità differente dai comuni vicini. Valori positivi indicano similarità: il comune presenta un indice di vulnerabilità simile a quello dei comuni vicini.

Nelle mappe sotto riportate (Figure 4, 5 e 6) abbiamo evidenziato i comuni con valori significativamente simili della misura LISA. In particolare, nelle cartografie successive, i comuni con valori LISA significativamente simili ai vicini e valori elevati dell'indice di vulnerabilità sono evidenziati in grigio, quelli con valori LISA significativamente simili ai vicini e valori bassi dell'indice di vulnerabilità sono evidenziati in nero.

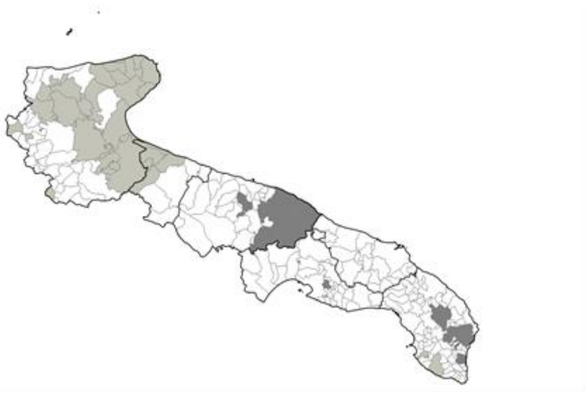
Nella regione Campania i cluster si concentrano nelle province di Caserta e di Napoli e nelle province di Benevento e Avellino. Nel primo caso si tratta di gruppi di comuni con incidenza elevata di vulnerabilità, mentre nel secondo con incidenza bassa. In Puglia i cluster di alta vulnerabilità si concentrano soprattutto nella provincia di Foggia, mentre quelli di bassa vulnerabilità nella zona sud delle province di Bari e di Lecce.

**Figura 4** – *LISA Cluster Map, per comune. Campania. Anno 2011*



*Fonte: nostre elaborazioni su dati 8milacensus.*

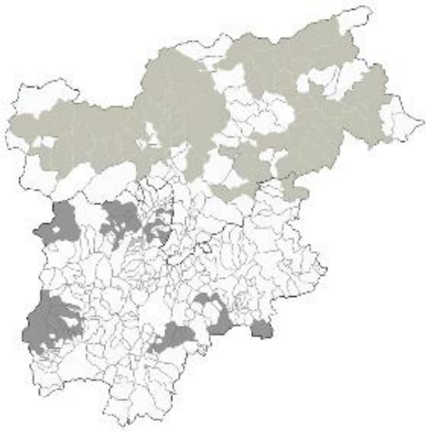
**Figura 5** – *LISA Cluster Map, per comune. Puglia. Anno 2011*



*Fonte: nostre elaborazioni su dati 8milacensus.*

Il Trentino-Alto Adige evidenzia forti differenzi tra le uniche due province che compongono la regione. La provincia di Trento presenta molteplici cluster composti da un numero limitato di comuni caratterizzati da bassa vulnerabilità sociale e materiale. Mentre la provincia di Bolzano presenta estesi cluster caratterizzati da alta vulnerabilità.

**Figura 6** – *LISA Cluster Map, per comune. Trentino-Alto Adige. Anno 2011*



*Fonte: nostre elaborazioni su dati 8milacensus.*

## 4. Conclusione

In questo lavoro si è trattato il tema della vulnerabilità sociale dei comuni italiani e si è cercato di mostrare la necessità, per una sua lettura più completa in termini di distribuzione territoriale, di un superamento delle tradizionali analisi che considerano uno alla volta diversi livelli di aggregazione territoriale: regionale, provinciale e comunale, riconducendole ad un approccio globale basato su due fasi.

Nella prima fase si è descritto come i pattern territoriali di vulnerabilità siano confinati nei limiti amministrativi provinciali all'interno delle singole regioni italiane. Sono stati utilizzati congiuntamente, da un lato, l'indice di Theil e la sua scomposizione in variabilità spiegata dalla provincia per quantificare l'effetto dei confini amministrativi provinciali sul fenomeno indagato e, dall'altro, l'indice di Moran per quantificare il livello di autocorrelazione del fenomeno indagato nella regione.

Questo tipo di analisi ha permesso di individuare alcune regioni, ad esempio Campania e Puglia, in cui il fenomeno è concentrato, sia in termini positivi (basso valore indice di vulnerabilità) sia negativi (alto valore indice di vulnerabilità) solo all'interno di alcune province della regione.

Nella seconda fase, si è focalizzata l'attenzione sul livello comunale ed utilizzando i LISA si è approfondita la configurazione dei cluster spaziali nelle regione individuate nella fase precedente, identificando cluster ben differenziati di comuni contigui con valori simili dell'indice di vulnerabilità.

La sequenza di tecniche proposte, oltre ad avere il vantaggio di studiare la distribuzione spaziale della vulnerabilità sociale e materiale considerando congiuntamente più confini amministrativi, ha anche una valenza in termini di implementazione di efficaci politiche locali, soprattutto se integrata con ulteriori analisi che pongano in relazione i risultati esposti in questo lavoro con le caratteristiche socio-demografiche del territorio.

Sarebbe, infine, utile un'analisi longitudinale per capire come il fenomeno evolve nel tempo, essendoci limitati a mostrare le potenzialità del nostro approccio solo per l'anno 2011.

**Riferimenti bibliografici**

ANSELIN L. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.

ANSELIN L. 1995. Local Indicators of Spatial Association – LISA, *Geographical Analysis*, Vol. 27, No. 2, pp. 93–115.

ANSELIN L. 2003. *An introduction to spatial autocorrelation analysis with Geoda*. Urbana-Champaign, USA: Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, https://personal.utdallas.edu/~briggs/poec6382/geoda_spauto.pdf, accesso 16/07/2020.

ANSELIN L., SYABRI I., KHO Y. 2004. *GeoDa: An introduction to spatial data analysis*. Urbana-Champaign, USA: Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, http://www.csiss.org/events/workshops/geodaGA.pdf, accesso 16/07/2020.

ELBERS C., LAJOUW P. F., MISTIAEN J. A., ÖZLER, B. 2008. Reinterpreting between-group inequality, *The Journal of Economic Inequality*, Vol. 6, No. 3, pp. 231–245.

FOTHERINGHAM A.S., BRUNSDON C., CHARLTON, M. 2000. *Quantitative geography – Perspectives on spatial data analysis*. London: SAGE.

ISTAT. 2015. *Indice di vulnerabilità sociale e materiale*, Roma: Istituto nazionale di statistica , http://ottomilacensus.istat.it, accesso 16/07/2020.

NETRDOVÁ P., NOSEK V. 2017. Exploring the Variability and Geographical Patterns of Population Characteristics: Regional and Spatial Perspectives, *Moravian*
*Geographical Reports*, Vol. 25, No. 2, pp. 85-94.

MAZZIOTTA M., PARETO A. 2014. A Composite Index for measuring Italian regions' development over time, *Rivista italiana di economia e demografia*, Vol. LXVIII, No. 3/4.

MAZZIOTTA M., PARETO A. 2015. On a Generalized Non-compensatory composite Index for Measuring Socio-economic Phenomena, *Social Indicators Research*, Vol. 127, pp. 983-1003.

## SUMMARY

### Combined use of Moran and Theil indices: an application on social and material vulnerability at territorial level

The variability of the characteristics of the population and families at different territorial levels is an important key to interpreting social phenomena. This work deals with the social vulnerability of Italian municipalities.

In the analyses we use the empirical indicators of vulnerability based on the 2011 census of the Italian population.

This work goes beyond the traditional analyses that consider, one at a time, different levels of territorial aggregation: regional, provincial and municipal, leading them back to a global approach based on two phases.

In the first phase we described how the patterns of vulnerability are confined to the provinces of the individual Italian regions, using jointly: the Theil index and its breakdown into variability explained by the province to quantify the effect of the provincial administrative boundaries on the investigated phenomenon and the Moran index to quantify the level of self-correlation of the investigated phenomenon in the region. This type of analysis made it possible to identify some regions, for example Campania and Puglia, where the phenomenon is concentrated only within some provinces of the region.

In the second phase, attention was focused on the municipal level and using LISAs, we improved the understanding of spatial clusters of municipalities in the regions identified in the previous phase



Simona BALLABIO, Istat - Ufficio territoriale per la Lombardia, ballabio@istat.it
Alberto VITALINI, Istat – Ufficio territoriale per la Lombardia, vitalini@istat.it

# SPATIAL INFORMATION COMPREHENSIVE WELL-BEING COMPOSITE INDICATORS: AN ILLUSTRATION ON ITALIAN VARESE PROVINCE[1]

Carlotta Montorsi, Chiara Gigliarano

## 1. Introduction

In recent years, the concept of well-being and its measurement has been at the forefront of the European research topic debates (Stiglitz *et al*., 2009). However, despite some great advancements, a unique definition of well-being has not been provided yet (Fiorillo *et al*., 2017). The well-being is thought as a multidimensional phenomenon that mirrors the values and preferences of a society and its citizens. Hence, its effective and faithful description requires relying on a set (or dashboard) of relevant indicators (Hall *et al.*, 2010). On the other hand, to enhance practicability, the complex information enclosed in such a dashboard of indicators must be synthesized through the construction of composite indicators (European Commission, 2008).

Within the Italian framework, the first project contextualized to this debate is the "Equitable and Sustainable Well-being (Bes)", jointly proposed in March 2013 by the National Council for Economics and Labor (Cnel) and the Italian National Institute of Statistic (Istat, 2018). In 2016, to complement this project with one more focused on the local level (NUTS3), a new project was started on "Well-being and planning measure at the municipal level (*"Misure di benessere e programmazione a livello comunale"*), coordinated by Istat, National Association of Italian municipalities (ANCI) and Union of Italian Provinces. The project aims to provide an integrated and harmonized data-set and information systems with a high local detail and to support local authorities in policymaking. The resulting dataset "*A Misura di Comune*" comes from an integration of different data sources, such as administrative archives or statistical surveys, which share the characteristic of being total and not sample sources.

[1] The authors gratefully acknowledge funding support from *Fondazione Giovanni Valcavi per l'Università degli Studi dell'Insubria*.

The main differential aspect of this data-set with respect to the others elaborated by Istat - “Bes” and “UrBes”- is the presence of particularly meaningful well-being domains, such as “Population and family” and “Mobility and Infrastructure”, for the use of their indicators in the “Single Programming document of local authorities”.

This paper exploits the “A Misura di Comune” dataset[2] for assessing the well-being of the Varese province by constructing composite indicators that synthesize, on a statistical basis, the atomic indicators information. We implement Bayesian factor analysis for spatially correlated data. Factor analysis for constructing composite indicators on BES dataset has already been proposed, for example in Chelli *et al.* (2015) and Ciommi *et al.* (2017). However, to the best of our knowledge, this is the first attempt to construct well-being composites indicators inclusive of spatial information in the underlying statistical model.

## 2. Methodology

The “A Misura di Comune” dataset is constituted by 50 atomic indicators, which are grouped into 10 macro-domains. Data are collected for four years, from 2014 to 2017, in all Italian municipalities.

The state of the art of aggregation methods entails a broad list of different approaches, form the simpler linear aggregation to more constructed ones. Constructed empirical indices, such as AMPI or GAMPI, are built on non-substitutable and non-compensatory indicators and allow for comparison across space (Mazziotta and Pareto, 2013). Nonetheless, they are based on several structural assumptions (Ciommi *et al*., 2017). For example, no adjustment is made for differential precision of the atomic indicators across local units that may have different population sizes. Moreover, it is usually assumed that for a specific area, information about well-being depends exclusively on variables from that area, and not on variables from neighboring areas. And finally, the traditional indices lack a posterior measure of uncertainty. This last feature can be problematic, for example, if decisions about policies or resource allocation are based on cutoff values or percentile of the index.

Hence, we try to move forward these shortcomings exploiting the methodology introduced in Hogan *et al*. (2004). We treat the “A Misura di Comune” atomic indicators as manifest variables while the well-being is the underlying latent factor. Hence, the well-being is defined as the posterior expectation of the latent factor given the manifest variables and the model parameters. Therefore, under the assumption

---

[2] The dataset is available at http://amisuradicomune.istat.it/aMisuraDiComune/

that adjacent areas have similar socioeconomic characteristics, we introduce spatial dependencies among the latent well-being variable of each municipality.

Given the presence of missing values, we have restricted our analysis to two years, 2014 and 2015. Moreover, we excluded from the analysis the indicators related to the domain "Population and family", since they are not strictly related to well-being. Our analysis hence is based on 32 atomic indicators and focuses on the 139 Varese province municipalities.

Our prior hypothesis is that neighboring municipalities share information on socio-economic development levels. Hence, accounting for their spatial information should increase the accuracy of our estimates. Therefore, before constructing the statistical model, we tested for spatial autocorrelation in the atomic indicators through the Global Moran I test (Moran, 1950), which provided significant results for almost all the indicators.

For municipality $i$, with $i = 1, \dots, N$ and $N = 139$, let $Y_{id}$ denotes the atomic indicator $d$ in municipality $i$, where $d = 1, \dots, D,$ and $D = 32$. Hence $Y_i = (Y_{i1}, \dots, Y_{iD})^T$ is the well-being profile of municipality $i$ . The general latent factor model assumes for each area a L dimensional $(L < D)$ latent variable $\delta_i = (\delta_{i1}, \dots, \delta_{iL})^T$, that fully characterizes socio-economic characteristics, which in turn manifest themselves through $Y_i$. We assume $L = 1$, hence reducing the model to one latent factor for each municipality and we represent the model in a hierarchical form.

At the first level we have:

$$\mathrm{Y_i} \mid \mu_i, \delta_i, \Sigma \sim \text{Multivariate-Normal}(\mu_i + \lambda\delta_i, \Sigma),$$

where $\mu_i$ is $a\, D \times 1$ mean vector, $\lambda$ is a $D \times 1$ vector of factor loading's and $\Sigma = \mathrm{diag}(\sigma_1^2, \dots, \sigma_D^2)$ is a diagonal matrix measuring residual variation in $Y_i$. Assuming $\Sigma$ diagonal implies independence among the elements of $Y_i$ conditionally on $\delta_i$.

Spatial autocorrelation is introduced at a second level. Let $\delta = (\delta_1, \dots, \delta_N)^T$ the municipalities' latent indexes vector. Thus, it is assumed:

$$\delta \sim \text{Multivariate-Normal}(0_N, \Psi),$$

where $\Psi$ is a $N \times N$ spatial variance-covariance matrix having 1's on the diagonal and $\psi_{ij} = corr(\delta_i, \delta_j)$ on the off-diagonal. When $\Psi = I_N$ the model assumes spatial independence. The well-being composite index for municipality $i$ is summarized by the conditional distribution of the latent factor $\delta_i$ given $Y$ and $\mu, \lambda, \Sigma$. Hence the posterior distribution of vector $\delta$ is a Multivariate normal distribution:

$$(\delta \mid Y, \mu, \lambda, \Sigma) \sim \text{Multivariate-Normal}(d, D),$$

$$\text{where } D = \{\Psi + \Lambda^T \Sigma^{-1} \Lambda\}^{-1} \text{ and } d = D \Lambda^T \Sigma^{-1} (Y - \mu).$$

We have chosen a conditional parametrization of the spatial variance-covariance matrix $\Psi$, through conditional auto-regressive specifications of spatial dependency. The more general structures are the Gaussian CAR models (Besag J.,1974; Sun *et al.*, 1999). Generally, these models require to construct a set $\mathcal{R}_i$ ,denoting the set of indices $\delta_j$ for areas that are neighbors of the area $i$. Then, they assume that:

$$\delta_i \mid \{\delta_j : j \in \mathcal{R}_i\} \sim N\left(\sum_{j \in \mathcal{R}_i} \beta_{ij} \delta_j , \frac{\nu}{\alpha_i}\right),$$

so that

$$(\delta_1, \dots, \delta_N)^T \sim \text{Multivariate-Normal}(0_N, \nu B^{-1}),$$

where $B$ is $N \times N$ matrix with $\{\alpha_1, \dots, \alpha_N\}$ along the diagonal and $-\alpha_i \beta_{ij}$ on the off-diagonal, provided that $B$ is symmetric and positive definite (Sun *et al.*, 1999). In order to ensure these conditions to hold one must constrain one or more parameters, but the constraints are model specific.

According to how we have defined the $\mathcal{R}_i$ and $\beta_{ij}$, different CAR models are specified. For this analysis we have defined: $R_{ij} = I(j \in \mathcal{R}_i)$, which is the indicator function that area $j$ is a neighbor of area $i$, $\beta_{ij} = \omega R_{ij}$, $\alpha_i = 1$ and $\nu = 1$; then $B = I_N - \omega R$, where $R$ is an adjacency (weight) matrix with $R_{ii} = 0$ and indicators $R_{ij} = R_{ji}$. One necessary condition for $B$ to be positive definite is that the ordered eigenvalues $\xi_1, \dots, \xi_N$ of $R$ have to satisfy: $\xi_1^{-1} < \omega < \xi_N^{-1}$.

Finally, a characteristic of the Bayesian framework is the introduction of prior distributions on all the model's parameters. In our model we have set: $\lambda_j \sim \text{Normal}(g, G) I(\lambda_1 > 0)$; $\sigma_j^2 \sim \text{Inverse-Gamma}(\alpha/2, \beta/2)$; $\mu_j \sim \text{Normal}(0, V_\mu)$.

The scope of prior distributions is to include subjective opinions on the parameters of interest. However, we let the data “speak for them-self” and choose uninformative priors with $g = 0$, $G = 10000$, $\alpha = 1/1000$, $\beta = 1/1000$, $V_\mu = 1000$.

The model is estimated through a Gibbs sampling that includes Metropolis Hasting steps for the spatial parameter $\omega$.[3]

[3] We have written the sampling algorithm in the R software and made it available in GitHub through the link https://github.com/CarlottaMnt/Bayesian-factor-analysis-sampler.

## 3. Results

Following this methodology, we have first computed a uni-dimensional overall well-being composite indicator which synthesizes, for each municipality, the 32 “A Misura di Comune” atomic indicators.

In Table 1 we report the factor loadings' distributions, which represent the covariances among the “A Misura di Comune” atomic indicators and the composite indicator (latent variable). Factor loadings with negative sign impact negatively the latent well-being, such that an increase in the corresponding atomic indicator leads to a decrease in the overall well-being. On the other hand, factor loadings with a positive sign would raise the value of the overall well-being. When the estimated factor loading is around zero all along its distribution, we consider the associated indicator meaningless for the well-being.

The main contributor with a positive impact on the composite indicator is the “Gross Income per capita” while the one with the greater negative impact is “Household with gross income less than the social allowance benefit”. Having zero impact are the “Self-containment index”[4] and the “Leakage of drinking water”.

Figure 1 illustrates the estimate of the well-being composite indicator for the Varese municipalities in 2014 and 2015. For each municipality, the graph reports the mean value of the composite indicator and its posterior 95% credibility intervals. We have highlighted in red the most populated municipalities, i.e. Varese, Gallarate, Busto Arsizio and Saronno. Among the two years considered, the municipalities' rank in term of composite indicator slightly changes, while the polarization among municipalities does not change significantly.

Finally, the maps in Figure 2 report the spatial distribution of the composite indicator's mean across the Varese municipalities. The above-average values are in green while below-average values are in red. We notice clusterization in the well-being phenomenon, which is discriminated among northern and southern municipalities, whereas the first appreciate the lowest level of well-being and the latter are better off. This result is maintained throughout the two years considered.

Next, we proceed with our well-being assessment by constructing a composite indicator for each of the three, non-fungible, sustainable development domains: social well-being, economic well-being and environmental well-being (Ciommi et al., 2020). This will clarify the contribution each domain provides on the overall well-being and the relative importance of the atomic indicators with respect to the specific well-being composite indicator they interact with. It may be that, when the number of atomic indicators in each domain is too little, the uncertainty of the composite indicator as well in the factor loading's estimates increases.

[4] It represents the ratio among the monetary flows from those who work within the municipal boundaries and the overall monetary flows generated overall in the municipality.

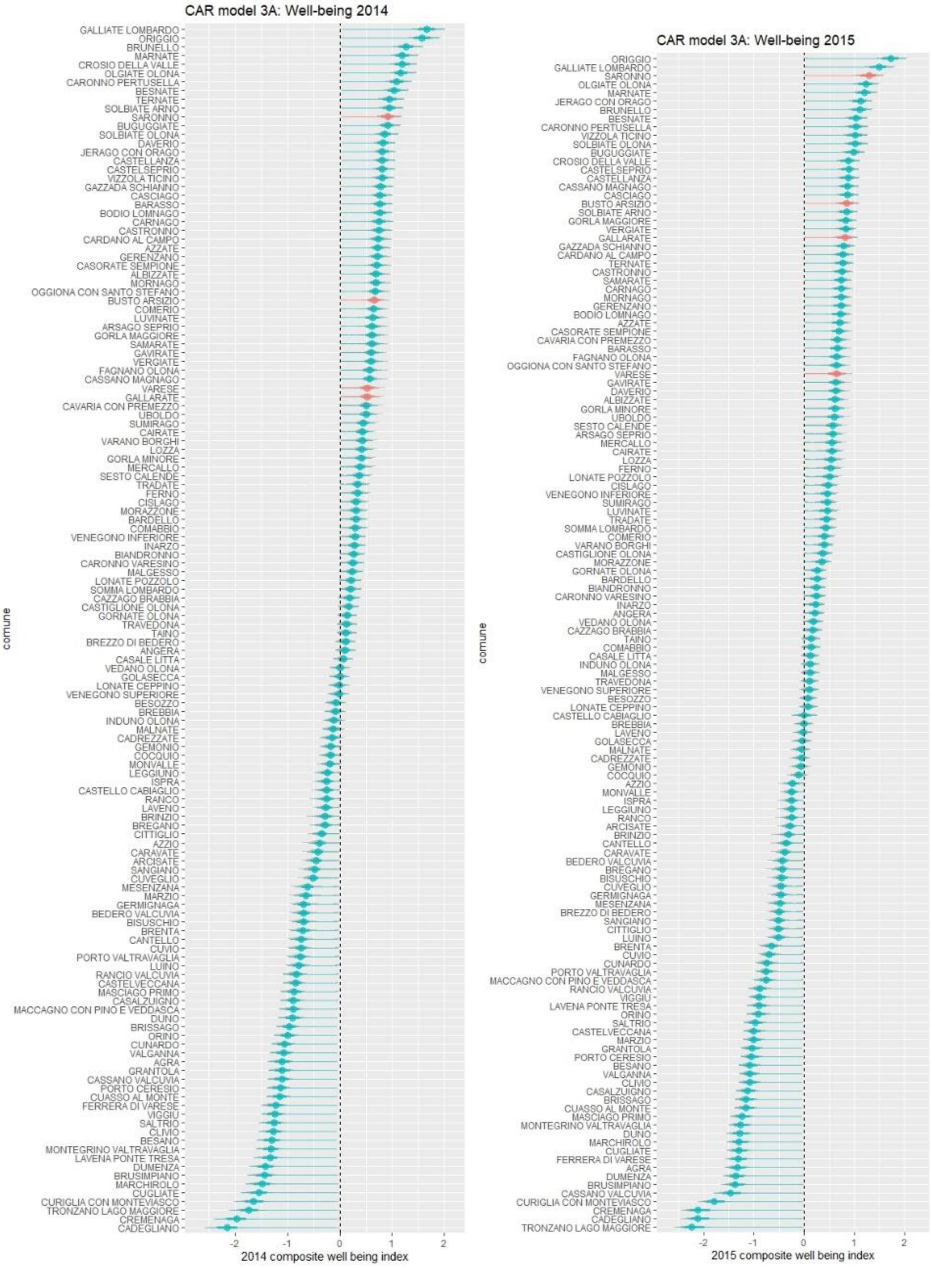
As a first result, Table 2 to Table 4 report the factor loading's distribution in the four domains.

**Table 1 -** *Summary of the factor loadings' distribution for each "A misura di comune" indicators, year 2014.*

| Indicator | Mean | 5% | 50% | 95% |
|---|---|---|---|---|
| Circulating polluting vehicles | -2.70 | -7.46 | -0.67 | -0.51 |
| Children in municipal childcare services | 0.83 | 0.05 | 0.21 | 2.27 |
| Household low labour intensity | -4.23 | -12.28 | -1.03 | -0.85 |
| Soil consumption | 3.53 | 0.67 | 0.84 | 11.65 |
| IRPEF taxpayers with total income < 10.000 | -4.38 | -11.52 | -1.00 | -0.83 |
| Local units' density | 3.25 | 0.59 | 0.75 | 9.86 |
| High school graduates (25-64) | 2.47 | 0.48 | 0.63 | 6.98 |
| Leakage of drinking water | 0.40 | -0.06 | 0.09 | 0.67 |
| Gross income differences | 4.41 | 0.83 | 1.00 | 12.74 |
| Women and political representation - City Council | -0.20 | -0.20 | 0.01 | 0.20 |
| Women and decision-making - Municipal council | 0.92 | 0.07 | 0.22 | 1.99 |
| Mean age local administrator | -0.84 | -2.17 | -0.19 | -0.04 |
| Mean age municipal councillors | -0.49 | -0.73 | -0.09 | 0.07 |
| Household with gross income less then social allowance benefit | -4.39 | -12.96 | -1.01 | -0.84 |
| Single-income households with children (age < 6) | -2.56 | -7.20 | -0.58 | -0.43 |
| Neet | -4.08 | -12.42 | -0.99 | -0.81 |
| Attraction Index | 2.25 | 0.37 | 0.52 | 6.18 |
| Self-containment index | 0.12 | -0.18 | 0.02 | 0.21 |
| Harmfulness of road accidents | -1.26 | -3.31 | -0.30 | -0.15 |
| Mortality index of road accidents | 0.29 | -0.01 | 0.15 | 0.95 |
| Employed (20-64) | 4.58 | 0.85 | 1.02 | 13.37 |
| Not stable employed | -1.53 | -3.27 | -0.31 | -0.16 |
| Graduates (30-34) | 2.36 | 0.40 | 0.56 | 6.66 |
| Museum, galleries, monuments | -1.50 | -4.28 | -0.39 | -0.23 |
| Electoral participation | 2.73 | 0.48 | 0.63 | 8.00 |
| Separate collection of municipal waste | 1.06 | 0.17 | 0.33 | 3.76 |
| Gross per capita income | 4.33 | 0.84 | 1.02 | 13.02 |
| Production specialization in high-tech sectors | 1.95 | 0.34 | 0.49 | 5.96 |
| Rate of entrepreneurship | 2.86 | 0.56 | 0.73 | 8.73 |
| Number of road accidents | 0.78 | 0.02 | 0.18 | 1.68 |
| Jobs' transformation from not stable to stable | 1.09 | 0.10 | 0.26 | 2.81 |
| Visitors to museum, galleries, monuments | -0.62 | -1.65 | -0.14 | 0.01 |

*Source: our elaboration on "A misura di comune" data.*

**Figure 1 –** *Overall well-being indicator estimate and its 95% credibility interval for 2014 and 2015***.**




*Source: our elaboration on "A Misura di Comune" data.*

**Table 2 –** *Social well-being: factor loadings' distribution, year 2014.*

| Indicator | Mean | 5% | 50% | 95% |
|---|---|---|---|---|
| Children in municipal childcare services | -0.31 | -0.22 | -0.07 | -0.01 |
| High school graduates (25-64) | 0.85 | 0.45 | 0.64 | 0.85 |
| Women and political representation - City Council | 0.55 | 0.10 | 0.29 | 0.50 |
| Women decision-making - Municipal council | 0.57 | 0.15 | 0.33 | 0.52 |
| Mean age local administrator | -0.60 | -0.67 | -0.45 | -0.24 |
| Mean age municipal councillors | -0.26 | -1.09 | -0.26 | 0.53 |
| Neet | -0.56 | -0.65 | -0.45 | -0.25 |
| Harmfulness of road accidents | -0.33 | -0.45 | -0.26 | -0.08 |
| Mortality index of road accidents | -0.37 | -0.32 | -0.12 | 0.06 |
| Graduates (30-34) | 0.85 | 0.48 | 0.65 | 0.86 |
| Museum, galleries and monuments | 0.12 | -0.25 | -0.04 | 0.17 |
| Number of road accidents | -0.56 | -0.33 | -0.13 | 0.05 |
| Visitors to museum, galleries and monuments | 0.26 | -0.17 | 0.01 | 0.21 |

*Source: our elaboration on "A misura di comune" data*

**Table 3 –** *Economic well-being: factor loadings' distribution, year 2014.*

| Indicator | Mean | 5% | 50% | 95% |
|---|---|---|---|---|
| Household low labour intensity | -6.84 | -44.49 | -1.05 | -0.89 |
| IRPEF taxpayers with total income < 10.000 euros | -6.72 | -45.48 | -1.03 | -0.89 |
| Local unit density | 4.44 | 0.57 | 0.68 | 28.21 |
| Gross income differences | 6.51 | 0.84 | 0.99 | 44.25 |
| Household with gross income < social allowance | -6.66 | -44.62 | -1.05 | -0.89 |
| Single-income households with children (age < 6) | -4.15 | -27.65 | -0.63 | -0.51 |
| Attraction Index | 3.25 | 0.40 | 0.51 | 20.70 |
| Self-containment index | -0.44 | -1.53 | -0.06 | 0.03 |
| Employed (20-64) | 6.78 | 0.89 | 1.04 | 45.30 |
| Not stable employed | -2.21 | -14.35 | -0.35 | -0.26 |
| Electoral participation- municipal elections | 3.99 | 0.54 | 0.65 | 28.30 |
| Gross per capita income | 6.40 | 0.84 | 0.99 | 42.29 |
| Production specialization in high-tech sectors | 3.27 | 0.40 | 0.50 | 22.16 |
| Rate of entrepreneurship | 0.63 | 0.47 | 0.63 | 0.79 |
| Jobs transformation from not stable to stable | 1.72 | 0.19 | 0.28 | 10.97 |

*Source: our elaboration on "A misura di Comune" data*

**Table 4 –** *Environmental well-being: factor loadings' distribution, year 2014.*

| Indicator | Mean | 5% | 50% | 95% |
|---|---|---|---|---|
| Circulating polluting vehicles | -0.90 | -0.84 | -0.66 | -0.46 |
| Soil consumption | 0.98 | 0.49 | 0.65 | 0.83 |
| Leakage of drinking water | 0.10 | 0.04 | 0.21 | 0.40 |
| Separate collection of municipal waste | 0.05 | 0.05 | 0.26 | 0.44 |

*Source: our elaboration on "A misura di Comune" data*

The leading domain is "economic well-being" (Table 3), where all the factors' loadings, in absolute value, are far from being zero. In this domain, the indicators driving the composite indicator to greater positive values are “Gross per capita income” and “Employed (20-64)”. Table 2 reports the social well-being indicators' factor loadings. With greater positive impact on the social well-being, despite being small and near zero, are “High school graduates” and “Graduates (25-64)”, revealing the importance of education in boosting the estimated well-being. Lastly, the environmental well-being (Table 4) is mainly explained by “Soil consumption”, which is the ratio among the soil consumed and the overall municipal soil, and “Circulating vehicles with standard emissions lower than euro 4”, that points out the negative role played by motor vehicles on air pollution.

**Figure 2 –** *Overall well-being indicator's spatial distribution in Varese province.*

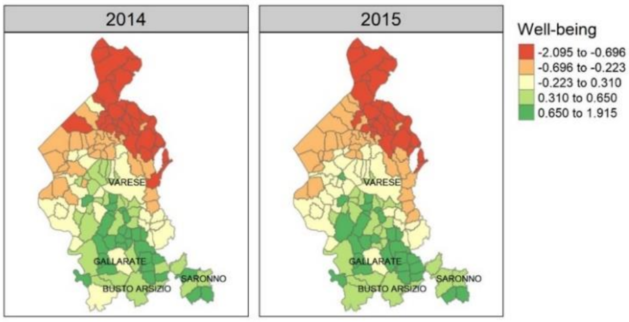
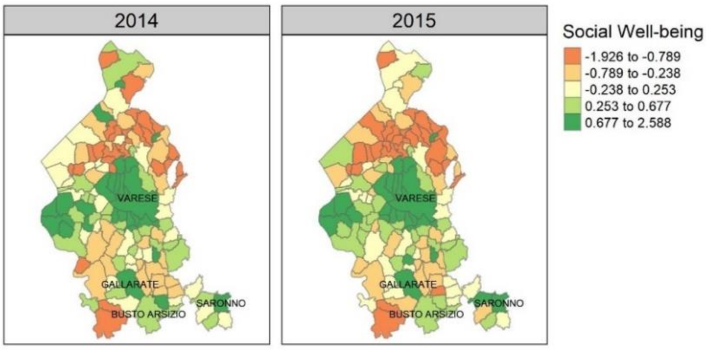



**Figure 3 –** *Social composite indicator's spatial distribution in Varese province.*




Finally, we move directly to the spatial distribution of the composite indicators across the Varese province. Figures from 3 to 5 map respectively the mean of the composite indicator distribution for social well-being, economic well-being and environmental well-being. The social well-being composite indicator interestingly shrinks from 2014 to 2015: the worse off municipalities, despite remaining below the average, increase their social well-being, indeed they become “clearer”, while

the better off municipalities slightly reduce their social well-being. The spatial distribution for the economic well-being composite indicator detects the presence of three separated groups, from the Southern to the Northern municipalities, with high, medium and low economic well-being. This figure looks like Figure 1, corroborating the greater importance of economic well-being in driving the overall well-being.
Finally, the environmental well-being spatial distribution also highlights differences among municipalities in the North and in the South, with the latter performing better

**Figure 4 –** *Economic composite indicator's spatial distribution in Varese province.*

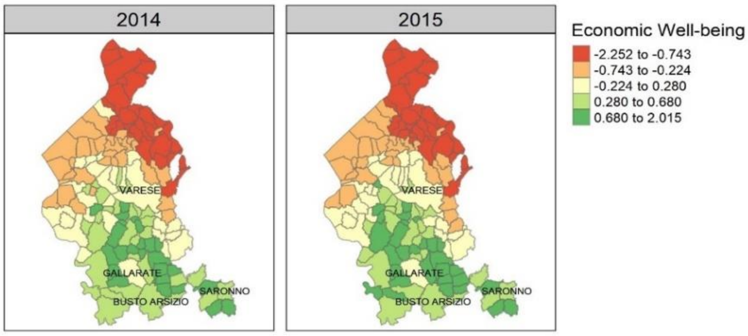
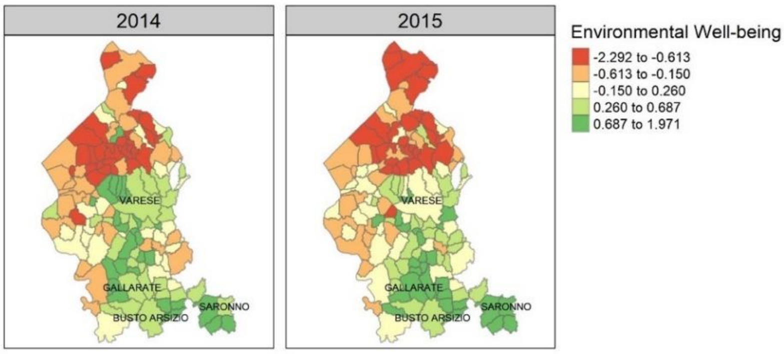



**Figure 5 –** *Environmental composite indicator's spatial distribution in Varese province.*




## 4. Concluding remarks

We have constructed well-being composite indicators by adopting a Bayesian latent variable approach which includes spatial information. From the overall well-being assessment on the Varese province, we have estimated heterogeneous well-being levels across the province's municipalities. Notably, when considering all the "A Misura di Comune" atomic indicators, the resulting composite indicator is clustered among Northern and Southern municipalities, where the former enjoys, on average, a lower well-being level with respect to the latter.

Next, we have analyzed the well-being within each of the three sustainable development domains. We have highlighted the greater importance of economic well-being in driving the overall municipalities' well-being. Within this domain, the leading indicators are related to income and occupational levels.

Given the severe presence of missing values, our analysis focuses only on two years, 2014 and 2015. However, as soon as updated data becomes available, future research will consider the inclusion of temporal information within the model.

## References

BESAG J. 1974. Spatial interaction and the statistical analysis of lattice systems, *Journal of the Royal Statistical Society: Series B*, Vol. 36, No. 2, pp. 192-225.

CHELLI F.M., CIOMMI M., EMILI A., GIGLIARANO C., TARALLI S. 2015. Comparing equitable and sustainable well-being (Bes) across the Italian provinces: a factor analysis-based approach, *Rivista Italiana di Economia Demografia e Statistica*, Vol. 69, pp. 61-72.

CIOMMI M., GIGLIARANO C., EMILI A., TARALLI S., CHELLI F.M. 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*, Vol. 76, pp. 281-296.

CIOMMI M., GIGLIARANO C., TARALLI S., CHELLI F.M. 2017. The equitable and sustainable well-being at local level: A first attempt of time series aggregation, *Rivista Italiana di Economia Demografia e Statistica*, Vol. 71, No. 4, pp. 131-142.

CIOMMI M., GIGLIARANO C., CHELLI F.M., GALLEGATI M. 2020. It is the total that does [not] make the sum: nature, economy and society in the Equitable and Sustainable Well-Being of the Italian Provinces, *Social Indicators Research*, https://doi.org/10.1007/s11205-020-02331-w.

EUROPEAN COMMISSION, 2008. *Handbook on constructing composite indicators: methodology and user guide*. OECD publishing.

FIORILLO F., MUSCILLO C., TARALLI S. 2017. Misure di benessere dei territori e programmazione strategica: il livello comunale, *Economia pubblica,* Vol. 1, pp. 61-96.

HALL J., GIOVANNINI E., MORRONE A., RANUZZI G. 2010. A framework to measure the progress of societies, *OECD Statistics Working Papers 05*, OECD Publishing, Paris, https://doi.org/10.1787/5km4k7mnrkzw-en.

HOGAN J. W., TCHERNIS R. 2004. Bayesian factor analysis for spatially correlated data, with application to summarizing material deprivation from census

data, *Journal of the American Statistical Association*, Vol. 99, No. 466, pp. 314-324.
ISTAT 2018. *Il benessere equo e sostenibile in Italia*. Roma: Istat.
MAZZIOTTA M., PARETO A. 2013. A non-compensatory composite index for measuring well-being over time, *Cogito, Multidisciplinary Research Journal*, Vol. 5, No. 4, pp. 93-104.
MORAN P.A. 1950. Notes on continuous stochastic phenomena, *Biometrika*, Vol. 37, No.1/2, pp. 17-23.
STIGLITZ J.E., SEN A., FITOUSSI J.P. 2009. *Report by the commission on the measurement of economic performance and social progress*. Paris: Commission on the Measurement of Economic Performance and Social Progress
SUN D., TSUTAKAWA K., SPECKMAN P.L. 1999. Posterior distribution of hierarchical models using car(1) distributions, *Biometrika*, Vol. 86, pp. 341-350.

**SUMMARY**

**Spatial information comprehensive well-being composite indicators: an illustration on Italian Varese province**

This analysis bears upon the European "Beyond GDP initiative", which promotes multi-dimensional approaches going beyond the traditional and uni-dimensional GDP macroeconomic indicator to monitor the living condition and the well-being of a territory. We assess the well-being within the Varese province by applying factor analysis with integration of spatial information in a Bayesian framework. To summarize the large number of indicators within the 10 domains that constitutes "A misura di comune" dataset we construct four composite indicators for each Varese municipality. The first is comprehensive of all the "A misura di comune" indicators but not the one related to the Population and Family domain. The last three composite indicators assess the municipalities in term of their social, economic and environmental well-being. We highlight differentials across Northern and Southern municipalities in all the well-being domains, with the former performing usually better. We also identify in the economic domain the leading well-being domain, that drives the overall wealth to higher values.

Carlotta MONTORSI, Università dell'Insubria, carlotta.montorsi@uninsubria.it
Chiara GIGLIARANO, Università dell'Insubria, chiara.gigliarano@uninsubria.it

# EVERYTHING YOU ALWAYS WANTED TO KNOW ABOUT NORMALIZATION (BUT WERE AFRAID TO ASK)[1]

Matteo Mazziotta, Adriano Pareto

## 1. Introduction

Researchers of social science often use statistical techniques with data that have different units of measurement and ranges (e.g., "Life expectancy" and "Gross national income per capita"). A typical example is the construction of a composite index, where a set of individual indicators of different nature must be aggregated (Salzman, 2003). A common practical problem the researcher faces is how to normalize indicators in order to make them comparable. Unfortunately, most of the reports in the literature describe the main normalization methods (OECD, 2008), but do not explain how to choose the method that best suits the needs of the researcher. So, many researchers choose the normalization method almost solely on the basis of how they want to present the results (e.g., indicators are converted to a common scale with range [0, 1] or to a common scale where a reference is set equal to 100). However, the normalization method has a strong impact on results for two important reasons: (1) it creates a 'correspondence system' between different variables (McGranahan, 1970), (2) it assigns them 'implicit weights' (Booysen, 2002). The 'correspondence system' defines what level of any one variable tends to go with (corresponds to) given levels of other variables (e.g., what level of "Life expectancy" should be found normally with a given level of "Gross national income per capita" or of "Hospital beds per 100,000 inhabitants" and vice versa). These correspondences are particularly important when a non-compensatory approach (i.e., an approach based on the concept of 'unbalance' or disequilibrium among variables) is followed (Casadio Tarabusi and Guarini, 2013)[2]. In such a case, in fact, it is necessary to define what is meant by 'balance' and this definition depends on the normalization method adopted. For example, if indicators are converted to a common

[1] The paper is the result of the common work of the authors: in particular M. Mazziotta has written Sections 2.1-2.3 and 4 and A. Pareto has written Sections 1, 2.4 and 3.

[2] We say that an approach is non-compensatory when it is not full-compensatory. Note that the arithmetic mean is full-compensatory, the geometric mean is partially-compensatory and the minimum is non-compensatory.

scale with range [0, 1], then the set of the maximum values and the set of the minimum values will be considered 'balanced'[3], whereas the set of the mean values could be considered 'unbalanced'. By contrast, if indicators are converted to a common scale where the mean value is set equal to 100, then the set of the mean values will be considered 'balanced'; whereas the set of maximum values and the set of minimum values could be considered 'unbalanced'. Moreover, the range (i.e., the variability) of normalized indicators acts as implicit weight during the aggregation. The wider the minimum and maximum values are apart, the higher the implicit weighting and vice versa. For example, if indicators are converted to a common scale with range [0, 1], then implicit weights are practically the same. By contrast, if indicators are converted to a common scale where the mean value is set equal to 100, but a normalized indicator ranges between 99 and 101 and other ranges between 50 and 200, the composite index will be dominated by the second indicator.

Therefore, an incorrect choice of the normalization method can lead to an unacceptably large degree of distortion of results. This paper discusses the differences among the main normalization methods, and proposes an alternative, denoted as 'Re-scaling with a reference' that combines the advantages of some of them. Some issues on the 'effect' of normalization and suggestions for a correct choice of the normalization method are also reported.

## 2. Normalization methods and their properties

### *2.1. Standardization (or transformation in z-scores)*

Standardization is the method most commonly used by statisticians. It converts variables to a common scale with a mean of 0 and a standard deviation of 1. For a generic unit *i* and variable *j*, the formula is:

$$y_{ij} = \frac{x_{ij} - \mathrm{M}_{x_j}}{\mathrm{S}_{x_j}} \qquad (1)$$

where $x_{ij}$ is the original value of variable *j* for unit *i*, and $\mathrm{M}_{x_j}$ and $\mathrm{S}_{x_j}$ are, respectively, the mean and standard deviation of variable *j*. If variable *j* has negative polarity[4], then formula (1) can be multiplied by -1. Transformed scores are known as z-scores and most of them (i.e., about 90%) lay between the values −3 and +3,

[3] Note that this is a strong and less plausible assumption, because the minimum and the maximum of a distribution often are 'outliers' (i.e., 'abnormal' values).

[4] The polarity of a variable is the sign of the relation between the variable and the phenomenon to be measured (Mazziotta and Pareto, 2017).

regardless of the shape of the original distribution. They may be further adjusted if calculations yield awkward values. For example, we can multiply each score by 10 and add 100 to obtain positive and more visually manageable scores (Booysen, 2002)[5].

Standardized variables have the same variance and similar (but not equal) range. Standardization has the advantage of 'centering' the variables around a common average and 'normalizing' their variability (Abdi, 2007).

The method allows to compare the values of the units, both in space and time, with respect to the mean and variance of the distribution. So an increase in the standardized value of a given unit, from one period to another, indicates that the original value has increased compared to the new mean and variance (which could also be decreased), but does not necessarily correspond to an increase of the original value. This can be a limitation of the method when different periods have to be compared.

### *2.2. Re-scaling (or Min-Max method)*

Re-scaling is the method most commonly used by sociologists. It converts variables to a common scale with range [0, 1]. For a generic unit *i* and variable *j*, the formula is:

$$y_{ij} = \frac{x_{ij} - \min_i(x_{ij})}{\max_i(x_{ij}) - \min_i(x_{ij})} \tag{2}$$

where $x_{ij}$ is the original value of variable *j* for unit *i*, and $\min_i(x_{ij})$ and $\max_i(x_{ij})$ are, respectively, a minimum and a maximum that represent the possible range of variable *j* (*goalposts*). If variable *j* has negative polarity, then the complement of (2) with respect to 1 is calculated[6]. Also, in this case, transformed scores may be further adjusted to facilitate reading. For example, we can multiply each score by 1,000 in order to obtain values between 0 and 1,000.

Re-scaled variables have the same range and similar (but not equal) variance. Re-scaling has the disadvantage of 'not centering' the variables around a common average. This can be a big problem, if the researcher is interested in constructing a non-compensatory composite index with an unbalance adjustment method (Casadio Tarabusi and Guarini, 2013).

The particularity of the method consists in the possibility of setting the goalposts

[5] In this case we obtain a common scale with a mean of 100 and an approximate range of 70-130.

[6] The 'complement with respect to 1' is the number to add to make 1.

regardless of the values of the variable in a given period[7]. This allows to compare the values of the units, both in space and time, with respect to a common reference (the goalposts) that does not change from one period to another (in contrast to the standardization, where the reference are the mean and variance of each period). So an increase in the normalized value of a given unit, from one period to another, corresponds to an increase of the original value.

### *2.3. Indicization (or transformation in index numbers)*

Indicization[8] is the method most commonly used by economist. It converts variables to a common scale where a reference is set equal to 1. For a generic unit *i* and variable *j*, the formula is:

$$y_{ij} = \frac{x_{ij}}{x_{oj}} \tag{3}$$

where $x_{ij}$ is the original value of variable *j* for unit *i*, and $x_{oj}$ is the reference value (or base) for variable *j* – generally, the maximum, the mean or an external benchmark. If variable *j* has negative polarity, then a non-linear transformation, such as the reciprocal, could be preliminarily applied; however, indicization is recommended only for indicators with positive polarity[9]. Formula (3) can also be adjusted to set the base equal to 100 or 1,000.

Index numbers have the same reference (e.g., the mean), but can have very different range and variance, because they have the same CV[10] of original variables. Therefore, indicization has the disadvantage of 'not normalizing' the variability of variables and introducing implicit weights. This can be a big problem, if the researcher is interested in constructing a composite index with equal or other explicit weights (Mazziotta and Pareto, 2017).

The base of index numbers can be set regardless of the values of the variable in a given period. This allows to compare the values of the units, both in space and time, with respect to a common reference (the base) that does not change in the various periods. So an increase in the index number of a given unit, from one period to another, corresponds to an increase of the original value.

---

[7] Usually, the goalposts are the minimum and maximum of the variable over an extended period of time, in order to take into account its evolution. Alternatively, they can be fixed by experts.

[8] This method is also known as 'Distance from a reference' (OECD, 2008).

[9] A non-linear transformation of variables causes distortions in the data, because correlations among transformed variables are not equal to correlations among original variables.

[10] The coefficient of variation (CV) is a measure of dispersion, often expressed as a percentage, defined as the ratio between standard deviation and mean.

### *2.4. Re-scaling with a reference (or Constrained Min-Max method)*

Re-scaling with a reference is a method that 'normalizes' the variables – similarly to re-scaling – but use a common reference that allows to 'center' them – like indicization. It converts variables to a common scale where a reference is set equal to 0 e the range is 1. For a generic unit *i* and variable *j*, the formula is:

$$y_{ij} = \frac{x_{ij} - x_{oj}}{\max_i(x_{ij}) - \min_i(x_{ij})} \tag{4}$$

where $x_{ij}$ is the original value of variable *j* for unit *i*, $\min_i(x_{ij})$ and $\max_i(x_{ij})$ are, respectively, a minimum and a maximum that represent the possible range of variable *j* (goalposts) and $x_{oj}$ is the reference value for variable *j*. If variable *j* has negative 'polarity', then formula (4) can be multiplied by -1. Transformed scores may be further adjusted as in the previous methods. For example, we can multiply each score by 60 and add 100 to obtain the normalization formula used in the Adjusted Mazziotta-Pareto Index (Mazziotta and Pareto, 2016).

Transformed scores have the same reference (e.g., the mean) and equal range. This allows to have the advantages of index numbers without introducing implicit weights.

Similarly to re-scaling and indicization, the goalposts and the reference can be set regardless of the values of the variable in a given period. This allows to compare the values of the units, both in space and time, with respect to a common reference that remains constant over all periods. So an increase in the normalized value of a given unit, from one period to another, corresponds to an increase of the original value.

## 3. Comparing normalization methods

### *3.1. The 'effect' of normalization*

To illustrate the effect of normalization on the distributions of individual indicators, we consider the three indicators represented in Figure 1. The first has an exponential distribution (Exp), the second has a normal distribution (Nor) and the third has a Beta distribution (Bet). The indicators have different mean and variance, as they represent the most disparate phenomena. Figure 2 shows the distributions of normalized indicators by 'standardization' (a), 're-scaling' (b), 'indicization' with base=mean set to 100 (c), 're-scaling with a reference' with reference=mean set to 100 and range 60 (d).

**Figure 1 –** *Individual indicators with different distributions.*

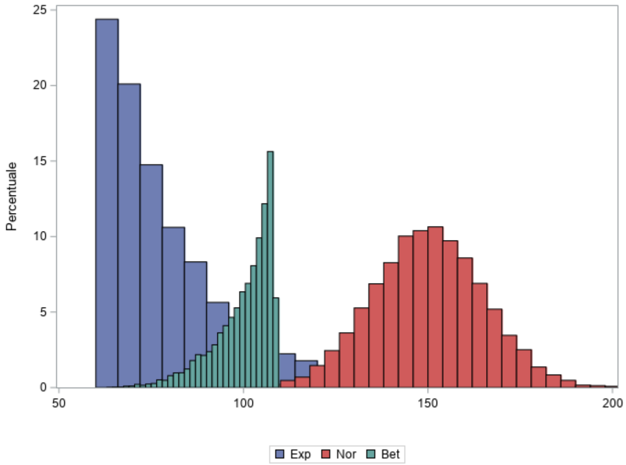
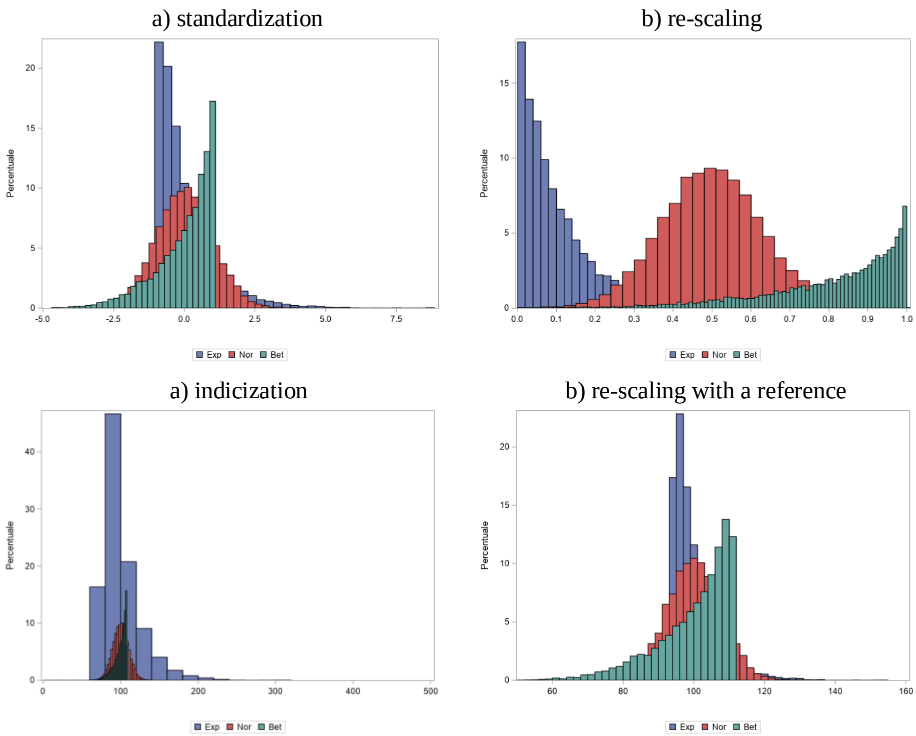



**Figure 2 –** *Normalized indicators with different distributions.*



As we can see, the distributions of indicators transformed into z-scores (Figure 2.a) are 'centered' around the origin (mean=0) and 'elongated' or 'shortened' to have the same variability (variance=1). So, no implicit weighting is introduced.

Re-scaling also makes the variances more homogeneous (but not equal), bringing all the values into a common interval (Figure 2.b). However, the distributions of indicators are not 'centered' and this leads to the loss of a common reference value, such as the mean. It follows that equal normalized values (i.e. balanced normalized values) can correspond to very unbalanced original values. For example, the normalized value 0.2 for the Exp indicator corresponds to a high original value; whereas for the Nor and Bet indicators it corresponds to a very low original value. Therefore, the use of a simple re-scaling for aggregating individual indicators with an unbalance adjustment method, such as the geometric mean, can lead to biased results. Moreover, the normalized value 0.5 is the mean of the range, but not of distributions, and then it cannot be used as a reference for reading results (e.g., if the normalized value of a given unit is 0.3., we cannot know if its original value is above or below the mean).

Indicization with the mean as a base set to 100 (Figure 2.c) 'centers' all distributions around the mean, but do not 'normalize' their variability (e.g., the range of the Bet indicator is very short, whereas the range of the Exp indicator is very large). So, a significant implicit weighting is introduced when different indicators are compared and aggregated.

Re-scaling with a reference is very similar to standardization when the mean is chosen as a reference (Figure 2.d). In fact, it 'centers' indicators (like indicization) and 'normalizes' them (like re-scaling). Nevertheless, it allows to keep fixed the goalposts and the reference when different periods have to be compared, contrary to standardization. So, the researcher can compare indicators over an extended period of time with respect to a reference, without introducing an implicit weighting.

### *3.2. The correspondence grid*

A correspondence grid emerges when for each indicator the original values that are identified with each level of the common scale are shown on a table. For example, at level 2 of the correspondence grid of standardization are given the original values (for each indicator) that have the standardized value of 2 and that correspond to each other. The result is a list of 'correspondence points' each of which represents a set of original values that will be considered 'balanced'. The correspondence grid must be carefully constructed and evaluated by the researcher, because it can yield an 'artificial' or 'inconsistent' model of balance of original indicators. Table 1 shows the correspondence grids for the four normalizations of Figure 2.

**Table 1 –** *Correspondence grid for different normalization methods.*

| Standardization | | | | Re-scaling | | | |
|---|---|---|---|---|---|---|---|
| **Scale** | **Exp** | **Nor** | **Bet** | **Scale** | **Exp** | **Nor** | **Bet** |
| 2.5 | 128.5 | 187.6 | 119.4 | 1.0 | 250.0 | 209.5 | 108.4 |
| 2.0 | 118.8 | 180.1 | 115.5 | 0.9 | 231.0 | 197.7 | 103.9 |
| 1.5 | 109.1 | 172.6 | 111.6 | 0.8 | 212.1 | 185.8 | 99.4 |
| 1.0 | 99.4 | 165.0 | 107.8 | 0.7 | 193.1 | 174.0 | 94.9 |
| 0.5 | 89.7 | 157.5 | 103.9 | 0.6 | 174.2 | 162.2 | 90.5 |
| 0.0 | **80.0** | **150.0** | **100.0** | 0.5 | 155.2 | 150.3 | 86.0 |
| -0.5 | 70.3 | 142.5 | 96.1 | 0.4 | 136.3 | 138.5 | 81.5 |
| -1.0 | 60.6 | 135.0 | 92.2 | 0.3 | 117.3 | 126.7 | 77.0 |
| -1.5 | 50.9 | 127.4 | 88.4 | 0.2 | 98.4 | 114.8 | 72.5 |
| -2.0 | 41.2 | 119.9 | 84.5 | 0.1 | 79.4 | 103.0 | 68.1 |
| -2.5 | 31.5 | 112.4 | 80.6 | 0.0 | 60.5 | 91.2 | 63.6 |
| **Indicization** | | | | **Re-scaling with a reference** | | | |
| **Scale** | **Exp** | **Nor** | **Bet** | **Scale** | **Exp** | **Nor** | **Bet** |
| 200 | 160.0 | 300.0 | 200.0 | 125 | 158.9 | 199.3 | 118.7 |
| 180 | 144.0 | 270.0 | 180.0 | 120 | 143.2 | 189.4 | 114.9 |
| 160 | 128.0 | 240.0 | 160.0 | 115 | 127.4 | 179.6 | 111.2 |
| 140 | 112.0 | 210.0 | 140.0 | 110 | 111.6 | 169.7 | 107.5 |
| 120 | 96.0 | 180.0 | 120.0 | 105 | 95.8 | 159.9 | 103.7 |
| **100** | **80.0** | **150.0** | **100.0** | **100** | **80.0** | **150.0** | **100.0** |
| 80 | 64.0 | 120.0 | 80.0 | 95 | 64.2 | 140.1 | 96.3 |
| 60 | 48.0 | 90.0 | 60.0 | 90 | 48.4 | 130.3 | 92.5 |
| 40 | 32.0 | 60.0 | 40.0 | 85 | 32.6 | 120.4 | 88.8 |
| 20 | 16.0 | 30.0 | 20.0 | 80 | 16.8 | 110.6 | 85.1 |
| 0 | 0.0 | 0.0 | 0.0 | 75 | 1.1 | 100.7 | 81.3 |

There are a number of points of interest in the table. In particular, all normalization methods, except re-scaling, consider ‘balanced’ the set of mean values. Standardization considers balanced a set of values when they are ‘equidistant’ from the mean in terms of standard deviations. For example, at level 3 of the correspondence grid are given all the original values that are equal to the mean plus 3 standard deviations. Indicization considers balanced a set of values when they are ‘equidistant’ from the mean (the base) in percentage terms. For example, at level 200 of the correspondence grid are given all the original values that are double the mean. In this case, also the set of null values is considered balanced, so the transformation in index numbers should be applied only to variables that have an ‘absolute zero’ point (e.g., “height” and “weight”). Re-scaling with a reference is similar to standardization, but it considers balanced a set of values when they are ‘equidistant’ from the mean (the reference) with respect to the range. Finally, classical re-scaling considers balanced the two sets of extreme values and it creates 'artificial' correspondence points (i.e., artificially balanced sets of values) in the middle. For example, the set of the mean values corresponds approximately to the set of normalized values (0.1; 0.5; 0.8), and then it will be considered very unbalanced. The greater the differences between the indicator distributions, the greater the distortion of the correspondence points.

### *3.3. The implicit weighting*

The second issue that the researcher should take into account is the implicit weighting induced by normalization. Suppose we have the following indicators:

- "Life expectancy (years)" ($X_1$) with a mean of 80 and a standard deviation of 1 (approximate range of 77-83 and CV of 1.25%);
- "Hospital beds per 100,000 inhabitants" ($X_2$) with a mean of 1,000 and a standard deviation of 200 (approximate range 400-1,600 and CV of 20%).

If we consider a significant change in $X_1$ from 80 to 82 years (z-score changes from 0 to 2) and a not significant change in $X_2$ from 1,000 to 1,025 beds (z-score changes from 0 to 0.125), the percentage variations of the two indicators coincide and are equal to 2.5% (both index numbers change from 100 to 102.5). So, normalizing by indicization, the two variations will be considered of equal importance. By contrast, if we consider two variations of equal importance where both indicators increase by 1 standard deviation, that is $X_1$ changes from 80 to 81 and $X_2$ changes from 1,000 to 1,200 (both z-scores change from 0 to 1), the percentage variation of $X_1$ is 1.25% (index number changes from 100 to 101.25), whereas the percentage variation of $X_2$ is 20% (index number changes from 100 to 120). So, normalizing by indicization, the variation of the indicator with greater CV ($X_2$) will be considered more important than the variation of the indicator with less CV ($X_1$).

Table 2 provides an example of aggregating z-scores and index numbers for indicators $X_1$ and $X_2$. The aggregation function is a simple arithmetic mean (full compensatory approach). Consider unit 1 and unit 5. In unit 1, $X_1$ is 1.5 standard deviations below the mean (78.5 years), and $X_2$ is 1.5 standard deviations above the mean (1,300 beds). Conversely, in units 5, $X_1$ is 1.5 standard deviations above the mean (81.5 years), and $X_2$ is 1.5 standard deviations below the mean (700 beds). If we assign the same importance to the indicators, the two units are in a similar situation and therefore must have the same position in a ranking according to the mean of normalized values. However, if we use index numbers, $X_2$ has a greater weight than $X_1$ in the computation of the mean, and unit 1 obtains a greater score than unit 5 (114.1 vs. 85.9) because it has the higher value (1,300) on $X_2$. On the contrary, if we use z-scores, the two units have the same score (0.0).

**Table 2 –** *Example of implicit weighting*

| Unit | Original indicators | | Z-scores | | | Index numbers | | |
|---|---|---|---|---|---|---|---|---|
| | **$X_1$** | **$X_2$** | **$X_1$** | **$X_2$** | **Mean** | **$X_1$** | **$X_2$** | **Mean** |
| 1 | 78.5 | 1,300 | -1.5 | 1.5 | 0.0 | 98.1 | 130.0 | 114.1 |
| 2 | 80.5 | 1,100 | 0.5 | 0.5 | 0.5 | 100.6 | 110.0 | 105.3 |
| 3 | 80.0 | 1,000 | 0.0 | 0.0 | 0.0 | 100.0 | 100.0 | 100.0 |
| 4 | 79.5 | 900 | -0.5 | -0.5 | -0.5 | 99.4 | 90.0 | 94.7 |
| 5 | 81.5 | 700 | 1.5 | -1.5 | 0.0 | 101.9 | 70.0 | 85.9 |
| **Mean** | **80.0** | **1,000** | **0.0** | **0.0** | | **100.0** | **100.0** | |
| **Std** | **1.0** | **200** | **1.0** | **1.0** | | **1.25** | **20.0** | |
| **CV** | **1.25** | **20** | | | | **1.25** | **20.0** | |

This simple example shows that indicization makes indicators independent of the unit of measurement, but not of their variability (original values and index numbers have the same CV). The higher the CV, the greater the weight, in terms of normalized values, on the aggregation function. Therefore, in order to assign the same 'importance' to each indicator, the researcher should apply a normalization method that also makes the indicators independent of the variability.

## 4. Conclusions

Values measured with different units of measurement alone do not explain so much, as each value is meaningful only relative to the mean and variability of the distribution. Therefore, how to compare a life expectancy of 82 years with 800 hospital beds per 100,000 inhabitants? Values from different distributions can be normalized in order to provide a way of comparing them that includes consideration of their respective distributions. This is normally done by transforming the values into z-scores which are expressed as standardized deviations from their means (Abdi, 2007). Nevertheless, standardization is not the best method for comparisons over different periods. In this paper, we propose a similar normalization method, denoted as 'Re-scaling with a reference' (or 'Constrained Min-Max method') that can be used when different periods have to be compared.

It is good to specify that, in general, the perfect normalization method does not exist. Each method has strengths and weaknesses and the choice depends on the aims of the research and/or on the aggregation function used for constructing the composite index. The paper shows that when choosing the normalization method an implicit weighting must always be avoided and, above all, a realistic 'correspondence grid' (not artificial or meaningless) must be constructed in order to consider a correct 'balancing model' of the values.

For this reason, many composite indices based on the classical Min-Max method should be revised. This is the case of the Human Development Index - HDI (UNDP, 2019), where however the goalposts are 'reasoned' and have been set by experts. In other cases, as in the composite indices summarizing the SDGs, published recently by Istat (Istat, 2020), the computation procedure is based on a re-scaling with 'observed' goalposts and percentage variations of indicators with different CV are considered of equal importance (-80%; +80%). This can lead to strong distortions of the 'balancing model' of indicators and, therefore, to incorrect or misleading results. In this regard, we must remember that the first criterion to be followed in the construction of a composite index (as in any statistical model) is the *principle of parsimony* (Mazziotta e Pareto, 2020). This principle states that the composite index must be as simple as possible, to allow an easy interpretation of results, both in space and time. In order to construct a composite index as simple as possible, the processing to be performed on the data must be reduced to the minimum necessary. Therefore, only one normalization method must be applied to the data matrix and no further transformation of the obtained scores should be carried out, as they are already normalized (for example, it does not make sense to calculate index numbers on values already normalized with re-scaling, because the percentage variations of these values have no meaning).

In conclusion, the construction of a composite index must follow a precise work paradigm and international literature is unanimous in this sense. Methodological shortcuts or even fanciful approaches, such as normalizing data several times, are absolutely to be avoided since a composite index has a great responsibility: measuring multidimensional phenomena to better understand the reality.

## References

ABDI H. 2007. Z-scores. In SALKIND N. (Ed) *Encyclopedia of Measurement and Statistics*. Thousand Oaks: Sage.

BOOYSEN F. 2002. An overview and evaluation of composite indices of development, *Social Indicators Research*, Vol. 59, pp. 115-151.

CASADIO TARABUSI E., GUARINI G. 2013. An Unbalance Adjustment Method for Development Indicators, *Social Indicators Research*, Vol. 112, pp. 19-45.

ISTAT 2020. *Rapporto SDGs 2020*. Roma: Istituto nazionale di statistica.

MAZZIOTTA M., PARETO A. 2016. On a Generalized Non-compensatory Composite Index for Measuring Socio-economic Phenomena, *Social Indicators Research*, Vol. 127, pp. 983-1003.

MAZZIOTTA M., PARETO A. 2017. Synthesis of Indicators: The Composite Indicators Approach. In MAGGINO F. (Ed) *Complexity in Society: From Indicators Construction to their Synthesis*, Social Indicators Research Series 70, Cham: Springer, pp. 159-191.

MAZZIOTTA M., PARETO A. (a cura di) 2020. *Gli indici sintetici*. Torino: Giappichelli.

MCGRANAHAN D. 1970. The interrelations between social and economic development, *Social Science Information*, Vol. 9, pp. 61-77.

OECD 2008. *Handbook on Constructing Composite Indicators. Methodology and user guide*. Paris: OECD Publications.

SALZMAN J. 2003. *Methodological Choices Encountered in the Construction of Composite Indices of Economic and Social Well-Being*. Ottawa: Center for the Study of Living Standards.

UNDP (2019). *Human Development Report 2019*. New York: United Nations Development Programme.

## SUMMARY

### Everything you always wanted to know about normalization (but were afraid to ask)

The solution to the problem of normalization of variables with different units of measurement is of primary interest in data analysis. Most of the reports in the literature present a wide variety of normalization methods, but do not explain how to choose the ‘right’ method. Researchers cannot avoid the question simply by choosing the normalization on the basis of how they want to present the results, as each method has its pros and cons. In this paper, a comparison among the main normalization methods is presented and an alternative method, denoted as ‘Re-scaling with a reference’ is proposed. Some issues on the ‘effect’ of normalization and suggestions for a correct choice of the normalization method are also reported.

---

Matteo MAZZIOTTA, Istat, mazziott@istat.it
Adriano PARETO, Istat, pareto@istat.it

# SOCIAL INDICATORS TO MEASURE THE WELL-BEING OF THE POPULATION. BENCHMARKING COUNTRIES

Vincenzo Marinello, Guglielmo L.M. Dinicolò, Chiara Di Puma

## 1. Introduction

The various criticisms made about GDP have facilitated the development of a set of indicators aimed at starting new processes for measuring well-being and subjective well-being.

The aim of this work is to provide an examination of the main well-being indicators in the literature, and their trends with a focus on the Italian context and the main European partner countries.

Within the framework of the United Nations Development Programme, a social indicator, the HDI, is developed, which includes different evaluation dimension with respect to GDP, considering, in fact, social and environmental aspects in the process of evaluating the well-being of a State; it also includes health and education. In Italy in 2010, Istat launched the Bes project to measure Equitable and Sustainable Well-being, with the aim of evaluating the progress of society not only from an economic, but also from a social and environmental point of view. To this end, the traditional economic indicators, GDP first of all, have been integrated with measures of the quality of people's life and of the environment. The purpose is to prepare statistical reports and analysis necessary for the verification of Fair and Sustainable Welfare at the local level to support local authorities in the process of developing intervention strategies. BES project has certainly represented an important moment in terms of analysis and measurement of well-being; it has also given rise to many initiatives at national and local level, such as the "Urbes" and "BES of the provinces" projects, which propose and suggest more precise measures of Equitable and Sustainable Welfare at territorial level.

## 2. Literature review

The contributions provided by recent literature on well-being and sustainable development, highlight how it is possible to pursue the objective of growth by damaging the social and ecological aspects of a country, favouring all those factors of an economic nature (Spangenberg and Lorek, 2014). The relationship between the environment and social objectives, or green society; inclusive growth that precludes the link between growth and social factors (Gupta et al., 2015). In a broad perspective, social inclusion implies the need to take into consideration the economy of less developed nations, developing nations and the economic realities of post-conflict countries (Okafor, 2008). The classic dichotomous bipartition on the topic varies between two extremes in which the authors have tried to consider whether higher equity facilitates (Schulze et al., 2018; Ostry et al., 2014) or penalizes GDP growth in the long term (Maibom and Andersen, 2016). Moreover, the studies carried out on “behavioural economics” have shown that each individual subject does not always operate maximizing their economic interest as a result of the presence of information asymmetries that significantly hinder the decision-making process (Stutzer et al., 2007; Verhofstadt et al., 2011).

Masur et al. (2010), on the basis of the framework of classical utilitarian theory, have demonstrated how the measurement of well-being, cannot be separated from the determination, analysis and study of the concept of “subjective well-being”. In the latter case, it is possible to count the indices of happiness, which have the advantage of providing an aggregate value of individual well-being. Undoubtedly the so-called developed nations show a higher happiness of their populations than the developing countries (Wolfers and Leigh, 2006; Diener and Oishi, 2000). But happiness in emerging countries is constantly growing (Oishi et al., 2011). In these countries it follows a factor that affects happiness, i.e. the economic condition of each individual, whereas in rich countries the elements that condition the level of happiness are the emotional relationships (Qu, 2015).

In agreement with Doyle (2018) one of the main concerns for the future years is the rapid overheating of the planet; so, this issue must be considered as the starting point for other sustainability hypotheses. Forecasts of economic effects and climate change are based on the “opulence principle”, while risks caused by a slow-down in growth are underestimated. As this principle shows various limitations, the estimates are approximate, not considering the costs of dealing with adverse climatic conditions and limiting the resulting economic inequalities. For this reason, several estimates based on this criterion have been revised (Ackerman and Stanton, 2008) with the aim of including these types of costs in the new forecasts. In any case, the ethical aspect of sustainability bypasses the concept of human well-being, attributing more responsibility to protecting the environment with particular attention to the protection of biodiversity (Knox, 2017; Boyce et al., 2010).

Among the other deficiencies of the GDP as an indicator of human well-being, there is the omission to account for environmental costs and values of non-market goods of economic activity (Costanza et al. 2017). Some authors argue that in a reality where ecological constraints are becoming increasingly evident, an increase in GDP can produce a so-called "uneconomic growth" (Lawn, 2016) produced by an increase in social costs that exceed the economic benefits (Bache and Reardon, 2016; Bleys and Whitby, 2015). Over the years, many attempts have been made to implement new indicators and improve existing ones, focusing on the correlation between indicators and their actual participation and application in decision-making processes (Hayden and Wilson, 2017). Some researchers argue that alternative indices to GDP can be a useful way to show and interpret public costs as "smart investments" (Daly and McElwee, 2014).

## 3. Human Development Index in the Italian context

The most significant indicator to measure the level of long-term well-being is the Human Development Index (HDI), which highlights the quality of life and human development in three key perspectives: a long and healthy life, access to knowledge and a "decent" standard of living. The first dimension measured is life expectancy; it is referred to the average number of years of a newborn child. In other words, it is an indicator that allows to determine the number of people, of different ages, who die in the reference to the born year to have a view of the mortality characteristics for the population. The second dimension, access to knowledge, is measured by the education index. It indicates the level of education and literacy rate and is constructed on the basis of adult literacy; the latter dimension expresses the attitude of registered students (students and pupils) of individuals over the age of 25, the percentage of students in higher and secondary education institutions and the number of people with higher education. Living standards are measured based on gross national income (GNI) per capita. This index is useful to show how countries with the same level of GNI per capita obtain different levels of human development.

It is essential to affirm that HDI could integrates the results obtained through the use of the traditional macroeconomic indicator. In 2018, the ranking of human development index, classify Italy in the country with a high level; this ranking places Italy in 29th place among 189 countries, with a rating of 0.883 (the index range is from 0 to 1). Between 2010 and 2018 the value of Italian HDI went from 0.871 to 0.883, with an increase of 1.1% and an average annual growth of 0.17%.

To better understand these aspects in Table 1, was show the level of the index in Italy between the 2015 and 2018 (there has also been included the 2010 like base year) (Figure 1).

Comparing the value of HDI in Italy with that of other main economic partner countries (Spain, United Kingdom, France, Malta) like a benchmark, in Table 2 it is possible to observe how Italy shows a low value of this index equal to 0.883

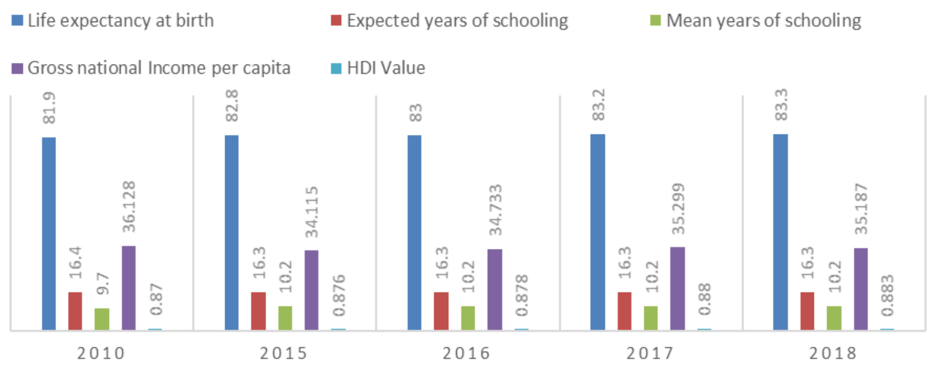
compared to the average of other countries that have a high index of human development. Despite the fact that Italy has a higher GNI per capita than Spain, it is possible to note that the value of the Human Development Index is lower; on the other hand, Montenegro, Greece Cyprus Portugal and Turkey show very low performance in HDI index and in other dimension collected in Table 2.

**Table 1 –** *Human Development Index in Italy from 2010 to 2018.*

| | Life expectancy at birth | Expected years of schooling | Mean years of schooling | Gross National Income per capita | HDI value |
|---|---|---|---|---|---|
| 2010 | 81.9 | 16.4 | 9.7 | 36,128 | 0.871 |
| 2015 | 82.8 | 16.3 | 10.2 | 34,115 | 0.875 |
| 2016 | 83.0 | 16.3 | 10.2 | 34,733 | 0.878 |
| 2017 | 83.2 | 16.3 | 10.2 | 35,299 | 0.881 |
| 2018 | 83.3 | 16.3 | 10.2 | 35,187 | 0.883 |

*Source: United Nations Development Programme – 2019 own elaboration.*

**Figure 1 –** *Value of each dimension of the HDI.*




*Source: United Nations Development Programme – 2019, own elaboration.*

### *3.1. Corrections to the Human Development Index*

HDI is an average of a country's human development outcomes, but it does not take into account the inequalities of a country's population. The United Nation Development Programme, in its Human Development Report (HDR), to overcome this gap, in 2010 introduced the Inequality-adjusted Human Development Index (IHDI), which is the index of human development net of inequalities. The IHDI combines a country's average achievements in health, education and income with how those achievements are distributed among country's population by "discounting" each dimension's average value according to its level of inequality. Thus, the IHDI is distribution-sensitive average level of human development. As the inequalities of a country decrease, the level of human development of the population increases.

**Table 2 –** *Values of the Human Development Index of Italy and other countries compared. Year 2018to 2018.*

| | HDI value | Life expectancy at birth | Expected years of schooling | Mean years of schooling | GNI per capita |
|---|---|---|---|---|---|
| Italy | 0.883 | 83.2 | 16.3 | 10.2 | 35.299 |
| Spain | 0.893 | 83.4 | 17.9 | 9.8 | 35.041 |
| United Kingdom | 0.920 | 81.7 | 17.4 | 12.9 | 39.116 |
| France | 0.891 | 82.5 | 15.5 | 11.4 | 40.511 |
| Germany | 0.939 | 81.2 | 17.1 | 14.1 | 46.946 |
| Greece | 0.872 | 82.1 | 17.3 | 10.5 | 24.909 |
| Cyprus | 0.873 | 80.8 | 14.7 | 12.1 | 33.110 |
| Malta | 0.885 | 82.4 | 15.9 | 11.3 | 34.795 |
| Montenegro | 0.816 | 76.8 | 15.0 | 11.4 | 17.511 |
| Portugal | 0.850 | 81.9 | 16.3 | 9.2 | 27.935 |

*Source: United Nations Development Programme – 2019 own elaboration.*

**Table 3 –** *Values of IHDI of Italy and other countries compared. Year 2018.*

| | IHDI value | Overall loss (%) | Human ineq. coeff (%) | Inequality in life expect at birth (%) | Ineq. in educ (%) | Ineq. in income (%) |
|---|---|---|---|---|---|---|
| Italy | 0.776 | 12.1 | 11.8 | 3.1 | 11.0 | 21.3 |
| Spain | 0.765 | 14.3 | 14.0 | 3.0 | 17.1 | 21.9 |
| United Kingdom | 0.845 | 8.2 | 8.0 | 4.1 | 2.8 | 17.0 |
| France | 0.809 | 9.2 | 9.1 | 3.8 | 9.1 | 14.4 |
| Germany | 0.861 | 8.3 | 8.1 | 3.8 | 2.7 | 17.7 |
| Greece | 0.766 | 12.2 | 11.9 | 3.5 | 12.8 | 19.5 |
| Cyprus | 0.788 | 9.7 | 9.6 | 3.6 | 11.0 | 14.3 |
| Malta | 0.815 | 8.0 | 7.9 | 4.6 | 6.7 | 12.5 |
| Montenegro | 0.746 | 8.6 | 8.5 | 3.6 | 7.4 | 14.6 |
| Portugal | 0.742 | 12.7 | 12.4 | 3.5 | 15.8 | 18.1 |

*Source: United Nations Development Programme – 2019 own elaboration.*

The data for 2018 it is shown in table 3. As it possible to read Italy's HDI is equal to 0.883, consequently, subtracting the value of inequality, the same index decreases to 0.776, with a 12.1% decrease. Spain and the United Kingdom show an overall loss of 14.0% and 8.0% respectively. The average loss resulting from the inequality of each country is 10.7, while for OECD countries it is 11.7%. Human inequality coefficient for Italy is 11.8% much higher than the United Kingdom and France which show respectively a coefficient of 8.0% and 9.1%. Comparing the other European countries, Italy register better performance also considering the Mediterranean region; Malta and Montenegro have a HDI index equal to 7.9% and 8.5%.

The difference between the IHDI and HDI is the human development cost of inequality, also termed – the overall loss to human development due to inequality. The IHDI allows a direct link to inequalities in dimensions, it can inform policies towards inequality reduction, and leads to better understanding of inequalities across population and their contribution to the overall human development cost; a high inequality can determine negative consequences for social cohesion, the quality of institutions and policies, slowing down human progress.

## 4. Better Life Index (BLI). Benchmarking between Italy and other European countries

Analysing the level of well-being in Italy through the Better Life Index, it is possible to observe how unstable trend was recorded in relation to the various indicators that make up this index. Italy is above the mean of the other OECD countries in terms of income, wealth, work-life balance, civic commitment, social relations and health status, while it has a low average for housing, subjective well-being, quality of the environment, level of employment and education. in order to better understand the positioning of Italy in the European panorama, in the analysis of this indicator it was decided to introduce only United Kingdom, Turkey, Greece, France, Germany, Portugal and Spain to obtain a broader perspective of continental and Mediterranean Europe areas.

In Italy, the average per capita income is significantly lower than the OECD average of 33,604 USD per year compared to 26,588 USD in Italy. In detail, as can be seen from Table 4, Germany and France show a higher mean disposable income per capita than Italy, Greece and Spain. Income is an important component because the level of economic wealth has a decisive influence on choices in terms of education and health care. In terms of employment, 58% of the working population in Italy is legally employed, while it is much higher in Germany and France, where the employment rate is 75% and 65% respectively. Spain also outperforms Italy with 62% employment rate, only Greece and Turkey showing an even lower level of employment than Italy. As Table 4 shows, a country with a high employment rate achieves income benefits, facilitates social inclusion and promotes the development of work skills. In fact, Germany and France with a high employment rate are among the countries with the highest average disposable income per capita.

An important element that allows a country to achieve an advantage in terms of social and economic well-being is a high level of education. The education variable indicates the percentage of adult population, between 25 and 64 years of age, that have at least an upper secondary education degree; an excellent level of education makes it easier to find a job and to have a substantial income. All this is confirmed by the data reported in the table below, where countries such as Germany and France, which report a high level of education, 75% and 65% respectively, result in a high employment rate of 87% and 78%.

**Table 4 –** *Some indicators compared to the Better Life Index – Year 2018.*

| | Mean disposable income per capita (USD) | Employment (%) | Education (%) | Health (%) | Life satisfaction (score from 0 to 10) |
|---|---|---|---|---|---|
| Italy | 26.588 | 58 | 61 | 83 | 6 |
| Spain | 23.999 | 62 | 59 | 83 | 6.3 |
| United Kingdom | 28.715 | 75 | 81 | 81 | 6,8 |
| France | 31.304 | 65 | 78 | 82 | 6.5 |
| Germany | 34.297 | 75 | 87 | 81 | 7.7 |
| Greece | 17.700 | 53 | 73 | 82 | 5.4 |
| Portugal | 21.203 | 68 | 48 | 81 | 5.4 |
| Turkey | 18.302 | 52 | 39 | 78 | 5.5 |

*Source: OCSE, own data processing.*

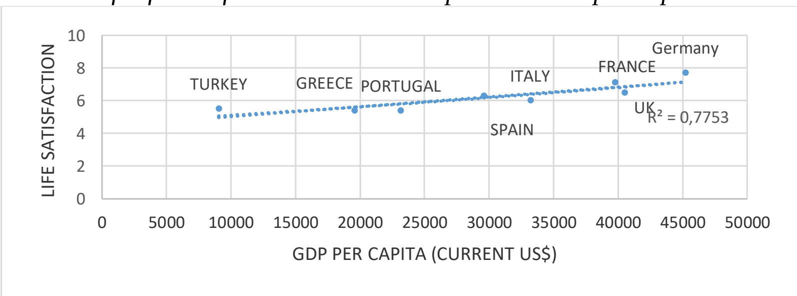
In recent years, a higher number of OECD countries have reported a clear improvement in life expectancy thanks to better living conditions, the quality of the public health system and the evolution of medicine. In Italy, life expectancy at birth is 83 years, which is higher than the mean of some OECD countries such as Germany, France and Greece, which record an average of 81, 83 and 82 years respectively. The increase in hope is related to a higher cost sustained by the population in terms of health.

Another indicator that makes it possible to measure the subjective well-being perceived by the population of a country is the level of life satisfaction, measured on average. As can be seen from Table 4, the people who believe they are most satisfied live in Germany, France e Spain; while Greece shows the worst performance with a score of 5.4. Italy shows an average income per capita of $26.588 higher than in Spain $23.999, but the level of life satisfaction is higher for the latter country, with a value of 6,3 compared to the value 6 of Italy.

### 4.1. Correlation between some BLI aggregates

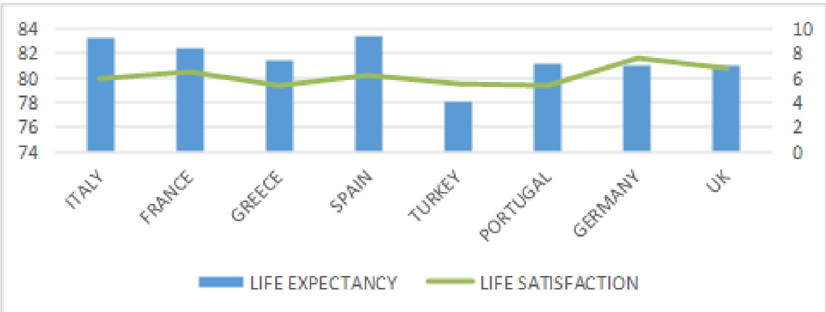
After analysing the data on some aggregates that make up the better life index, it is possible to identify the correlation that exists between some factors, such as income, health and the level of life satisfaction. Life Satisfaction is related to family relationships, health, quality of life, wealth and heritage level, but also confidence in governance. If we consider the data concerning the level of life satisfaction of various countries, all of them covering the same period, it can be affirmed that countries with high average national incomes, consequently, have a higher life satisfaction score.

**Figure 2 –** *Level of Life satisfaction in relationship to the GDP per capita.*




*Source: own elaboration on Word Bank data 2019.*

**Figure 3 –** *Trend of life satisfaction and life expectancy level.*




*Source: own elaboration on Word Bank data 2019.*

Figure 2 shows the data of various countries (France, Germany, Greece, Italy, United Kingdom, Turkey, Portugal and Spain) with different economic and social conditions expressed from the GDP per capita. In the axes of the ordinate, the levels of satisfaction are reported, on a scale from 0 to 10 for each individual country examined, while in the axes of the abscised are reported the GDP per capita. As can be seen from the graph, income and satisfaction are a combination that vary in the same direction. In fact, Germany, UK e France which have a GDP per capita of 45.229,25$, 39.753,24$ and 38.605,67$ respectively, they show a very high level of life satisfaction equal to 7.7; 7.1 and 6.64. This level is lower than in Spain and Greece which in 2018 had a GDP of 34.272,36 and 27.936,9. While Italy with a GDP of 35220,08 has a level of satisfaction of 6.2. Similar data show Spain Greece and Portugal where the Life satisfaction index and GDP per capita is lower compare to the other European country.

Another factor that has a considerable influence on the level of life satisfaction is the state of state of health of citizens. Figure 3 shows a positive correlation between life expectancy and satisfaction levels. In fact, the people who live longer affirm to be satisfied. In other words, the level of satisfaction is on the increase in countries with the lowest mortality rates. The analysis of the data shows that the level of well-being varies according to both socio-economic and demographic differences

## 5. Conclusions

In the present work we tried to evaluate well-being and sustainable development in Italy and in the main partner countries of the European Union, such as France, Germany, Spain and Greece, using both complementary and integrative indices to the GDP. In this perspective, the HDI, the IHDI and the Better Life Index have been analysed; all these indices allow to amplify the traditional economic analysis including both environmental and social factors, allowing a more efficient use of resources, but also to come up with new policies aimed at improving the well-being of the population. The complementary indices, expanding the traditional information base underlying the measurement of GDP, make it possible to increase the qualitative and quantitative aspect of the information, providing disaggregated data useful for making decisions and carrying out the related general planning activities on a country's environmental and social policies.

In this paper, the use of different indices has been emphasized compared to traditional development indicators, which merely take into account the aggregation of simple economic indicators. In this perspective, the IHDI and the Better Life Index provide a wide perspective to evaluate the policy options considering other factors through which it is possible to arrive at a more complete framework of the well-being of a population, compared to what is not possible to do using GDP as the only indicator of surveying. Their application is useful in highlighting the social benefits of policies such as increasing investment in education, increasing the minimum wage.

Starting a process of measuring well-being and using indicators that allow to highlight the variations related to the multiplicity of aspects that affect the concept of well-being, is relevant for the public administration sector, for economic operators and policy makers, as it makes possible the knowledge of information previously not determinable.

## References

ACKERMAN F., STANTON E.A., 2008. The Social Cost of Carbon, real-world economics review, Vol. 26, n. 53, pp. 129-143.

BACHE, I., REARDON, L., 2016. The Politics and Policy of Wellbeing: Understanding the Rise and Significance of a New Agenda, Sheffield, Edward Elgar.

BLEYS, B., WHITBY, A., 2015. Barriers and opportunities for alternative measures of economic welfare, Ecological Economy, Vol. 107, pp.162-172.

BOYCE D. G., LEWIS, M.R., Worm B., 2010. Global phytoplankton decline over the past century, Nature, Vol. 466, n. 7306, pp. 591–596.

COSTANZA, R., DALY, L., FIORAMONTI, L. GIOVANNINI, E., KUBISZEWSKI, I., 2017. Modelling and measuring sustainable wellbeing in connection with the UN Sustainable Development Goals, Ecological Economics, vol. 130, pp. 350-355.

DALY, L., MCELWEE, S., 2014. Forget the GDP. Some States Have Found a Better Way to Measure Our Progress, The New Republic, 4 February 2014, https://newrepublic.com/article/116461/gpi-better-gdp-measuring-united-states-progress.

DIENER, E., OISHI, S., 2000. Money and happiness: Income and subjective well-being across. nations, Culture and subjective well-being across nations, Cambridge, MA, The MIT Press, pp. 185-218.

DOYLE, A., 2018. Warming set to breach Paris accord's toughest limit by mid-century: draft, Reuters, 11 Gennaio 2018.

GUPTA, J., BAUD, I., BEKKERS, R., 2015. Sustainable development goals and inclusive development, Encyclopedia of global environmental politics and governance, Edward Elgar Publishing, pp. 61-72.

HAYDEN, A., WILSON, J., 2017. “Beyond GDP” Indicators: Changing the Economic Narrative for a Post-Consumerist Society?, Social Change and the Coming of Post-Consumer Society: Theoretical Advances and Policy Implications, New York, Routledge, pp. 170-191.

KNOX, J., 2017. Report of the Special Rapporteur on the issue of human rights obligations relating to the enjoyment of a safe, clean, healthy and sustainable environment, UN Special Rapporteur, Human Rights Council - Biodiversity Report.

LAWN, P. 2016. The Genuine Progress Indicator: An indicator to guide the transition to a steady state economy. In: WASHINGTON H. and TWOMEY P (Eds) *A Future Beyond Growth*. Routledge, pp. 182-199.

MAIBOM, J., ANDERSEN, T. M., 2016. The big trade-off between efficiency and equity - is it there?, CEPR Discussion Paper.

MASUR, J., BRONSTEEN, J., BUCCAFUSCO, C., 2010. Welfare as Happiness, Georgetown Law Journal, Vol. 98, pp. 1583-1641.

OKAFOR, E., 2008. Inspirations, challenges and possibilities, International Community Law Review, Vol. 10, n. 4, pp. 371-378.

OISHI, S., KESEBIR, S., DIENER, E., 2011. Income Inequality and Happiness, Psychological Science, Vol. 22, n. 9, pp. 1095-1100.

OSTRY, J.D., BERG, A., TSANGARIDES, C.G., 2018. Redistribution, inequality, and growth: new evidence, Journal of Economic Growth, Vol. 23, n. 3, pp. 259-305.

QU, L., 2015. Life satisfaction across life course transitions, Australian Family Trends n. 8.

SCHULZE, AMROMING, G., DE NARDI, M., 2018. Inequality and Recessions, Chicago Fed Letter, Federal Reserve Bank of Chicago, n. 392.

SPANGENBER, J. H., LOREK, S., 2014. Sustainable consumption within a sustainable economy – beyond green growth and green economies, Journal of Cleaner Production, Vol. 63, pp. 33- 44.

STUTZER, A., FREY, B. S., 2007. What Happiness Research Can Tell Us About Self-Control Problems and Utility Misprediction, CESifo seminar series. Economics and psychology: A promising new cross-disciplinary field, MIT Press, pp. 169–95.

VERHOFSTADT, G., SCHOKKAERT, E., VAN OOTEGEM, L., 2011. Preferences and Subjective Satisfaction: Measuring Well-being on the Job for Policy Evaluation, CESifo Economic Studies, Vol. 57, n. 4, pp. 683–714.

WOLFERS, J., LEIGH, A., 2006. Happiness and the Human Development Index: Australia is Not a Paradox, The Australian Economic Review, Vol..39, n. 32, pp. 176-184.

## SUMMARY

### Social Indicators to measure the well-being of the population. Benchmarking Countries

The issue of the inadequacy of GDP as an indicator of sustainable economic well-being has been the subject of political and economic debate for years. The most relevant literature on the subject focuses on the correlation between GDP and the main elements of human development.

Amromin (2018), Ostry et al. (2014) argue that a high degree of equity facilitates GDP growth in the long term, on the other hand, Maibom and Andersen (2016) affirm that uncontrolled and rapid GDP growth causes inequalities and inequities.

In this work, a selection of summary measures used in literature was considered to compare: the Human Development Index (HDI), which aims to go beyond the concept of economic growth and GDP, to allow the social field to be integrated with the economic dimension by evaluating three factors: longevity, level of education and quality of life standard; the Better Life Index, made up of 11 indicators, gives importance to the concept of "sustainability of well-being"; the Genuine Progress Indicator (GPI) measures the increase in quality of life by taking into account the production of goods and services that do not originate a market transaction, such as volunteering, domestic work and thus balancing the value of private consumption by considering inequalities in the distribution of income. GDP, in fact, indicating the growth of a country in quantitative terms, is not very consistent with the concept of "sustainable economic well-being" that has qualitative nature, so we need a set of indicators that take into account social and environmental factors. In this context, the objective of this work has been to explain how the logic of strict measurement of GDP has been abandoned through the determination of aggregated welfare indices.



Vincenzo MARINELLO, Università degli Studi di Enna "Kore", vincenzo.marinello@unikore.it

Guglielmo L.M. DINICOLÒ, Università degli Studi di Enna "Kore", guglielmo.dinicolo@unikore.it

Chiara DI PUMA, Università degli Studi di Enna "Kore", chiara.dipuma@unikorestudent.it

# EUROPE 2020 STRATEGY FOR A SMART, INCLUSIVE AND SUSTAINABLE GROWTH: A FIRST EVALUATION

Elena Grimaccia

## 1. Introduction

Here we are: it is 2020 and the target planned by the Europe 2020 strategy should have been reached. Of course, many and unforeseen events have occurred, the latest one – and perhaps the worst and most unpredictable – is the Covid-19 pandemic, whose effects we will be able to fully evaluate only in due time. However, it is now worth analysing the level of development reached by the European Union just before the pandemic, also in order to state how Italy performed in comparison with other European Countries.

In this paper, an analysis of the Europe 2020 strategy indicators has been carried out, in order to identify in what socio-economic conditions European Countries have reached 2019.

The "Europe 2020 Strategy", proposed by the European Commission, has been adopted by the European Council in 2010 (European Commission, 2010). The ten years strategy, defined three priorities for growth in the European countries: Smart growth (developing an economy based on knowledge and innovation), Inclusive growth (fostering high employment, and ensuring social and territorial cohesion) and Sustainable growth (promoting a more resources-efficient, greener and more competitive economy).

To reach these objectives, a number of benchmarks have been set, and the indicators identified to measure these goals have been and still are subject to regular statistical monitoring and reporting (Eurostat, 2019). In this way, the policy objectives (smart, sustainable and inclusive growth) were declined in measurable, and thus well-defined, numerical targets, making governments "accountable" to the citizens and to the Commission (Rondinella and Grimaccia, 2017). The five targets are currently measured by eight headline indicators, concerning employment, research and innovation, climate change, renewables and energy, education and poverty.

Europe2020 is perhaps not a complete set of indicators for measuring the progress of societies and the quality of life of their citizens, but it is a very important recognition of European institutions that GDP alone is not enough and that it must necessarily be integrated with measures that take into account equity and sustainability (Grimaccia and Rondinella, 2018). The Europe2020 Indicators have been included in the "GDP and beyond" initiative and present the characteristics of simplicity and reliability needed to monitor the crisis.

In this study, an analysis of Europe 2020 indicators is carried out, in order to show the trends of the eight indicators from the beginning of the Strategy to the most recent data, underlying the convergence (or divergence) processes among European countries, and Italy more specifically. The particular condition of Italy, that appeared in 2019 well behind the other European countries in the development process, has been analysed measuring the distances between national and European targets, and the Europe 2020 indicators values at the beginning of process and in 2019.

## 2. Data

Europe 2020 is the EU's ten-year growth strategy that puts forward three mutually reinforcing priorities (European Commission, 2010):

– Smart growth: developing an economy based on knowledge and innovation;

– Sustainable growth: promoting a more resource efficient, greener and more competitive economy;

– Inclusive growth: fostering a high-employment economy delivering social and territorial cohesion.

These three mutually reinforcing priorities should help the EU and the Member States deliver high levels of employment, productivity and social cohesion (Steurer and Hametner, 2013).

This strategy should provide a road map for economic recovery and guarantee the EU a strong position in international relations. It is essential that five headline targets are met and have been agreed for the whole EU (European commission 2010; Ruser and Anheier, 2014).

The policy objectives (smart, sustainable, and inclusive growth) were declined in well-defined, measurable, numerical targets. Inclusive growth is monitored by the Employment rate (it should be 75% of the population aged 20–64 in 2020), and by the number of People at risk of poverty or social exclusion (there should have been 20 million people at risk of poverty less). Smart growth is measured by the share of Early leavers from education and training (that should decrease to under 10%), by the share of Tertiary educational attainment (that should reach 40% of the population aged from 30 to 34 years old), and by the Gross domestic expenditure on research and development (R&D) that should increase to 3% of the EU's GDP. Sustainable growth: the ''20/20/20'' climate/energy targets should be met reducing by 20%

Greenhouse gas emissions (30% of emissions reduction if the conditions are right), raising the Share of renewable energy by 20%, and enhancing Energy efficiency, reducing energy consumption by 20% (Table 1).

**Table 1 –** *The Europe 2020 strategy's key priorities and headline targets.*

| **Priorities** | **Headline targets** |
|---|---|
| *Smart growth* | Increasing in R&D to 3 % of GDP<br>Reducing school drop-out rates to less than 10 %<br>Increasing the share of the population aged 30–34 having completed tertiary education to at least 40 % |
| *Inclusive growth* | Increasing the employment rate of the population aged 20–64 to at least 75 %<br>Lifting at least 20 million people out of the risk of poverty and social exclusion |
| *Sustainable growth* | Reducing greenhouse gas emissions by at least 20 % compared to 1990 levels<br>Increasing the share of renewable energy in final energy consumption to 20 %<br>Moving towards a 20% increase in energy efficiency |

*Source: Eurostat 2019.*

In order to raise effectiveness in implementing the strategy, each Member State (MS) should develop their own policies, such as: action plans and defining goals and short-term, mid-term and long-term actions; preparing qualitative and quantitative measures allowing to make comparisons between the MSs and between countries and the whole Europe Union (EU); translate strategic objectives and schedules into national action plans; periodic reviews of the strategy implementation in terms of realisation its objectives and exchanging experiences (Barder et al., 2013; Ruser and Anheier, 2014). Each MS set its own targets, according to their level of development, measured by the eight Europe 2020 indicators. Italy chose very unambitious targets, in soe cases as lower as a half of the EU goals.

The EU will not accomplish its goals if the individual member states do not pursue them. The Union is a community of great economic potential but with many problems to overcome. It is composed of 27 countries with different levels of socio-economic development[1].

[1] In this paper, since the United Kingdom exit from the Union in 2019, the EU is considered to be composed of 27 Member States.

## 3. Analysis of Europe 2020 Indicators

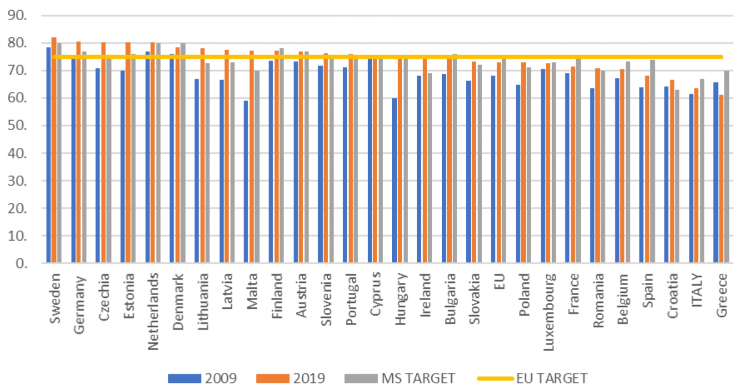
In this Section, an analysis of the Europe 2020 indicators in the ten years considered in the Strategy is presented. Afterwards, a more specific study on Italy's target and achievement is presented.

### *3.1. Europe 2020 Strategy indicators in the European Union*

The employment target has a pivotal role in the strategy, among education-related targets, and the goal of poverty reduction. For the monitoring of the target, the employment rate of people aged between 20 and 64 was chosen as the key indicator, setting a target of 75 percent.

In 2010, the average European employment rate of people aged between 20 and 64 was 6.8 percentage points lower than the target set for 2020. However, this average summarizes strong differences between Member States: some countries, including Sweden, Denmark and Germany, had already reached the target set for 2020, while many others, including Italy and Spain, had differences of more than 10 percentage points with the European target.

**Figure 1 -** *Employment rate by sex, age group 20-64 in the UE –2009, 2019 and 2020 national and EU targets (percentages).*

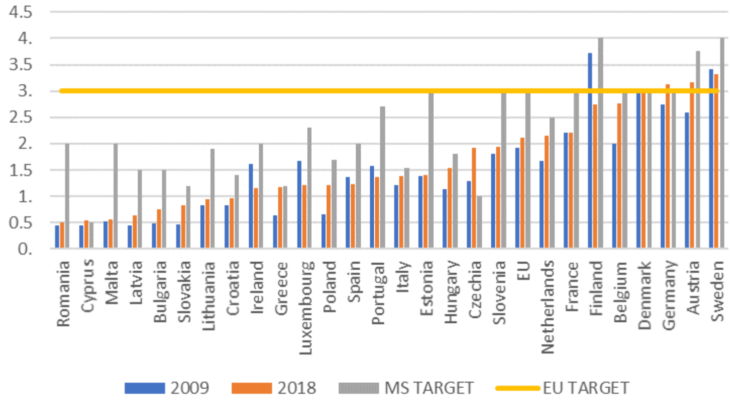



*Source: Author's elaboration on Eurostat data*

The increase in the employment rate in the period considered was not as high as was expected. The target for the EU as a whole has not been achieved, although some countries (such as Germany and Sweden) have exceeded expectations. Italy has not achieved the EU target and not even its national target, that was lower and set at 67%.

Research and Development, and innovation are key policy components of the Europe 2020 strategy and they contribute to a well-functioning knowledge-based economy and industrial competitiveness (Eurostat, 2019). R&D intensity in the EU is growing too slowly to meet the Europe 2020 target. Moreover, the national target set by Italy for 2020 is clearly less ambitious and is only half of the European one (1.53%). The national target – until 2018 – has not been reached, nor the EU one.

**Figure 2 -** *Gross domestic expenditure on R&D. –2009, 2019 and 2020 MS and EU Europe 2020 targets (percentages).*

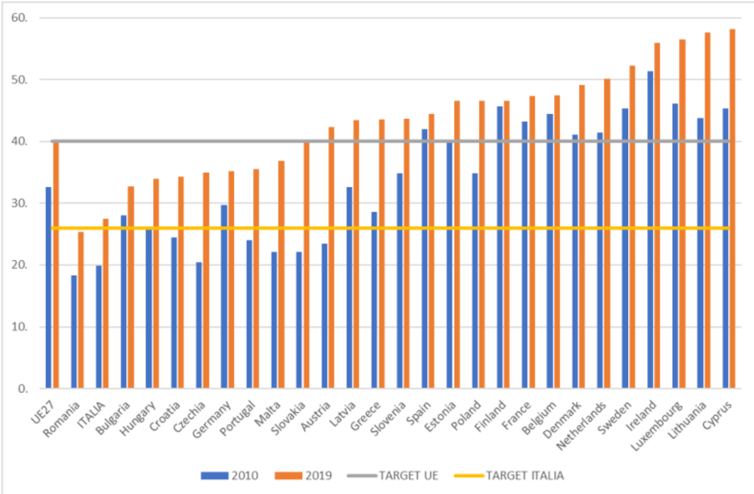



*Source: Author's elaboration on Eurostat data*

Closely linked to the objectives relating to the labour market, and to R&D and innovation, it is that of university education. The target set for the share of graduates between 30 and 34 years for the Union is 40 percent.

In 2009, almost half of the countries of the Union had already reached the target set: some northern European countries (Sweden, the Netherlands, Ireland, Denmark, and Finland) but also France, Belgium, and Spain, had rates close to or above 40 percent. Thanks to the positive dynamics observed between 2009 and 2019, many other European countries have reached this share. Among these, the Baltic republics and Poland, with progress even over 10 percentage points (Figure 3). Italy, on the other hand, had a fairly low indicator value (19.0 percent) in 2009, ranking fourth last in the ranking of the European Union, and lost further ground, limiting itself to reach 27.5 percent of graduates between 30 and 34 years of age and finishing at the bottom of the ranking.

**Figure 3 -** *Tertiary educational attainment by sex, age group 30-34 in EU countries (Isced 5-8) –2009, 2019 and 2020 MS and EU Europe 2020 targets (percentages).*

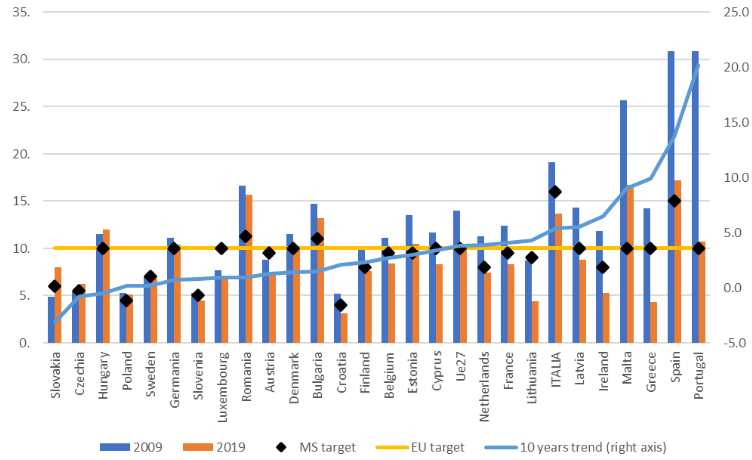



*Source: Author's elaboration on Eurostat data*

For this indicator, the efforts made by the countries are highly diversified. Some countries like Slovakia set very high targets and reached them. Some countries such as Austria, Czechia and Lithuania have exceeded their very ambitious targets. Italy, instead, has established an unambitious target, and has reached and surpassed it.

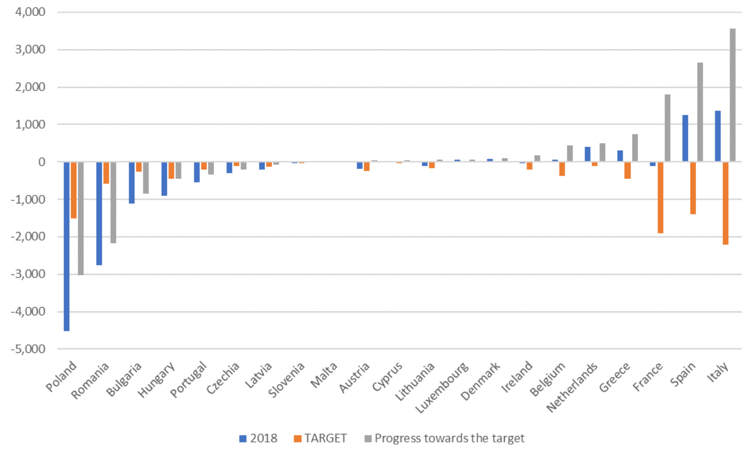
Reducing the dropout rate to less than 10 percent by the end of the decade is the second goal for education set by the Europe 2020 Strategy. The term "early school leaving" means all forms of drop out from school or training before completion of upper secondary education or its equivalents in vocational training. The indicator chosen for monitoring is the share of the population belonging to the age group between 18 and 24 years old who dropped out of studies without having obtained a qualification higher than the upper secondary education diploma (level 3C short of the classification international education level Isced). In 2009, Italy was among the countries of the Union with the highest school dropout rate (19 percent of young people, compared to a European average of 14 percent). In this context, Italy is among the countries that has made the greatest progress, decreasing the drop-out rate by 5 points and exceeding the set national target (the least ambitious among all 27 countries, equal to 16 percent). However, the school dropout rate for young people between 18 and 24 is still higher than the European average in Italy (13% compared to 10).

**Figure 4 -** *Early leavers from education and training in EU countries, 2009, 2019, MS and EU 2020 targets (percentages).*




*Source: Author's elaboration on Eurostat data*

**Figure 5 -** *People at risk of poverty or social exclusion in EU countries, 2018 differences from 2008, MS 2020 targets and distance from the target.*




*Source: Author's elaboration on Eurostat data*

Poverty deserves a separate analysis: in Italy, instead of the foreseen reduction of the number of poor people by 2.2 million, an increase of more than one million poor has occurred (Figure 5).

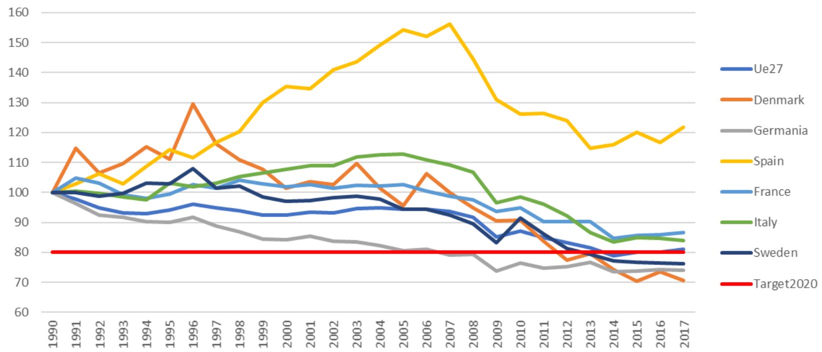
Italy, together with Spain, France and Greece, is the only EU MS presenting this trend, and it is the worst performing.

With regard to environmental aspects of the strategy, Italy is in line with European levels and targets. The section of the Europe 2020 Strategy dedicated to sustainable growth from an environmental point of view identifies specific targets to be reached by the end of the decade: a 20% reduction in greenhouse gas emissions compared to 1990 levels; the 20% increase in the share of final consumption from renewable energy; a 20 percent increase in energy efficiency.

In the EU countries, the level of greenhouse gas emissions in 2009 was 17.6% lower than in 1990, therefore not far from the 20% target set by Europe 2020. This reduction, however, was largely due to the significant contraction that has occurred since 2008 (Figure 6), that is, in correspondence with the economic crisis.

In fact, despite the slowdown in the economy, only some of the main European countries (Denmark, Sweden and Germany) achieved the emissions reduction target, while many others remained above 80% of 1990 emissions, including Italy.

**Figure 6 –** *Greenhouse gas emissions – Anni 1990- 2017 (numeri indici 1990=100).*



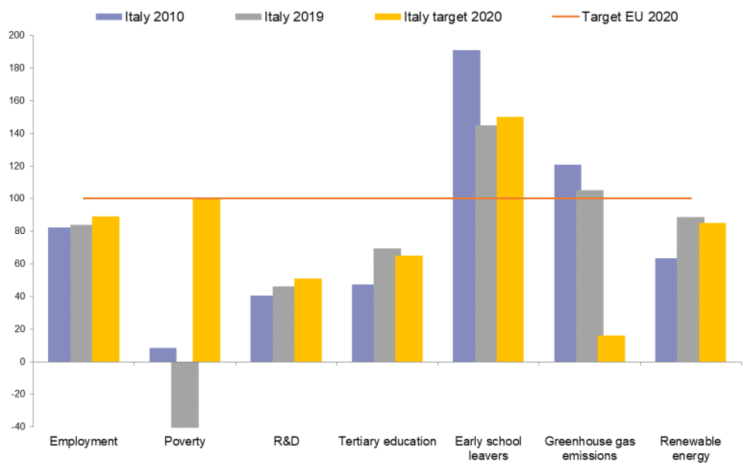

*Source: Author's elaboration on Eurostat data*

## *3.2. Europe 2020 target and indicators for Italy*

A dynamic analysis of the distance from the European targets shows, in summary, that Italy has achieved its objectives for tertiary education and school dropouts, shows trends in terms of the employment rate and expenditure on research and development that are positive but not sufficient to reach their targets, with positive trends for the two environmental indicators but with particularly high deviations from the target value compared to the risk of poverty (Figure 7).

In the past ten years, the Italian objectives related to education have been achieved and exceeded, but the values of the indicators are still far from the level of the targets of the European Union as a whole. For school dropouts, a target of 15% was foreseen, exceeded in 2018 with 14.5% of dropouts, marking a significant progress compared to 19% in 2010, but still with a level 50% higher than the EU27 targets, established at 10%. Same trend for university degrees: the target of 26% of young graduates was already achieved in 2018 (with a share of 27.9% of young people between 30 and 34 years of age), but it is still a much lower level to that established - and achieved - by the European Union in 2019.

**Figure 7 -** *Distances from Europe 2020 Targets- ITALY- 2009 and 2019*[2] *(percentages).*




*Source: Author's elaboration on Eurostat data*

Different speech must be made - instead - for the employment rate. Compared to this important indicator, Italy has not even reached the national target, further increasing the distance from the European average.

The trend of the indicator relating to the risk of poverty is even worse. The number of people at risk of poverty and social exclusion, which had been established to decrease by 2.2 million, instead increased by 1.3 million. An opposite trend, therefore, to that expected in 2010.

Overall, it is important to highlight how the target levels for Italy were all less ambitious than those set by the Union (Figure 1). For employment, for investment in

[2] Data refer to 2017 for poverty and R&D, and to 2018 for emissions and energy.

research and development, for education Italy had set targets between 10 and 50% lower than in the EU27. The environmental objectives were indicated at 20% for all, while the reduction of poverty was established individually for each country. When the objectives were established, the main problems (i.e. the greatest distances between the Italian levels in 2009 and the European objectives) referred to the sectors most linked to competitiveness, highlighted by the research and development indicators, higher education and of school dropouts. For the representation of the indicator relating to the "poverty" objective, which in the Europe 2020 strategy envisages a reduction of 20 million people at risk of poverty and exclusion as the EU aggregate value, the quota assigned to Italy was a reduction of 2.2 million people.

## 4. Conclusions and policy remarks

Public policy has an important impact on people’s wellbeing: income, work, environment are all factors that affect the main determinants of quality of life. Therefore, monitoring of objective indicators, such as poverty, employment or greenhouse emissions, policy makers could measure the modifications of the wellbeing of different populations (Grimaccia and Lima, 2015).

This study shows that – according to the Europe 2020 indicators, in the EU social and economic disparities have stopped narrowing, poverty and exclusion have increased, including in the richer Member States. EU27 as a whole shows a progress in sustainability and smartness indicators but it failed in the inclusion indicators (no improvement in employments level and poverty increased). More specifically, Italy set very unambitious targets, well below the other MSs and the EU’s ones. Moreover, the employment rate did not increase enough to reach even the national target (8 percentage points lower than the EU one). The number of poor people presents a growing trend that is opposite to what was expected in 2010. This increase in the number of poor is a trend that is not common in Europe, with Poland and Romania reducing the number of poor of 4.8 and 3 million respectively.

The European Union, in 2010, set clear development goals - ambitious and measurable. This should have allowed policy makers to follow up on achieving these goals. Analysing the situation 10 years later (but with data that do not take into account - for temporal reasons - the shock due to Covid-19 pandemic) it can be seen that Europe has not achieved its objectives. Looking at the detail of single MSs, apart from virtuous countries such as Sweden and Austria, many MSs did little to achieve the goals that were considered important by the countries themselves. Italy - which had set itself much lower targets than the rest of Europe - in many cases did not achieve them.

For the future, UN Sustainable Development Goals certainly identify an important pattern for fair and sustainable development, but the use of more than 200 indicators (some of which are currently not measurable ) certainly makes progress more difficult to monitor, and countries less accountable for their development's progress efforts.

## References

BARDER, O., CLARK, J., LEPISSIER, A., REYNOLDS, L., ROODMAN, A. 2013. Europe beyond aid: assessing European countries' individual and collective commitment to development, *Journal of International Development*, Vol. 25, pp. 832–853

EUROPEAN COMMISSION. 2010. *Europe 2020. A strategy for smart, sustainable and inclusive growth*. COM(2010), Brussels, 2010

EUROSTAT. 2019. *Smarter, greener, more inclusive? Indicators to support the Europe 2020 strategy*. Luxembourg.

GRIMACCIA E., LIMA R. 2015. Quality of Life in Europe and Italy: Regional disparities according to the Europe2020 indicators on inclusion and smartness. In *Giornate di studio sulla popolazione 2015*, POPDAYS, University of Palermo, 4-6 February 2015.

GRIMACCIA E., RONDINELLA T. 2018. A novel perspective in the analysis of sustainability, inclusion and smartness of growth through Europe 2020 indicators In CIRA PERNA, MONICA PRATESI, ANNE RIUZ-GAZEN (Eds.) *Studies in Theoretical and Applied Statistics*. Springer.

RONDINELLA T., GRIMACCIA E. 2017. Joint Analysis of Structural Models and Performance: Merging Clustering and Composite Indicators in the Analysis of Europe 2020 Strategy. In MAGGINO F. (Ed.) *Complexity in Society: From Indicators Construction to their Synthesis*. Social Indicators Research Series, 70.

RUSER, A., ANHEIER, H.K. 2014. The EU's future role on the global stage. *Global Policy*, Vol. 5, No.1, pp. 58–67

STEURER, R., HAMETNER, M. 2013. Objectives and indicators in sustainable development strategies: similarities and variances cross Europe. *Sustainable Development*, Vol. 21, No.4, pp. 224–241.

## SUMMARY

### Europe 2020 Strategy for a Smart, Inclusive and Sustainable Growth: A First Evaluation

In this paper, an analysis of the Europe 2020 strategy indicators has been carried out, in order to identify in what socio-economic conditions European Countries have reached 2019. The strategy defined three priorities for growth in the European countries: Smart growth (developing an economy based on knowledge and innovation), Inclusive growth (fostering high employment, and ensuring social and territorial cohesion) and Sustainable growth (promoting a more resources-efficient, greener and more competitive economy). These goals were declined in eight headline indicators, and thus measurable targets, making governments "accountable" to the citizens and to the Commission. The indicators are related to employment, research and innovation, climate change, renewables and energy, education and poverty. For each indicator, a benchmark has been set, and the eight indicators are subject to regular statistical monitoring and reporting. Europe2020 is perhaps not a complete set of indicators for measuring the progress of societies and the quality of life of their citizens, but it is a very important recognition of European institutions that GDP alone is not enough and that it must necessarily be integrated with measures that take into account equity and sustainability . After 10 years, the EU has not reached the target identified in 2010, and Italy in 2019 was well behind the other countries in the EU.


Elena GRIMACCIA, Istat, elgrimac@istat.it

# PUBLIC SUPPORT FOR AN EU- WIDE SOCIAL BENEFIT SCHEME: EVIDENCE FROM ROUND 8 OF THE EUROPEAN SOCIAL SURVEY (ESS)

Paolo Emilio Cardone

## 1. Introduction

An important aspect of most democratic societies is a welfare state with government funded services that offer financial protection to its citizens, paid for by taxes. This can encompass a whole set of services including healthcare provision, unemployment benefits, housing costs and pensions.

In the past decades, the European Union has gradually assumed a more active role in social policymaking (Falkner, 2016), but the extended European-style welfare state became substantially challenged due to a number of major economic, social and political developments.

Furthermore, longer-term challenges have been exacerbated by the shock of the banking crisis in 2008, which was quickly followed by an economic recession in 2009, and a longer-lasting fiscal and debt crisis in many European states.

In many welfare states, the challenges posed by the nearly universal trends of growing inequality, migration, ageing, globalisation and digitalization of work have been further aggravated by the recent economic crisis. At the same time, these trends put the sustainability of social policies under pressure and thus bring back to the political agenda discussions about policy reforms (Bonoli, 2005).

Many countries have experienced government-imposed austerity measures since the initial shadow of the 2008 economic crisis, and many areas of public expenditure have been stagnant, scaled back or cut completely.

This raises the question whether European respondents support this evolution, or whether they see the development of a Social Europe as a threat to their national welfare arrangements in order to assess whether financial restrictions on the welfare state in many countries have changed public attitudes towards it.

Furthermore, there is an ongoing European Union debate, ignited substantially by the unequal degree to which the economic crisis has hit the different countries in Europe. It regards the solidarity between Europeans, addressing the question of whether a redistribution of welfare from richer to poorer Europeans would be necessary to create cross-European social cohesion, and would be politically and economically feasible.

Thus, the possibility to track, capture and investigate individuals' behaviours, values, beliefs and attitudes over time and across space has become increasingly relevant for the scholarly understanding of a rapidly changing social world.

The ESS Round 8 module "*Welfare Attitudes in a Changing Europe: Solidarities under Pressure*" makes it possible to shed scientific light on these debates.

## 2. Data and methods

The analysis is carried out using microdata from the quantitative research "*European Social Survey*" (ESS Data, 2016)[1].

The ESS is an academically driven cross-national survey that has been conducted in over thirty countries across Europe since its establishment in 2001.

It collects information on people's attitudes, beliefs and behaviour patterns in many European countries. It does so every two years in order to measure stability or change over time. Subjects covered in the ESS questionnaire include participation in society, religious and political beliefs, and (specific to the eighth round) welfare as well as climate change and energy.

In 2013, the European Social Survey (ESS) became a European Research Infrastructure Consortium (ERIC). The ESS ERIC is hosted by the United Kingdom with its headquarters at City University London. Other institutions that are part of the Core Scientific Team behind the ESS are: Leibniz Institute for the Social Sciences (GESIS, Germany), the University of Leuven (KU Leuven, Belgium), the Norwegian Centre for Research Data (NSD - Norwegian Centre for Research Data, Norway), the Netherlands Institute for Social Research (SCP, The Netherlands), the University of Ljubljana (UL, Slovenia) and the Universitat Pompeu Fabra (UPF, Spain).

The ESS source questionnaire contains a "core" module, which largely remains the same each round[2].

In each round, there are also two short "rotating" modules, which are developed by competitively selected, multinational questionnaire design teams in collaboration with the Core Scientific Team (CST).

In Round 8 these modules focus on:

• Public Attitudes to Climate Change, Energy Security and Energy Preferences, (new).

[1] For more details: https://www.europeansocialsurvey.org.

[2] For more details: www.europeansocialsurvey.org/methodology/questionnaire.

• Welfare Attitudes in a Changing Europe: Solidarities under Pressure (repeat module with a number of new items).

The ESS Round 8 module "*Welfare Attitudes in a Changing Europe: Solidarities under Pressure*", fielded in 2016/17, only partly repeats the ESS Round 4 Welfare Attitudes module (fielded in 2008/09).

In particular, the core and rotating modules that form the backbone of the ESS questionnaires have addressed multiple topics, including attitudes toward the media, social trust, politics, democracy and citizen involvement; subjective well-being and human values; attitudes towards immigration; family, work and well-being, the timing of life and gender roles; economic morality, welfare attitudes and justice; public attitudes toward climate change.

More in details, the inclusion of the Welfare Attitudes in Europe module during Round 8 of the ESS, first of all allowed attitudes towards these services to be assessed in 23 countries, but also it addresses new solidarity questions fielded for the first time, most notably items assessing the introduction of a universal basic income (UBI) scheme and the implementation of a European Union-wide social benefit scheme.

Round 8 of the ESS (about 44,000 individuals aged 15 or older) was fielded in 23 countries: Austria, Belgium, Czech Republic, Estonia, Finland, France, Germany, Hungary, Iceland, Ireland, Israel, Italy[3], Lithuania, Norway, the Netherlands, Poland, Portugal, Russia, Slovenia, Spain, Sweden, Switzerland and the United Kingdom.

## 3. Results

The Welfare Attitudes module evaluates whether respondents think *the level of social benefits and services in their country would become higher or lower if more decisions were made by the European Union rather than by national governments* (item E37).

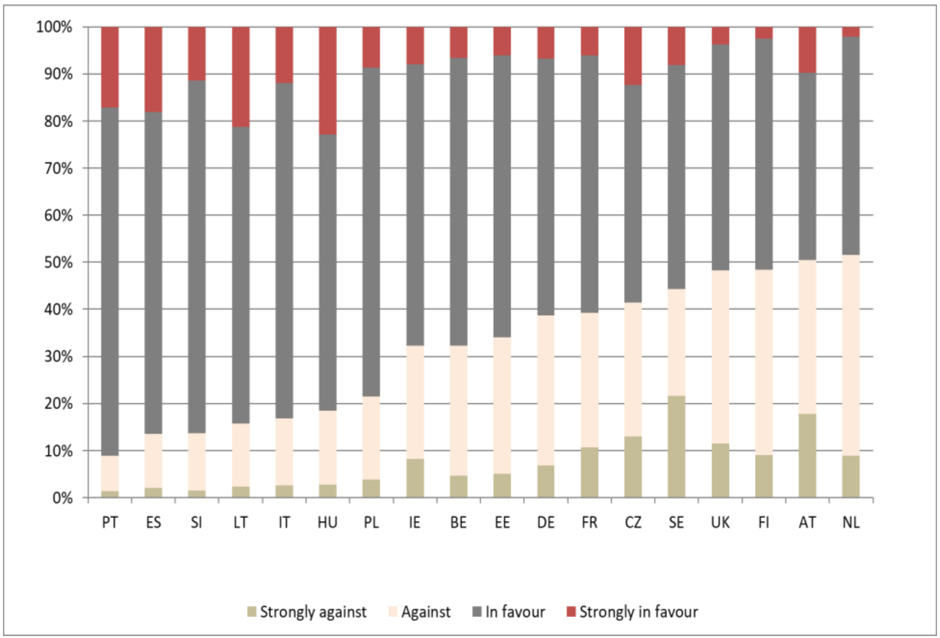
In general, 67,1% of Europeans express their support for an EU-wide social benefit scheme that would guarantee a minimum standard of living for the poor (in favour 58,2% and strongly in favour 8,9%)[4].

---

[3] Italy has participated in the ESS on four occasions: in rounds 1, 2, 6 and in the recent round 8, collected between 2016 and 2017 and released in May 2018. In 2017, thanks to Inapp to carry out the survey, Italy returns to the ESS ERIC with the status of "full member". INAPP is National Institute for Public Policy Analysis, former ISFOL (National Research Institute for Vocational Education and Training Employment), that changed its company name in INAPP (Istituto Nazionale per l'Analisi delle Politiche Pubbliche – Public Policy Innovation) on December 1st 2016 (www.inapp.org).

[4] Please note that five extra-EU countries (Iceland, Israel, Norway, Russian Federation and Switzerland) are obviously excluded from this item.

**Figure 1** – *Public support for an EU-wide social benefit scheme (% values).*




*Source: own elaboration on ESS data Round 8.*

Legend: PT=Portugal; ES=Spain; SI=Slovenia; LT=Lithuania; IT=Italy; HU=Hungary; PL=Poland; IE=Ireland; BE=Belgium; EE=Estonia; DE=Germany; FR=France; CZ=Czech Republic; SE=Sweden; UK=United Kingdom; FI=Finland; AT=Austria; NL=Netherlands.

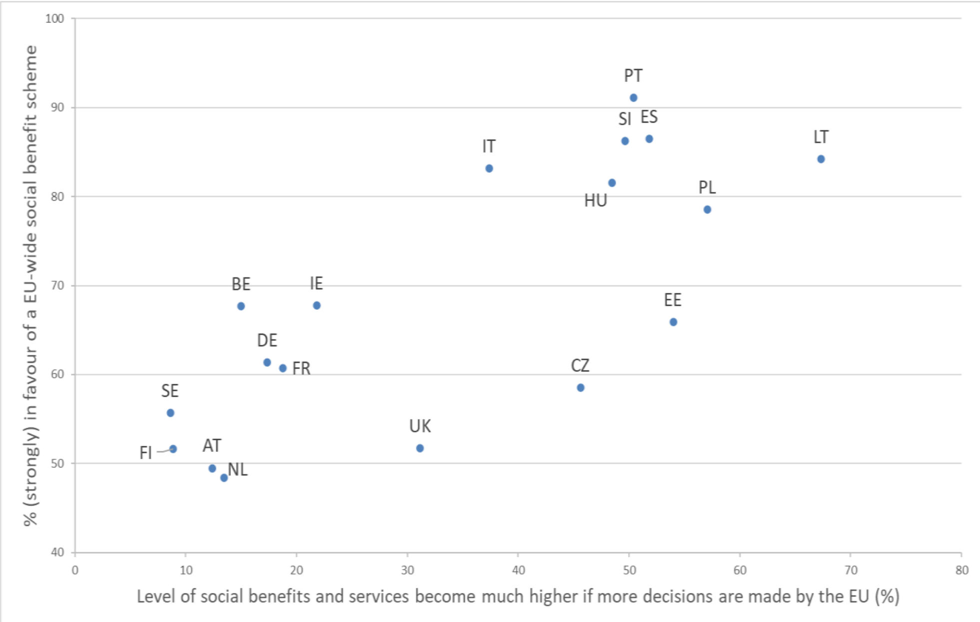
As shown in figure 1, Portugal is the most European country clearly in favour of a public support for an EU-wide social benefit scheme (individuals "in favour" and "strongly in favour" are more than 90%) followed by Spain, Slovenia, Lithuania, Italy and Hungary with a percentage between 80 and 90%. On the other side, the Netherlands are the most opposed country with less than 50% in favour (Cardone *et al.*, 2019). This is why we have chosen this country as reference category in the logistic model shown below.

On average, three in ten Europeans (30,5%: people who is higher 27,2% and much higher 3,3%) believe that increased EU involvement would lead to higher or much higher levels of social protection (item E38). By contrast, 69,5% expect benefit levels to stay the same or become lower as a result of more European decision-making.

Despite these relatively widespread concerns about Social Europe, as explained previously, 67,1% of Europeans express their support for an *EU-wide social benefit scheme that would guarantee a minimum standard of living for the poor* (item E37).

Both attitudes are neatly aligned: in countries with strong expectations that Europeanisation will increase benefit levels, public support for an EU-level benefit scheme is comparatively strong as well (Figure 2).

**Figure 2** – *Public support for an EU-wide social benefit scheme and expectations that Europeanisation will increase benefit levels (% values).*



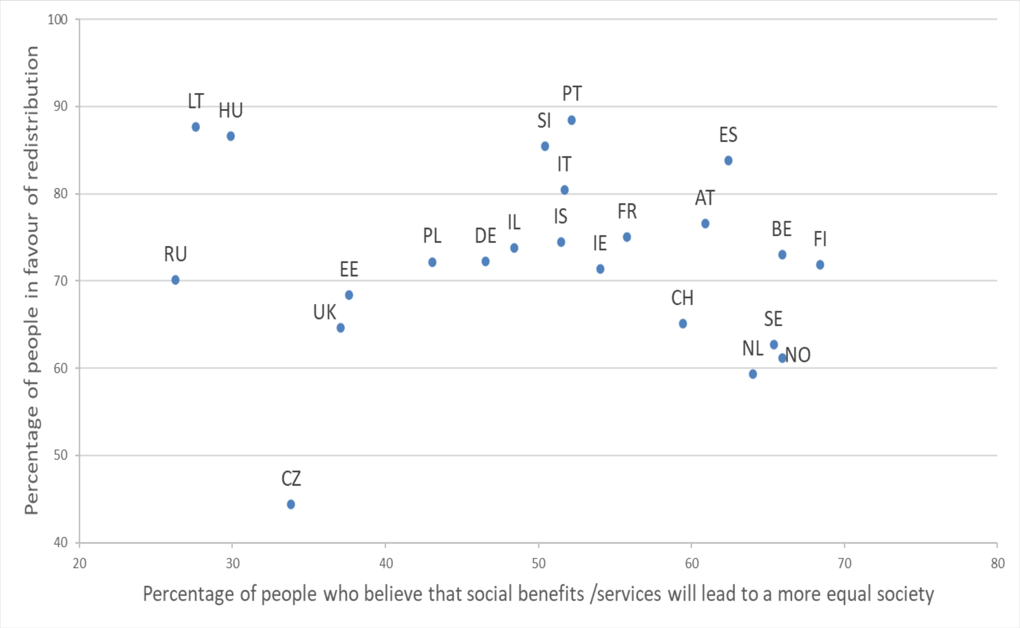

*Source: own elaboration on ESS data Round 8.*
*Note: Results are weighted for age, gender and education (pspweight).*
N (item E38) = 31.764. "*More decisions made by EU: level of benefits in [country] become higher or lower*". Question was answered on a 5-point scale: 'much higher', 'higher', 'neither higher nor lower', 'lower' or 'much lower'. The numbers refer to the percentage of people who is much higher or higher with the respective statement.
N (item E37) = 32.587. "*Against or in favour of European Union-wide social benefit scheme*". Question was answered on a 4-point scale: 'strongly in favour', 'in favour', 'against', 'strongly against'. The numbers refer to the percentage of people who is strongly in favour or in favour with the respective statement.

The generosity of national welfare systems is a crucial driver of the sizeable crossnational differences in attitudes towards Social Europe.

In the strongly developed Nordic welfare states, few respondents expect improvement from Europeanisation of social policy, and support for EU-level benefits is relatively low. In the Eastern and Southern European countries, where social expenditure is considerably lower, respondents more often see the EU as an agent that could improve social protection.

Solidarity with poor people, measured as the agreement with the statement "*the government should take measures to reduce income inequality"*, receives strong support all over Europe.

**Figure 3 –** *Average support for redistribution vs. the percentage of people who believe that social benefits/services lead to a more equal society (% values).*




*Source: own elaboration on ESS data Round 8.*
*Note: Results are weighted for age, gender and education (pspweight).*

N (item E11) = 42.894. The belief in the effectiveness of the social benefits system is measured by the agreement with the statement: "*Social benefits and services in [country] lead to a more equal society*".

N (item B33) = 43.715. The preference for redistribution is measured by the agreement with the statement: "*The government should take measures to reduce differences in income levels*".

Both questions were answered on a 5-point scale: 'agree strongly', 'agree', 'neither agree nor disagree', 'disagree' or 'disagree strongly'. The numbers refer to the percentage of people who agree or strongly agree with the respective statement.

In all countries except for the Czech Republic, more than 60 per cent of the population are in favour of redistribution (Figure 3).

If we look at how people perceive the effectiveness of the social benefits system in their country (i.e., whether they believe that social benefits will lead to more equality) we again find regional patterns, but not a strong overall relationship. People support the idea that the government is responsible for reducing income inequality (item B33) independently of whether they think the government's social services will be successful (item E11). People in the Nordic countries are quite confident that social benefits will lead to more equality, whereas people in the Eastern European countries are at the other end of the scale.

In other words, Russians and Eastern Europeans are less confident that social benefits lead to more equality. In particular Russia stands out as a distinctive case, with only a quarter of the population (26,3%) believing in the equalising effects of their welfare state, which may also contribute to Russians not being as favourable of the idea of income redistribution as one would expect (only 70,1%), given the high level of inequality (the highest Gini Index) in this country (Ochsner *et al*., 2018).

Using multivariate analysis (logistic regression models with Stata software) it was possible to estimate the different attitudes among countries for an EU-wide social benefit scheme more accurately. The model has been developed for EU citizens only and includes, first of all adults' socio-demographic characteristics (age, gender, number people living in the household, citizenship, domicile, education level, voted or not), secondly, economic and work-related (worked or not, total household income).

In order to achieve this goal, the dependent variable of this study is the "*social benefit scheme*" (equal to 1 if the individual is in favour, otherwise against).

Concretely, in the study analyzed variables are:

- *Gender*. Categorical. Dummy variable: Female, Male (reference cat.).
- *Country*. Categorical. Eighteen countries. Netherlands (reference cat.), Portugal, Spain, Slovenia, Lithuania, Italy, Hungary, Poland, Ireland, Belgium, Estonia, Germany, France, Czech Republic, Sweden, United Kingdom, Finland, Austria.
- *Domicile*. Categorical. Four levels. A big city/Suburbs or outskirts of big city (reference cat.); Town or small city; Country village; Farm or home in countryside.
- *Work*. Categorical. Dummy variable: Yes, No (reference cat.).
- *Income*. Categorical. Ten levels: 1st decile (reference cat.), 2nd decile, 3rd decile, 4th decile, 5th decile, 6th decile, 7th decile, 8th decile, 9th decile, 10th decile.
- *Household*. Categorical. Five levels: 1 individual (reference cat.), 2 ind., 3 ind., 4 ind., 5 ind. or more.
- *Vote*. Voted in the last election. Categorical. Dummy variable: No, Yes (reference cat.).
- *ISCED*. Categorical. Three levels: Low (Isced 0-1-2), Medium (Isced 3-4), High (Isced 5-6, reference cat.).
- *Age group*. Categorical. Three intervals. From 15 to 40; 40 to 60; over 60 (reference cat.).

First of all, we test the goodness-of-fit using a postestimation tool, the Hosmer-Lemeshow statistic. This test follows a chi-square distribution with the degrees of freedom equal to the number of groups minus 2. A not significant *p* value indicates that the model fits the data well since there is no significant difference between the observed and expected data (Liu, 2016). In this case, the Hosmer-Lemeshow chi-square test has a value of 8,06 with the degrees of freedom equal to 8. The associated *p* value is 0,4272 which is not significant. Therefore, the model fits the data well.

Logistic model for "*social benefit scheme*", goodness-of-fit test:
Number of observations = 32.042
Number of groups = 10
Hosmer-Lemeshow chi2(8) = 8,06
Prob > chi2 = 0,4272

Table 1 shows odds ratios of logistic model. The coefficients (*Beta*, not showed) can be expressed in odds by getting rid of the natural log. This is done by taking the exponential for both sides of the equation, because there is a direct relationship between the coefficients produced by logit and the odds ratios produced by logistic: a logit is defined as the natural log (base *e*) of the odds.

This fitted model says that, holding covariates at a fixed value, the odds of being in favour of a public support for an EU-wide social benefit scheme for female over the odds of being in favour of a public support for an EU-wide social benefit scheme for male (reference category) is 1,10. In terms of percent change, we can say that the odds for female are 10% higher than the odds for male. In other words, the hazard to be in favour of a public support for an EU-wide social benefit scheme is higher for female rather than male.

Regarding the citizenship, the odds of being in favour of a public support for an EU-wide social benefit scheme for all countries over the odds of being in favour of a public support for an EU-wide social benefit scheme for The Netherlands (reference category) is getting higher except for Austria (OR=0,89). Please note that the odds for Finland and United Kingdom are not significant (*p* value > 0,05).

The choice of the reference category is always for the extreme one since moving away from these categories, the risk increases (country, age, Isced, vote and household ) or decreases (income and domicile).

**Table 1 -** *Logistic regression model.*

| Variables | | ODDS | Sign. |
|---|---|---|---|
| • Gender (Male=base) | Female | 1,10 | 0,000 |
| • Country (NL=base) | AT | 0,89 | 0,050 |
| | BE | 2,03 | 0,000 |
| | CZ | 1,47 | 0,000 |
| | DE | 1,60 | 0,000 |
| | ES | 6,23 | 0,000 |
| | FI | 1,06 | 0,429 |
| | FR | 1,44 | 0,000 |
| | UK | 0,97 | 0,711 |
| | HU | 4,39 | 0,000 |
| | IE | 2,11 | 0,000 |
| | IT | 4,92 | 0,000 |
| | LT | 5,05 | 0,000 |
| | PL | 3,63 | 0,000 |
| | PT | 10,97 | 0,000 |
| | SE | 1,24 | 0,003 |
| | SI | 6,23 | 0,000 |
| • Domicile | Town or small city | 0,90 | 0,001 |
| (A big city/Suburbs=base) | Country village | 0,89 | 0,000 |
| | Farm or home in countryside | 0,81 | 0,000 |
| • Work (No=base) | Yes | 0,87 | 0,000 |
| • Income (J - 1st decile =base) | R - 2nd decile | 0,93 | 0,252 |
| | C - 3rd decile | 0,90 | 0,094 |
| | M - 4th decile | 0,88 | 0,041 |
| | F - 5th decile | 0,78 | 0,000 |
| | S - 6th decile | 0,83 | 0,006 |
| | K - 7th decile | 0,83 | 0,007 |
| | P - 8th decile | 0,70 | 0,000 |
| | D - 9th decile | 0,72 | 0,000 |
| | H - 10th decile | 0,70 | 0,000 |
| • Household | 2 | 1,12 | 0,002 |
| (Single person / lone parent=base) | 3 | 1,11 | 0,016 |
| | 4 | 1,08 | 0,088 |
| | 5 or more | 1,12 | 0,056 |
| • Vote (Yes=base) | No | 1,09 | 0,014 |
| | Not eligible to vote | 1,29 | 0,000 |
| • ISCED (High=base) | Low | 1,17 | 0,000 |
| | Medium | 1,04 | 0,176 |
| • Age group (Over 60=base) | 15 – 40 | 1,14 | 0,001 |
| | 40 – 60 | 1,15 | 0,000 |
| | cons. | 1,06 | 0,443 |

*Note: Number of obs = 32.042; LR chi2(44) = 3251,77; Prob > chi2 = 0,0000; Log likelihood = -18752,16; Pseudo R2 = 0,0798*
*Source: own elaboration on ESS data Round 8.*

In particular, the hazard to be in favour of a public support for an EU-wide social benefit scheme is higher for young people (younger ones have more confidence than the elderly, "over 60" reference cat.) and for those who have a low education level (Isced 5-6 reference cat.). Moreover, it decreases with household income (1st decile reference cat.), for those who live in a small town/village or farm (big city reference cat.) and for workers (not workers reference cat.). On the contrary, the hazard to be in favour increases for those who do not vote or are not eligible (those who vote reference cat.) and for individuals who belong to families of 2 or more people (single persons/lone parents reference cat.).

## 4. Conclusions

The literature on the Europeanization of social protection has been mainly focused on labour market policies. In contrast, far less research has been devoted to the developments for an implementation of a European Union-wide social benefit scheme.

Besides the classic schemes of redistribution (i.e., towards the elderly, the unemployed, the sick) new solidaristic relationships are at the center of public debates. European respondents stand widely divided on new policy proposals, such as the implementation of an EU-wide benefit schemes.

The Welfare Attitudes module shows that 67,1% of Europeans express their support for an EU-wide social benefit scheme that would guarantee a minimum standard of living for the poor.

However, striking cross-national differences are present regarding these new proposals that challenge the foundations of the nationally bounded welfare state. In the more developed welfare states of Northern and Western Europe, there appears to be considerable reluctance to replace the existing arrangements. In Eastern and Southern Europe, dissatisfaction with current provisions is more widespread, and new proposals are looked at as an opportunity to improve living conditions. These findings evidence clear feedback effects of current institutional settings on welfare state legitimacy (see figure 2).

With regard to welfare opinions, we also find that countries often cluster in line with geographic regions, reflecting the fact that neighbouring countries tend to have similar levels of economic development and welfare systems (see figure 3).

The Eastern European group is less homogeneous, presumably because in the last 30 years the social policy reforms have rendered these welfare systems more as hybrids of different European regimes, rather than as an idiosyncratic regime type (Hacker, 2009).

These main research findings are also confirmed by the fitted logistic model. Finally, as previously seen, adults' socio-demographic, economic and work-related characteristics play an important role in order to be in favour of a public support for an EU-wide social benefit scheme.

## Acknowledgements

I would like to take this opportunity to thank the Amsterdam Institute for Advanced Labour Studies (AIAS) for the assistance I received during the visiting period spent in 2019. The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 730998, InGRID-2 – Integrating Research Infrastructure for European expertise on Inclusive Growth from data to policy.

## References

BONOLI G. 2005. The politics of new social policies: providing coverage against new social risks in mature welfare states. *Policy and Politics*, Vol. 33, No.3, pp. 431–449.

CARDONE P.E., DEIDDA M., MAROCCO M. 2019. *Le opinioni sulla condizionalità: i risultati in Italia dell'European Social Survey*. Inapp Paper no. 21, Roma, Inapp, p. 17.

EUROPEAN SOCIAL SURVEY Round 8 DATA 2016. *Data file edition 2.1*. NSD – Norwegian Centre for Research Data, Norway - Data Archive and distributor of ESS data for ESS ERIC.

FALKNER G. 2016. The European Union's social dimension. In M. CINI, & N.P.-S. BORRAGÁN (Eds.) *European Union Politics*, 5th ed., pp. 275–290. Oxford: Oxford University Press.

HACKER B. 2009. Hybridization instead of clustering: Transformation processes of welfare policies in Central and Eastern Europe. *Social Policy and Administration*, Vol. 43, No.2, pp. 152–169.

LIU X. 2016. *Applied Ordinal Logistic Regression using Stata.* Sage Publications, pp. 121-122.

OCHSNER M., RAVAZZINI L., GUGUSHVILI D., FINK M., GRAND P., LELKES O., VAN OORSCHOT W. 2018. *Russian versus European welfare attitudes: evidence from the 2016 European Social Survey*. London: European Social Survey.

STATA, *Software for Statistics and data Science*: https://www.stata.com.

## SUMMARY

## Public support for an EU-wide social benefit scheme: evidence from Round 8 of the European Social Survey (ESS)

Over the years, European Union has gradually assumed a more active role in social policymaking. This raises the question whether European respondents support this evolution, or whether they see the development of a Social Europe as a threat to their national welfare arrangements. The ESS Round8 module - Welfare Attitudes in a Changing Europe: Solidarities under Pressure - makes it possible to shed scientific light on welfare debates. The inclusion of this module addresses new solidarity questions as this new module also includes some questions fielded for the first time, most notably items assessing the introduction of a universal basic income scheme and the implementation of a European Union-wide social benefit scheme. Using logistic regression model, it is possible to estimate the different attitudes among countries for an EU-wide social benefit scheme more accurately. Striking cross-national differences are present regarding these new proposals that challenge the foundations of the nationally bounded welfare state. In the more developed welfare states of Northern and Western Europe, there appears to be considerable reluctance to replace the existing arrangements. In Eastern and Southern Europe, dissatisfaction with current provisions is more widespread, and new proposals are looked at as an opportunity to improve living conditions.


Paolo Emilio CARDONE, INAPP, p.cardone@inapp.org

# THE MARCHE REGION AND ITS INDUSTRY PATTERN: A QUANTITATIVE EVALUATION

Clio Ciaschini, Margherita Carlucci, Francesco Maria Chelli, Giuseppe Ricciardo Lamonica

## 1. Introduction

The economy of the Marche region is characterized by what Giorgio Fuà called the "Marche model", a set of small - medium sized businesses distributed throughout the territory, in particular on the coasts and valleys, and clustered in industrial districts, (Fuà, 1993). The main districts include furniture in Macerata and Pesaro, pharmaceutical and naval industries in Ancona, paper and appliance industries in Fabriano, musical instrument industry in the district of Castelfidardo and footwear, of considerable national, as well as regional, importance in the Macerata district. Starting from 2009 this so called “Marche model” begun to show some critical features which persist in the present times. As to the manufacturing industry, stagnation manifested in heterogeneous trends among the dimensional classes of firms also in relation to the typology of economic activity. Medium-large sized firms showed an increasing trend in the revenues as opposed to the decrease of revenues experienced by the smaller ones. The expansion of mechanic industry is in contraposition to the quinquennial footwear industry decay, (Banca d’Italia, 2019). The post-earthquake reconstruction has served as an engine for a partial recovery of the construction sector deeply impacted by the 2008 crisis, recovery that in any case does not reach the pre-crisis level. A not-brilliant frame is shaped also by the trend of the services sector. After having adapted to the Marche region the latest national Input-Output Table (NIOT) by means of the Flegg Location Quotient method (FLQ), an analysis on the position occupied by the economic sectors and the identification of the regional key sectors will be carried out. To this regards the Rasmussen (Rasmussen, 1956) approach will be used.

The FLQ contains a crucial unknown parameter ($0 \leq \delta < 1$) that must be estimated. On the basis of similar studies concerning Peterborough’s economy in 1968 (Morrison and Smith, 1974) and Scotland in 1989, (Flegg and Webber, 1997) an approximate value of $\delta = 0.3$ allows the derivation of closer multipliers to those obtained by surveys than multipliers obtained by the conventional cross-industry location quotients.

The present work is aimed to provide, in Section 2, a global vision of the industrial framework at a national level. Section 3 is devoted to the regionalization of the national matrix by means of the FLQ method (Lamonica and Chelli, 2018). Section 4 presents the empirical results for both Italy and the Marche region, while Section 5 is devoted to the discussion of the outcomes. Section 6 is dedicated to conclusions.

## 2. The Italian framework

World globalization in latest decades leads to the international development of manufacture industry. Starting from 2018, though, this development phase began to slowdown mainly due to conjuncture factors: inward-looking American commercial policies, a confused framework on the possible outcomes of Brexit, tensions between USA and China and risks connected to the results of the elections in Europe. Within this climate of instability, Italy has to face its weaknesses, especially those related to the internal market and to its various industries that found themselves deprived of the stimulus given by the international demand, which is now more fragile.

**Table 1 –** *Value added (2018) and Exports (2019) in percentage values of world's total.*

| | | | | |
|---|---|---|---|---|
| 1 | China | 28,5 | China | 15,1 |
| 2 | USA | 17,2 | Germany | 9,4 |
| 3 | Japan | 8,1 | USA | 8,1 |
| 4 | Germany | 6,1 | Japan | 4,4 |
| 5 | South -Korea | 3,1 | South- Korea | 3,9 |
| 6 | India | 3,0 | Hong Kong | 3,7 |
| 7 | **Italy** | **2,3** | France | 3,4 |
| 8 | France | 2,1 | The Netherlands | 3,2 |
| 9 | Unted Kingdom | 1,9 | **Italy** | **3,2** |
| 10 | Indonesia | 1,6 | United Kingdom | 2,8 |

*Source: Confindustria Report 2019*

This stagnation manifests itself not only under the point of view of the public component intrinsic in investments in infrastructures, but also in the private environment, even if supported by the incentives to the 4.0 digital conversion of manufacturing, (Confindustria Report, 2019). The concentration of industrial development towards new economic areas does not prevent Italy to be the seventh world manufacturing power, still in 2018 and the ninth country in relation to the export capacity (Table 1). The slowdown, which took place in 2017, has been common with the other European countries (France, Germany and Spain). A fundamental factor in favour of firms' development enriching the industry supply and improving the efficiency of productive systems is the 4.0 conversion of manufacturing. Technology 4.0 allows fastening the decision processes and new forms of interaction human-machine to connect the entire value chain within the firm.

Italy has joined the European framework of Industry 4.0 only in 2016, with a certain delay with respect to other European countries, through the National Plan Industria 4.0. The main measure used for the restart of firms has been that of hyper – amortisation which is estimated to account for 10 billion euros in investment.

## 3. The Methodology

Input-output table (IOT) provide information on the flows of goods and services among economic sectors of a country over a given period. It represents one of the most important tools for analysing the economic structure of a country and the relationships among specific economic sectors. A serious limitation in the construction of an IOT is the great volume of information required that is not always completely available. The same problems arise, in an amplified way, if the goal is to adapt the IOT of a country (NIOT) to a subnational region of interest (RIOT).

In this section, we briefly review the most used location quotient (LQ) methods to estimate a RIOT. In Table 2, we show the national and regional IOT for an economic system of k sector in block matrix notation:

**Table 2** – *Pattern of national and regional IOT.*

NIOT

$$
\begin{array}{|ccc|c|c|}
\hline
x_{11}^n & \cdots & x_{1k}^n & f_1^n & x_1^n \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
x_{k1}^n & \cdots & x_{kk}^n & f_k^n & x_k^n \\
\hline
v_1^n & \cdots & v_k^n & & \\
\hline
x_1^n & \cdots & x_k^n & & \\
\hline
\end{array}
$$

RIOT

$$
\begin{array}{|ccc|c|c|}
\hline
x_{11}^r & \cdots & x_{1k}^r & f_1^r & x_1^r \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
x_{k1}^r & \cdots & x_{kk}^r & f_k^r & x_k^r \\
\hline
imp_{11}^r & \cdots & imp_{1k}^r & f1_1^r & \\
\vdots & \ddots & \vdots & \vdots & \\
imp_{k1}^r & \cdots & imp_{kk}^r & f1_k^r & \\
\hline
v_1^r & \cdots & v_k^r & & \\
\hline
x_1^r & \cdots & x_k^r & & \\
\hline
\end{array}
$$

where:

– $\mathbf{X}^n = \left[x_{ij}^n\right]$ is the matrix whose entries are the total flows for intermediate use from the i-th sector to the j-th sector at national level;

- $\mathbf{X}^r=\left[x_{ij}^r\right]$ is the matrix whose entries are the flows for intermediate use from the i-th sector to the j-th sector at regional level (intraregional flows: both i and j sectors located in region r);
- $\mathbf{f}^n$ is the national final demand vector; $\mathbf{f}^r$ and $\mathbf{f1}^r$ are the regional final demand vectors of internal production and, respectively, imported from other regions;
- $\mathbf{IMP}^r=\left[imp_{ij}^r\right]$ is the matrix of imported intermediate inputs produced by the i-th sector of the other regions and acquired by the regional j-th sector;
- $(\mathbf{v}^n)'$ and $(\mathbf{v}^r)'$ are row vectors whose entries are the primary input (imports of goods and services and gross value-added components) by sector at national and regional level.

Moreover let $\mathbf{A}^n=\left[a_{ij}^n=\frac{x_{ij}^n}{x_j^n}\right]$, $\mathbf{R}=\left[r_{ij}=\frac{x_{ij}^r}{x_j^r}\right]$ and $\mathbf{M}^r=\left[m_{ij}^r=\frac{imp_{ij}^r}{x_j^r}\right]$ define the matrices whose entries are the national technical coefficients, the regional input coefficients and the regional import coefficients.

Assuming that only **NIOT** ($\mathbf{A}^n$**)** and the vector of the regional total sectorial output ($\mathbf{x}^r$) are known, the LQ methods estimate the matrix of the regional input coefficients **R** adjusting the national technical coefficient in the following way:

$$\hat{r}_{ij}=a_{ij}^n q_{ij} \tag{1}$$

where $q_{ij}$ represents the degree of modification of the national coefficient. Interregional import coefficients (the entries of $\mathbf{M}^r$) are estimated as difference between the national and the estimated regional input coefficient. The LQ methods are based on the assumption that the region has the same productive technologies of the nation:

$$a_{ij}^n=r_{ij}+m_{ij}^r \tag{2}$$

The most widely used LQ method is the *Simple Location Quotient* (SLQ) one. Here, the regional input coefficient is estimated as:

$$\hat{r}_{ij}=SLQ_i \cdot a_{ij}^n \tag{3}$$

where $SLQ_i$ is defined as:

$$SLQ_i=\frac{x_i^r/_{x^r}}{x_i^n/_{x^n}} \tag{4}$$

where $x_i^r$ and $x_i^n$ are the total output (production) of the i-th regional and national sector respectively. When the regional total output is not available, the sectorial employment can be used.

The previous ratio can be interpreted as the relative specialization of the region in the *i*-th sector compared to the nation. The $SLQ_i$ can be greater than, equal to, or less than one. When the Location Quotient is less than one, the corresponding regional sector is relatively less important than the same sector at national level. In this case, the regional sector will not be able to satisfy all local requirements, so that some of its products must be imported from other regions and no exports can be made. The interregional import coefficients ($m_{ij}$) are usually estimated from the difference between the national coefficient and the estimated regional input coefficient.

By contrast, if the Location Quotient is greater than or equal to one, the sector is judged able to fulfil all requirements of regional purchasing sectors. In other words, the region is self-sufficient for that activity or has a relative advantage. Hence, in these circumstances, the regional input coefficients are considered to be national technical coefficients. In this case, no adjustment is needed, and consequently the regional sector has the same input coefficient as the nation. Therefore, the regional input coefficients are adjusted in the following way:

$$\hat{r}_{ij}=\begin{cases} a_{ij}^n \cdot SLQ_i & \text{if } SLQ_i<1 \\ a_{ij}^n & \text{if } SLQ_i \geq 1 \end{cases} \tag{5}$$

One of the first enhancements of the SLQ method is the Cross-Industry Location Quotient (CILQ).

Indeed, the SLQ method is a uniform adjustment that takes into consideration only the supply side (the row side), i.e. only the size of the selling industry. Unlike the SLQ, the CILQ considers both supplying and purchasing sectors.

The CILQ formula can be written as follows:

$$CILQ_{ij}=\frac{x_i^r/x_i^n}{x_j^r/x_j^n}=\frac{SLQ_i}{SLQ_j}, \tag{6}$$

and

$$\hat{r}_{ij}=\begin{cases} a_{ij}^n CILQ_{ij} & \text{if } CILQ_{ij}<1 \\ a_{ij}^n & \text{if } CILQ_{ij} \geq 1 \end{cases}. \tag{7}$$

Contrary to the SLQ method, the CILQ method is a cell-by-cell adjustment.

The *symmetric cross-industry location quotient* (SCILQ) is a variant of the CILQ method. It was designed to take into account the possibility of deriving regional coefficients that exceed national values, thus overcoming the problem of asymmetric adjustments. It takes the following form:

$$SCILQ_{ij} = 2 - \frac{2}{CILQ_{ij}+1} . \tag{8}$$

The *semilogarithmic location quotient* (RLQ) incorporates the properties of both the SLQ and CILQ methods and takes the following form:

$$RLQ_{ij} = \frac{SLQ_i}{\log_2\left(1+SLQ_j\right)} = \frac{x_i^r/x^r}{x_i^r/x^r} \Big/ \left[\log_2\left(1+\frac{x_j^r}{x^r}\cdot\frac{x^n}{x_j^w}\right)\right]. \tag{9}$$

The RLQ has been criticized for underestimating imports from other regions when the size of the region is small. To overcome these drawbacks, the *Flegg location quotient method* (FLQ) was introduced. The key idea underlying the FLQ is that a region's propensity to import from other domestic regions is inversely and nonlinearly related to its relative size. By incorporating explicit adjustments for interregional trade, analysts should be able to gain more accurate estimates of regional input coefficients and hence multipliers. As with other non-survey techniques, the principal aim of the FLQ is to provide a means whereby regional analysts can construct regional tables that reflect a region's economic structure as much as possible:

$$FLQ_{ij} = \begin{cases} CILQ_{ij}\lambda & \text{for } i \neq j \\ SLQ_{ij}\lambda & \text{for } i = j \end{cases} , \tag{10}$$

where $\lambda$ stands for the relative size of the region and takes the following form:

$$\lambda = \left[\log_2\left(1+\frac{x^r}{x^n}\right)\right]^{\delta} . \tag{11}$$

Here, $\delta$ ($0 \leq \delta < 1$) is a sensitivity parameter that controls the degree of convexity in the previous equation. The larger the value of $\delta$, the lower the value of $\lambda$, so that greater adjustments of regional imports are made. Implementation of the FLQ formula is carried out in a manner similar to other LQ methods:

$$\hat{r}_{ij} = \begin{cases} a_{ij}^W FLQ_{ij} & \text{if } FLQ_{ij} < 1 \\ a_{ij}^W & \text{if } FLQ_{ij} \geq 1 \end{cases} . \tag{12}$$

The value of parameter δ is the focus of the method. McCann and Dewhurst (1998) pointed out that regional coefficients may exceed national coefficients when there is regional specialization (i.e., the regional coefficient becomes larger than the national coefficient). Thus, Flegg and Webber (2000) proposed the *augmented FLQ* (AFLQ). The AFLQ is defined as follows:

$$AFLQ_{ij}=\begin{cases} FLQ_{ij}\left[\log_2\left(1+SLQ_j\right)\right] & \text{for } SLQ_j>1 \\ FLQ_{ij} & \text{for } SLQ_j\leq 1 \end{cases}, \tag{13}$$

where $\log_2\left(1+SLQ_j\right)$ represents the regional specialization of sector j and has been included to allow for the effects of regional specialization. If $SLQ_j>1$ and $FLQ_{ij}\geq 1$, the national coefficients are scaled upwards. However, to avoid an excessive upward adjustment, the constraint $FLQ_{ij}\leq 1$ is imposed. Consequently, the regionalization is performed as follows:

$$\hat{r}_{ij}=\begin{cases} a_{ij}^{n}AFLQ_{ij} & \text{if } SLQ_j>1 \\ a_{ij}^{n}FLQ_{ij} & \text{if } SLQ_j\leq 1 \end{cases} \tag{14}$$

In consideration of the results obtained by Lamonica and Chelli (2018) and Lamonica et al. (2019), for the purpose of estimating the Marche IOT, the FLQ method with δ=0.3 was used. Moreover, due to the lack of regional data about the sectoral total employed, a reduced version of the Marche IOT for the year 2015 to 19 production sectors was considered.

## 4. Empirical results

After having estimated the Marche IOT by means of the FLQ method, to assess the position occupied by various economic sectors within the Marche economy the Rasmussen approach was considered. The core of this approach is the inverse Leontief matrix i.e. $\mathbf{L}=(\mathbf{I}-\mathbf{R})^{-1}$ where $\mathbf{R}$ is the matrix of regional direct input coefficients. The generic $L_{ij}$ entry of the $\mathbf{L}$ matrix measures the total requirement (multiplier), both direct and indirect, of goods and services produced by the i-th industry, which are necessary in order to satisfy one unit of final uses of the j-th sector. In other words, it measures the extent to which a unit increase in the final demand of the j-th sector causes a production increase in the i-th sector.

Consequently the j-th column-sum ($L_{.j}$) of $\mathbf{L}$ measures the total requirements needed by the j-th sector in order to produce one unit of final uses of its production;

or, the extent to which a one unit increase in the final demand of the j-th sector causes production increases in all sectors.

On the contrary, the row-sum of the **L** matrix ($L_{i.}$) measures the total production requirements of the i-th sector needed to off-set a unitary increase in final uses of each product. In other words, the magnitude of output increases in the i-th sector if final demand of all sectors increases by one unit.

Dividing $L_{.j}$ and $L_{i.}$ by the total number of sectors (k) yields the mean requirement (or the mean production increase) of the j-th sector ($L_{.j}/k$) and the mean requirement supplied by the i-th sector ($L_{i.}/k$). Alternatively, the mean impact on the economic system's production caused by a one unit increase in the final demand of the j-th sector and the mean impact on the i-th sector caused by a one unit increase in the final demand of all sectors. For the purposes of comparison, these two means are normalized with the general mean of all the elements in **L**:

$$\beta_j = \frac{L_{.j}/k}{i'\mathbf{L}i/k^2} \quad \text{for } j=i,..,k \tag{15}$$

$$\varphi_i = \frac{L_{i.}/k}{i'\mathbf{L}i/k^2} \quad \text{for } i=1,..,k \tag{16}$$

where $i$ is a vector of one, a prime (') denotes a row vector. The index (15), known as 'Backward linkage' (or power of dispersion), measures the degree of activation of an economic sector: the more this is greater than 1, the more the sector is important for the economy of the country considered, because it requires a production level by the other sectors in excess of the general mean. By contrast, the more the index falls below 1, the less important is the sector considered.

The other index (16), which is known as 'Forward linkage' (or sensitivity of dispersion), measures the level at which the output of one sector is used as input to the remaining productive sectors. It thus measures the degree of reaction of an economic sector. In this case, too, the more the index is greater than 1, the more important the corresponding sector is because it supplies its production to the others sectors at a level which exceeds the general mean. By contrast, the more the index falls below 1, the less important is the sector considered. The joint analysis of these two indices makes it possible to determine how an individual sector is woven into the economic structure of a country and how important it is.

The following Table 3 depict for the region Marche and the year 2015 the values of the two indices.

**Table 3 –** *Backward and forward indices for the Marche.*

| | | | |
|---|---|---|---|
| A(1) | Agriculture, forestry and fishing | 1.0341 | 0.8234 |
| B(2) | Mining and quarrying and other industry | 1.1057 | 0.728 |
| C(3) | Manufacturing | 1.4056 | 3.4121 |
| D(4) | Electricity, gas, steam and air-conditioning supply | 1.353 | 1.0357 |
| E(5) | Water supply, sewerage, waste management and remediation | 1.2637 | 0.9299 |
| F(6) | Construction | 1.2534 | 0.9372 |
| G(7) | Wholesale and retail trade, repair of motor vehicles and motorcycles | 0.978 | 1.2258 |
| H(8) | Transportation and storage | 1.0744 | 1.1732 |
| I(9) | Accommodation and food service activities | 1.0602 | 0.7105 |
| J(10) | Information and communication | 1.0592 | 0.9477 |
| K(11) | Financial and insurance activities | 0.8558 | 1.1708 |
| L(12) | Real estate activities | 0.6558 | 0.9012 |
| M(13) | Professional, scientific and technical activities | 0.9287 | 1.3225 |
| N(14) | Administrative and support service activities | 1.0849 | 0.9442 |
| O(15) | Public administration and defence, compulsory social security | 0.7206 | 0.5826 |
| P(16) | Education | 0.6839 | 0.5685 |
| Q(17) | Human health services | 0.9095 | 0.6281 |
| R(18) | Arts, entertainment and recreation | 1.1001 | 0.7894 |
| S(19) | Other services | 0.9173 | 0.6132 |
| T(20) | Activities of households as employers; undifferentiated goods and services-producing | 0.556 | 0.556 |

## 5. Discussion

In this paragraph we provide a short discussion of the results of the linkage analysis of the Marche region and compare them to the outcomes of a similar analysis performed for Italy. Results are shown in Figures 1 and 2.

In both cases the most relevant sector (or industry), so called *key industry* is sector C(3), Manufacturing, this sector is located in the first quadrant of both the graphs. Its relevance is higher regarding to Marche region, in fact in Figure 2 values of both backward and forward linkages are higher than those of Figure 1. This means that this sector sells to the other sectors materials for an amount higher than the average value of all the sectors and buys from the other sectors materials for a higher amount than all the other sectors. Another relevant issue to be mentioned is the fact that all the sectors linked to the Public Administration and public goods have both forward and backward linkages that lie under 1. Sectors K(11) - *Financial and insurance activities* and M(13) - *Professional, scientific and technical activities* in Figure 1 belong to the group of Forward Linkage Industries, that forward to the other sectors amounts higher than the average value of all sectors.

**Figure 1** – *Dispersion Analysis Backward and Forward Linkages – Italy.*

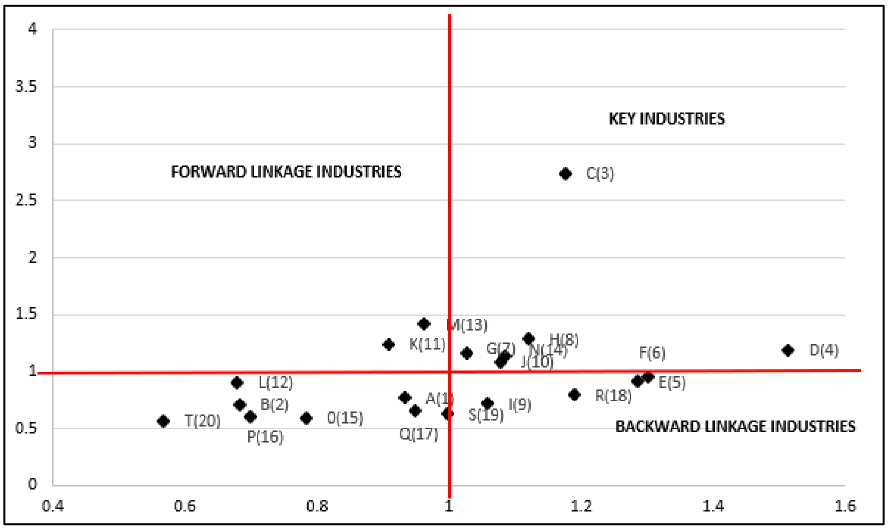
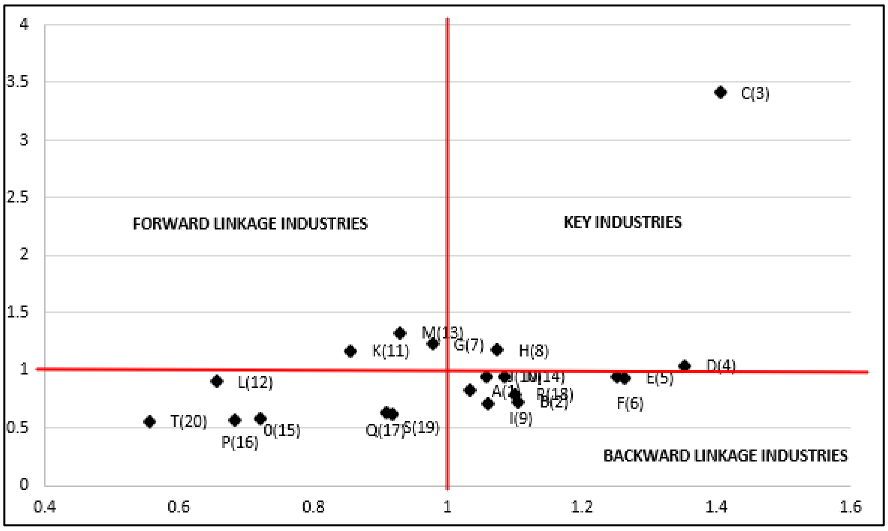



**Figure 2** – *Dispersion Analysis Backward and Forward Linkages – Marche.*



When moving to Figure 2 these two sectors remain in the group of Forward Linkage Industries but this group also includes sector G(7) - *Wholesale and retail trade, repair of motor vehicles and motorcycles*. In both the Figures we can see that sectors: E(5) - *Water supply, sewerage, waste management and remediation*, I(9) - *Accommodation and food service activities*, R(18) - Arts, entertainment and recreation and S(19) - *Other services* belong to the class of sectors that purchase intermediate goods in higher amounts per unit of product over the average (their Backward Linkage coefficient is higher than 1) but their Forward Linkage coefficient is lower than 1, these industries are called Backward Linkage Industries.

## 6. Conclusions

Our paper provides a multisectoral picture of the economic framework both in Italy and in the Marche region. The National Input-Output Table is regionalized using the Flegg Location Quotients method, since we think that this method is the most effective among the so-called *non-survey* methodologies of regionalization. Linkage analysis highlights manufacturing as the key sector, C(3), especially at the regional level, stressing its relevance in the Marche region also in relation to the already mentioned districts of the “Marche model”. The prominence of this industry both at national and regional level is expected to improve with the introduction within the production processes, of the Technology and Manufacturing 4.0. Industries linked to public goods are those that exhibit lower interactions; a possible introduction of Technology 4.0 is expected to improve all processes of Public Administration though higher digitalization and more efficient output supply.

## References

BANCA D’ITALIA 2019. *Economie regionali. L’economia delle Marche*. https://www.bancaditalia.it/pubblicazioni/economie-regionali/2019/2019-0033/1933-marche.pdf (accessed 21/07/2020).

CONFINDUSTRIA. 2019. *Dove va l’industria italiana. 2019. Rapporto sull’industria Italiana 2019*. https://www.confindustria.it/home/centro-studi/temi-di-ricerca/tendenze-delle-imprese-e-dei-sistemi-industriali/tutti/dettaglio/rapporto-industria+-italiana+-2019. (accessed 21/07/2020).

FLEGG, A. T., WEBBER, C. D. 1997. On the appropriate use of location quotients in generating regional input-output tables: reply, *Regional studies*, Vol.31, No.8, pp. 795-805.

FLEGG, A. T., WEBBER, C. D. 2000. Regional size, regional specialization and the FLQ formula, *Regional studies*, Vol. 34, No.6, pp. 563-569.

FUA', G. 1993. *Crescita economica: le insidie delle cifre*. Bologna. Il Mulino.

LAMONICA, G. R., RECCHIONI, M. C., CHELLI, F. M., SALVATI, L. 2019. The efficiency of the cross-entropy method when estimating the technical coefficients of input–output tables. *Spatial Economic Analysis*, Vol.15, No.1, pp. 62-91.

LAMONICA, G. R., CHELLI, F. M. 2018. The performance of non-survey techniques for constructing sub-territorial input-output tables. *Papers in Regional Science*, Vol. 97, No.4, pp. 1169-1202.

MC CANN, P., DEWHURST, J. H. L. 1998. Regional size, industrial location and input-output expenditure coefficients. *Regional Studies*, Vol. 32, No. 5, pp. 435-444.

MORRISON, W. I., SMITH, P. 1974. Nonsurvey input-output techniques at the small area level: An evaluation. *Journal of Regional Science*, Vol.14, No.1, pp. 1-14.

RASMUSSEN, P. N. 1956. *Studies in Intersectoral Relations*. North-Holland, Amsterdam.

## SUMMARY

## The Marche region and its industry pattern: a quantitative evaluation

The "Marche model" is the term used to define the industry structure of the Marche region, i.e. small- medium sized firms settled in coasts and valleys and organized in small districts. By means of Flegg Location Quotients, we regionalize the national Input – Output Table (NIOT) and obtain the Regional Table (RIOT). Through the linkage analysis we compare the relevance of the industrial pattern at a regional and national level. Manufacturing reveals as the key sector both at national and regional level. It has to be noted that the same industry results even more performing at the regional level.

---

Clio CIASCHINI, Università Politecnica delle Marche, c.ciaschini@staff.univpm.it
Margherita CARLUCCI, Università di Roma "La Sapienza", marherita.carlucci@uniroma1.it
Francesco M. CHELLI, Università Politecnica delle Marche, f.chelli@staff.univpm.it
Giuseppe RICCIARDO LAMONICA, Università Politecnica delle Marche, g.ricciardo@staff.univpm.it

# AN ANALYSIS ON CONSUMER PERCEPTIONS OF CORPORATE SOCIAL RESPONSIBILITY AND SUSTAINABLE CONSUMPTION

Gabriella Schoier, Giovanna Pegan

## 1. Introduction

The issue of corporate social responsibility (CSR) in the literature generally prioritized the point of view of managers or other stakeholders, such as regulators, corporate responsibility advocates, investors and the media (Pegan *et al.*, 2020; Sheth *et al.*, 2011; Martínez-Ferrero and García-Sánchez, 2015). CSR can be considered a mega trend that is widespread among executives who are increasingly aware of how their response to the challenge of sustainability can affect not only competitiveness but also the very survival of their business. Yet studies measuring CSR from the perspective of the consumer, a key stakeholder in pursuing sustainability objectives, are more limited and lack a homogeneous conceptualization of CSR that takes into account the three dimensions of sustainability: economic, environmental, and social (Jackson, 2009). The definition of the CSR construct, which will also be adopted in this research for the purpose of measuring consumers' perceptions, is as follows: "a firm's commitment to maximize long-term economic, societal and environmental well-being through business practices, policies and resources" (Du *et al*., 2011; Alvarado-Herrera *et al*., 2017). From this perspective, a company's objective of sustainability represents a triple responsibility in which the evaluation of the company's results focuses on the fusion of economic performance and environmental and social impact. This awareness seems widespread among many leading international companies in various sectors (Carroll, 1979; Frederick, 1986; Berens *et al.*, 2007). However, the emphasis placed on CSR research from a corporate perspective has overlooked the perception and engagement of the consumer, which by embodying multiple stakeholder identities such as citizen, parent, employee, and community member is crucial for sustainability efforts to succeed (Sheth *et al.* , 2011). On a global and cross-sector level, a consumer figure is becoming increasingly sensitive to the sustainability of consumption, and determined to influence business decisions with purchasing choices that reward brands willing to actively contribute to making the world a better and fairer place by supporting relevant social causes (Kotler *et al*., 2017; Schoier and De Luca, 2017). Therefore, consumer social responsibility can be interpreted as a pattern of behaviors that is adopted by the consumer to contribute to the achievement of sustainable development (Buerke *et al*., 2017), which also completes companies' CSR initiatives**.**

Given the limited number of studies exploring CSR from a consumer perspective and as an essential link for sustainability, the main objective of this paper is to contribute to the consumer perceptions of CSR and sustainable consumption by investigating the possible relationships between them. This study combines the three-dimensional conception of CSR (Alvarado-Herrera *et al.*, 2017) and the sustainability of consumption (Sharma and Jha, 2017).

This paper is formed by two parts. The first illustrates the theoretical background of the importance of investigating perceptions of CSR as a three-dimensional construct (environmental, social, and economic) from the perspective of the consumer and explores the crucial theme of sustainable consumption. In particular problems regard the society and its need for well-being, which, in this context, identifies social sustainability. The second presents the empirical research carried out on a sample of Italian consumers. The study investigates the relationships between sustainable consumption toward sportswear products and perceptions of the three dimensions of CSR of a well-known sustainable sportswear firm using, after a preliminary step, a two-step cluster analysis has been performed.

## 2. The Theoretical Framework

Nowadays, more and more often companies are seeking to resolve the consumer_citizens' accusations of being agents of strong social and natural imbalances through more ethical and transparent behavior aimed at the more sustainable development of society (Lee and Carroll, 2011; Martínez-Ferrero and García-Sánchez, 2015). For these reasons an increasing number of companies have started to promote strategies and actions based on CSR.

Today there is still no single definition of CSR; this is demonstrated by the numerous and varied terms and definitions that have been used over time to define it “corporate responsibility,” “corporate accountability,” and “corporate ethics” (European Commission, 2011; Arru and Ruggeri, 2016). The process of integrating CSR in strategies and actions is difficult and depends on several aspects, such as the size of the company, its business nature, and its corporate culture. In this perspective, the company's objective of sustainability should be translated into a triple responsibility, in which the evaluation of the company's results would depend on the synthesis of economic performance with its environmental and social impact (Pegan *et al.*, 2020).

The concept of triple responsibility was clearly explained by Elkington (1997), who defined it as a method of business management based on three specific needs: the company and its need to be profitable, which identifies economic sustainability; the environment and its need to be respected, which identifies environmental sustainability; and the society and its need for well-being, which, in this framework, identifies social sustainability.

One of the crucial aspects of CSR is that it must be communicable. Communicating means involving both primary and secondary stakeholders in the responsible actions the company undertakes. Therefore, it is essential for managers to understand what to communicate and to provide effective messages that are perceived by customer as truthful. To communicate CSR, companies can choose to adopt a general or specific message strategy. Whereas managers may prefer a general message, consumers tend to respond more favorably to specific messages. This paper analyses the theme of CSR from the perspective of consumer perception; it embraces the idea that CSR and consumer sustainability (consumption) are two sides of the same coin, and one must necessarily integrate and coordinate the other.

The adoption of sustainable behaviors is strongly determined by people's consumption choices in particular by sustainable consumption.

Consumer social responsibility for sustainable development is, therefore, to be understood as responsible consumer behavior, which must be coordinated with the efforts of CSR to contribute to the attainment of sustainable development (Buerke *et al.,* 2016).

Several authors (e.g. Brown and Dacin, 1997) have shown a positive relationship between CSR activities and the propensity to buy a brand or even reward it with loyalty only if the consumer perceives the corporate commitment as authentic and long term. Consumers are becoming more critical and skeptical about the authenticity of various firm proposals regarding environmental and social causes (Rozensher, 2013). So far the literature has not explored how the CSR of a company can be measured correctly to capture the real perception of the consumer in its many dimensions. It would also be interesting to investigate how CSR relates to sustainable purchasing attitudes and behaviors. The economic component may seem unrelated to the others because it concerns the company itself, whereas the social and environmental component may vary in intensity even on the basis of the sector under consideration.

## 3. The Methodological Analysis

In the present section we present the questionnaire, the descriptive analysis, the hypothesis development, and the results of the analysis.

### 3.1. The Questionnaire

From the methodological point of view a quantitative research has been developed using a structured questionnaire based on a seven-point Likert scale[1]; (see e.g. Tullis and Albert, 2013). The questionnaire has been tested on a small group of Italian consumers, and then revised for the final form, made via Internet on a randomly selected group, trying to ensure that the main social and demographic characteristics of the respondents (e.g., gender, age, and occupation) were diverse.

We have collected 207 questionnaires between August and November 2018.The questionnaire has been created using Google Drive modules and have the following structure:

- Part 1: a short description of the meaning of sustainable; generic environmental concern (EC) (three items); skepticism toward sustainable advertising (SSA) (three items) and perceptions of sustainable information usefulness (IU) (three items).
- Part 2: a short description of the meaning of sustainable sportswear products; attitudes toward sustainable sportswear products (ASSP) (four items); sustainable sportswear purchasing behaviors and sportswear purchasing habits. (SBSP) (four items).
- Part 3: knowledge of company X[2]; the multidimensional scale of CSR (CSRsoc (five items), CRSenv (six items), CSReco (six items) to measure consumer perception of CSR activities implemented by sportswear company X; the frequency of purchase of sportswear products in general and of the specific brand X; the frequency of purchase of sustainable sportswear products in general and of the specific brand X; and the degree of knowledge of the most recent CSR initiative communicated by the company X.
- Part 4: personal data, including gender, age, nationality, educational qualifications, profession, marital status, family members, perceptions of economic status compared to peers' average, and annual income (nine items).

[1] The scale presents the following answers: 1 = strongly disagree; 2 = disagree; 3 =somewhat disagree; 4 = neither agree nor disagree; 5 = somewhat agree; 6 = agree; 7 = strongly agree.

[2] The company name is omitted for privacy reasons.

In the next sections we consider a preliminary descriptive analysis followed by a two-step cluster analysis.

### 3.2. An Initial Descriptive Analysis

As regards the characteristics of the sample the 51.2% is formed by female and the 48.8% by male. Ages ranged from 17 to 85. The highest educational qualifications of the participants are: junior high school (5%); high school (53%); bachelor's degree and master's degree (36%); and PhD (6%). The jobs carried out by the sample are many and include: teachers, entrepreneurs, students, university professors, nurses, doctors, and managers. The 77% of the sample sample considers its economic situation, compared to its age peers in the average, 17% consider itself above the average and 6% below the average.

A number of descriptive statistical analyses were first done to provide an overview of the sample's responses in relation to the different parts of the questionnaire. The sample in general states a high level of Environmental concern (EC) with a median value of 6. Respondents are neutral about sustainable advertising Skepticism toward sustainable advertising (SSA).

Focusing on the consumer's perception of the three components of company X's CSR, the sample seems not to consider company X as socially responsible. In fact, CSRsoc achieved a median of 4, similar to CSRenv, whereas CSReco achieved a higher median of 5.5.

Regarding purchasing habits, the sample participants seem to be buyers of both sportswear products in general and sportswear products of brand X (median 5). Compared with the frequency of purchase of the category of sustainable products of company X, the respondents are more neutral with a median close to 4.

Some correlation analyses allowed to investigate the relationships between the measured constructs. For this purpose, Spearman's correlation coefficient and Tau-Kendal correlation coefficient have been used, the values are similar so in the next tables the first one has been presented.

The correlation analysis shows that the two groups of questions relating to ASSP and SBSP are strongly and positively correlated with each other, whereas the correlation between SBSP and EC is weaker.

**Table 1 –** *Correlations EC, ASSP, SBSP, IU, (Spearman index).*

| | EC | ASSP | SBSP | SSA | IU |
|---|---|---|---|---|---|
| EC | 1 | 0,491** | 0,539** | 0,105 | 0,416** |
| ASSP | | 1 | 0,716** | 0,051 | 0,393** |
| SBSP | | | 1 | 0,105 | 0,416** |
| SSA | | | | 1 | -0,017 |
| IU | | | | | 1 |

It is interesting to highlight the relationship between these measures and the perception of IU and SSA. As illustrated in Table 1, the SSA is not correlated with any of the examined constructs, whereas IU has significant and medium intensity correlations with EC, SASP, and SBSP.

Correlation analyses were then conducted to understand how the three responsibilities underlying the CSR components of the scale relate to one another, with the ASSP and SBSP (Table 2) and then with the perception of IU and SSA (Table 3). Table 2 shows that CSRsoc is strongly correlated with CSRenv, whereas both are weakly correlated with CSReco.

The EC is not related to any of the three responsibilities of company X, whereas there are significant but weak correlations between the first two dimensions of CSR with SASP. However, there is no correlation between CSRsoc, CSRenv, and CSReco.

The correlations that emerge with SBSP are similar; there is a significantly higher intensity correlation, but not with CSReco (Table 2). It is interesting to note that a general interest in protecting the environment among the respondents is not linked to the perception of specific activities that must be implemented by the company to pursue CSR. It is also important to underline the difference in the strength of the correlation between the two dimensions of CSRsoc and CSRenv, and between these and the CSReco dimension, where it is very weak.

**Table 2 –** *Correlations among EC, ASSP, SBSP, CSRsoc, CSRenv, CSReco, (Spearman index).*

| | EC | ASSP | SBSP | CSRsoc | CSRenv | CSReco |
|---|---|---|---|---|---|---|
| EC | 1 | 0,491** | 0,539** | 0,135 | 0,114 | -0,030 |
| ASSP | | 1 | 0,716** | 0,242** | 0,298** | 0,120 |
| SBSP | | | 1 | 0,283** | 0,320** | 0,094 |
| CSRsoc | | | | 1 | 0,837** | 0,245** |
| CSRenv | | | | | 1 | 0,212** |
| CSReco | | | | | | 1 |

The three components of company X's CSR are not correlated with the SSA, whereas there is a significant but weak correlation between these and the perception of UI by the respondents (Table 3).

**Table 3 –** *Correlations SSA, IU, CSRsoc, CSRenv, CSReco (Spearman index).*

| | SSA | IU | CSRsoc | CSRenv | CSReco |
|---|---|---|---|---|---|
| SSA | 1 | -0,017 | -0.012 | -0.052 | -0.143* |
| IU | | 1 | 0.261** | 0.253** | 0.032 |
| CSRsoc | | | 1 | 0,837** | 0,245** |
| CSRenv | | | | 1 | 0,212** |
| CSReco | | | | | 1 |

Further correlations were made to deepen the link between the three components of CSR and the purchases of both the brand in general and the sustainable product category of brand X. As can be seen from Table 4, there is a significant correlation of medium strength between CSRsoc and CSRenv of company X and the frequency of purchase of the sample sustainable sportswear product of brand X. Yet the correlations between the components and the frequency of purchase of brand X are significant but very weak. The CSReco is again not related at all to the purchase of the sustainable products of brand X, whereas there is a significant but very weak correlation with the general purchase of brand X.

**Table 4 –** *Correlations among customers' perception of CSR, frequency of brand X purchase, and purchase of sustainable sportswear products of brand X (Spearman index).*

| Constructs | Purchase brand X | Purchase sustainable sportswear products brand X |
|---|---|---|
| CSRsoc | 0,222** | 0,470** |
| CSRenv | 0,177* | 0,463** |
| CSReco | 0,191** | 0,064 |

### *3.3. A Cluster Analysis*

Various clustering algorithms have been developed to group data into clusters, however, they work effectively either on numerical or categorical data but most of them perform poorly on mixed categorical and numerical data. Two step cluster analysis allows to avoid this problem (Bacher *et al.*, 2004, Ming *et al.* , 2010).

SPSS two step clustering developed (Chiu *et al.* , 2001, Xu and Wunsch , 2005) for the analysis of large data sets consists of two steps:

• *Step1: Pre-clustering of cases.*

A sequential approach is used to pre-cluster the cases. The aim is to compute a new data matrix with fewer cases for the next step; in order to reach this aim, the computed pre-clusters and their characteristics (cluster features) are used as new cases. The pre-clusters are defined as dense regions in the analyzed attribute space.

The results may depend on the input order of cases therefore it is recommended to use random order.

• *Step2: Clustering of cases.*

A model based hierarchical technique is applied. Similar to agglomerative hierarchical techniques, the pre-clusters are merged stepwise until all clusters are in one cluster. In contrast to agglomerative hierarchical techniques, an underlying statistical model is used. The model assumes that the continuous variables are within clusters independent normal distributed and the categorical variables are within clusters independent multinomial distributed. Two distance measures are available: euclidean distance and a log-likelihood distance in case of mixed types of attributes.

As regards the cluster analysis some results are presented. The procedure determine the "best" number of clusters in two phases, in the first phase the Bayesian Information Criterion (BIC) or the Akaike's Information Criterion is computed ; the second phase uses the ratio change $R(k)$ in distance for $k$ clusters merged in $k-1$ clusters (see Bacher *et al.*, 2004). We have considered the Bayesian Information Criterion in so doing three clusters have been found.

Cluster 2 and Cluster 3 are mainly formed by females while Cluster 1 is formed mainly by males. Cluster 2 is composed mainly of young employees with a high school diploma or laurea while Cluster 3 is formed mainly by older people with a high school diploma working in different fields. Cluster1 is mainly formed by males both young and older with a high school diploma or laurea that work in different fields.

Cluster 2 is willing to sacrifice for the environment and is concerned about the environment itself; people belonging to this cluster want to try to improve the life of the local community.

People of Cluster 2 and Cluster 3 are more willing to make sacrifices for the environment, watch TV programs and websites dealing with the environment while not so people of Cluster 1.

Members of Cluster 1 and 3 like sustainable sportswear products more than those of Cluster 2. They are formed by consumers who take a more precise position with respect to the analyzed variables. They are composed mainly of consumers with more confidence in the commercial operations of the companies. They have the highest number of consensus as regards the credit in respect of the operations of marketing and CSR.

The groups composed mainly of adults that form Cluster 2 and Cluster 3 are more skeptical about the enviroment.

The results of the cluster analysis just exposed before show a certain tendency to the division of the opinions between the two generations identifiable as parents and children. The clusters “parents” are the most skeptical and cynical about the CSR activities of the company, probably due to a greater experience in life. Young people instead are those with a more positive vision of marketing, advertising and social, environmental and economical CSR, probably due to a greater propensity to trust the good conducts of companies. Indeed, young consumers are increasingly asking brands what they stand for, what values they embody, and are more likely to reward brands that take a stand on important issues, acting in the interests of society with concrete CSR actions.

## 4. Conclusions

In this paper we first illustrate the theoretical background of the importance of investigating perceptions of CSR as a three-dimensional construct (environmental, social, and economic) from the perspective of the consumer and explores the crucial theme of sustainable consumption. In particular problems regards the society and its need for well-being, which, in this context, identifies social sustainability. Then we present the empirical research carried out on a sample of Italian consumers.

The study investigates the relationships between sustainable consumption toward sportswear products and perceptions of the three dimensions of CSR of a well-known sustainable sportswear firm using, after a preliminary step, a two steps cluster analysis has been performed.

From a theoretical point of view, the results support the importance of studying CSR from a consumer perspective and its relationship with sustainable consumption in order to coordinate them together adopting a product-industry specific perspective (Pegan *et al.*, 2020). The results suggest that in order to better understand the multi-faceted phenomenon of sustainable consumption, it is better to study the relationships between consumer attitudes and behavior and adapt them to a specific context, as in the case of sportswear.

First of all, the fairly high value of the EC construct agrees with the trend that there is some sensitivity among consumers to environmental issues and consumption. This sensitivity diminishes when one moves toward measurements of specific sustainable purchasing attitudes and behaviors. A certain neutrality of uncertainty in perceiving the level of usefulness of sustainable UI and in judging the truthfulness and credibility of sustainable information within advertising SSA suggests a certain level of uncertainty. Similarly, the low frequency of purchases of sustainable sportswear products may also be due to the lack of clarity. These results may also express a difficulty for the consumer to translate the concept of sustainability in a concrete and industry-specific context. In addition, the results seem to support what scholars of attitudes have expressed about the importance of measuring sustainable attitudes and behaviors at the same level of abstraction. In fact, the correlations that emerged emphasize, on one hand, a general consistency between the various constructs of sustainable consumption (EC, ASSP, and SBSP), which reflect the greater attention to ethics in consumption; but, on the other hand, it also indicates that this consistency is much more evident between ASSP and SBSP.

Regarding CSR, the results of the analysis of customer perceptions concerning company X's CSR show that even though the company has been selected because of its award-winning and documented commitment to CSR, it is actually not perceived as such by the sample. Although they are buyers of the brand, respondents do not know how to assess its responsibility, as they attribute a neutral value to all CSRsoc and CSRenv items. This difficulty, as highlighted above, could be an expression of a lack of consumer knowledge and confusion due to too many ambiguous definitions of sustainable fashion-sportswear. The analysis of the correlations has highlighted how there is a significant correlation and quite high strength between CSRsoc and CSRenv and the propensity to buy sustainable products of the specific company X. The CSReco seems not to be related to purchases of the brand. The results also show a slightly different perception of respondents with respect to CSR, CSRsoc and CSRenv merge together, whereas CSReco is alone (Pegan *et al.*, 2020).

## References

ALVARADO-HERRERA A., BIGNE E., ALDAS-MANZANO J., CURRAS-PEREZ R. 2017. A Scale for Measuring Consumer Perceptions of Corporate Social Responsibility Following the Sustainable Development Paradigm, *Journal of Business Ethics*, Vol. 140, pp. 243-262.

ARRU B., RUGGIERI M. 2016. I benefici della Corporate Social Responsibility nella creazione di valore sostenibile: il ruolo delle risorse di competenza e del capitale reputazionale, *Economia aziendale online*, pp. 17-41.

BACHER J. K., WENZING M., VOGLER. M. 2004. SPSS Two Cluster – A First Evaluation, *Universitat Erlangen-Nurnberg*, pp. 1-20, www.statisticalinnovations.com/products/twostep.pdf cited July, 2020.

BERENS G., VAN RIEL C., VAN REKOM J. 2007. The CSR-quality trade-off: When can corporate social responsibility and corporate ability compensate each other? *Journal of Business Ethics*, Vol. 74, No. 3, pp. 233-252.

BROWN T. J., DACIN P. A. 1997. The company and the product: Corporate associations and consumer product responses, *Journal of Marketing*, Vol. 61, No.1, pp. 68-84.

BUERKE A., STRAATMANN T., LIN-HI N., MÜLLER K. 2017. Consumer awareness and sustainability focused value orientation as motivating factors of responsible consumer behavior, *Review of Managerial Science*, Vol .11, No. 4, pp. 959-991.

CARROLL A. B. 1979. A three-dimensional conceptual model of corporate social performance, *Academy of Management Review*, Vol 4, No. 4, pp. 497-505.

CHIU T., FANG D., CHEN J., WANG Y., JERIS C. 2001. A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Database Environment. In *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 263-268.

DU S., BHATTACHARYA C. B., SEN, S. 2011. Corporate social responsibility and competitive advantage: Overcoming the trust barrier, *Management Science*, Vol. 57, No. 9, pp. 1528-1545.

ELKINGTON J. 1997. *Cannibals with forks - Triple bottom line of 21st century business*. Stoney Creek, CT: New Society Publishers.

EUROPEAN COMMISSION 2011. Green Paper - Promoting a European framework for corporate social responsibility, Brussels European Commission. Renewed EU strategy for 2011-2014 on corporate social responsibility. Brussels: Belgium.

FREDERICK C. W. 1986. Towards CSR: Why ethical analysis is indispensable and unavoidable in Corporate affairs, *California Management Review*, Vol. 28, No. 2, pp. 126-141.

JACKSON T. 2009. *Prosperity without growth: Economics for a finite planet.* London: Earthscan.

KOTLER P., KARTAJAYA H., SETIAWAN I. 2017. *Marketing 4.0.* Milano, Italy: Hoepli.

LEE S. Y., CARROLL C. E. 2011. The emergence, variation, and evolution of corporate social responsibility in the public sphere 1980-2004: The exposure of firms to public debate, *Journal of Business Ethics*, Vol. 104 No. 1, pp. 115-131.

MARTÍNEZ-FERRERO J., GARCÍA-SÁNCHEZ I. 2015. Is corporate social responsibility an entrenchment strategy? Evidence in stakeholder protection environments, *Review of Managerial Science*, Vol. 9, No.1, pp. 89-114.

MING-YI S., JAR-WEN J., LIEN-FU L. 2010. A Two-Step Method for Clustering Mixed Categorical and Numeric Data, *Tamkang Journal of Science and Engineering*, Vol. 13, No. 1, pp. 11-19.

PEGAN G., SCHOIER G., DE LUCA P. 2020, The Importance of Consumer Perception of Corporate Social Responsibility to Meet the Need for Sustainable Consumption: Challenges in the Sportswear Sector. In C. Silvestri, M. Piccarozzi, and B. Aquilani (Eds.), *Customer Satisfaction and Sustainability Initiatives in the Fourth Industrial Revolution*, Hershey, PA: IGI Global, pp. 212-235.

ROZENSHER S. 2013. The growth of cause marketing: past, current, and future trends, *Journal of Business & Economics Research*, Vol. 11, No. 4, pp.181-186.

SCHOIER G., DE LUCA P. 2017. Cause-Related Marketing: A Qualitative and Quantitative Analysis on Pinkwashing. In Palumbo, F., Montanari A., Vichi, M. (Eds.), *Data Science. Innovative Developments in Data Analysis and Clustering*, pp. 321-332, New York, NY: Springer.

SHARMA R., JHA M., 2017, Values influencing sustainable consumption behaviour: Exploring the contextual relationship, *Journal of Business research*, Vol. 76, pp. 77-88.

SHETH J.N., SETHIA N.K., SRINIVAS S. 2011. Mindful consumption: A customer-centric approach to sustainability, *Journal of the Academy of Marketing Science*, Vol. 39, No. 1, pp. 21-39.

TULLIS T., ALBERT B. 2013. *Measuring the user experience*. Amsterdam: Elsevier.

XU R., WUNSCH D. 2005. Survey of clustering algorithms. *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 16, No. 3, pp. 645-678.

## SUMMARY

### An analysis on consumer perceptions of corporate social responsibility and sustainable consumption

The aim of this paper is to analyze consumer perception of corporate social responsibility (CSR) and sustainable consumption by investigating the possible relationships between them. A theoretical overview of the importance of investigating the perceptions of CSR as a three-dimensional construct from a consumer perspective and the crucial theme of sustainable consumption are presented. At this point the results of a quantitative empirical research carried out on Italian consumers are presented. The starting point has been a structured questionnaire. Through this a descriptive analysis also using a two step cluster analysis has been performed. The results confirm the relevance of deepening consumer perceptions of CSR by using a multidimensional scale and its relationship with sustainable consumption while focusing on a specific firm and category of sustainable products.

Gabriella SCHOIER, Università di Trieste, gabriella.schoier@deams.units.it
Giovanna PEGAN, Università di Trieste, giovanna.pegan@deams.units.it

# LA QUALITÀ NEI PROCESSI DI DATA CAPTURING. IL CASO DELL'INDAGINE SUGLI ASPETTI DELLA VITA QUOTIDIANA

Claudio Ceccarelli, Marco Fortini, Manuela Murgia, Alessandra Nuccitelli, Rita Ranaldi, Francesca Rossetti

## 1. Introduzione

La qualità dell'informazione prodotta da un'indagine statistica dipende da numerosi fattori: definizioni e concetti adottati, disegno di campionamento, tecniche di indagine, strumenti, via via fino alle modalità di analisi e rappresentazione dei dati. Il concetto di Total Survey Error (TSE) è sviluppato secondo questo principio per legare la qualità di ciascuna fase di processo a quella dell'informazione prodotta, in modo da poter monitorare e intervenire sulle attività di indagine che risultano più critiche in un'ottica di rapporto tra costi e benefici (Lyberg and Stukel, 2017). In questo approccio la qualità dell'informazione è rappresentata dalla somma degli scostamenti, in termini di distorsione e variabilità, tra il valore stimato e il valore vero, che intervengono durante le diverse fasi del processo.

Questo lavoro si concentra sul processo di raccolta dei dati, determinante nella misura in cui contenerne l'errore contribuisce in modo sostanziale a migliorare la qualità complessiva dell'informazione prodotta.

La prestazione del processo di raccolta dei dati può essere valutata tramite dati riguardanti l'esecuzione delle fasi di processo, i cosiddetti paradati (Kreuter, 2013). Prendono il nome di paradati tutte le informazioni sull'andamento del processo di produzione di un'indagine statistica raccolte nel corso del processo stesso. Il contatto, la partecipazione, il numero di solleciti, la modalità e la durata dell'intervista (o della compilazione) sono tutti esempi di paradati; ad essi si affiancano altre informazioni trattate in occasione di un'indagine e costituite da: dati; metadati; dati ausiliari; macrodati (o stime).

L'analisi dei paradati e della loro associazione con le altre informazioni disponibili si dimostra utile a costruire indicatori di prestazione delle fasi di indagine per orientare i necessari aggiustamenti di processo.

Il classico esempio di paradato è costituito dalla variabile indicatrice di risposta dell'unità statistica, a partire dalla quale si può calcolare il tasso di risposta, inteso come rapporto tra il numero di rispondenti e il totale delle unità statistiche selezionate nel campione. Se il tasso di risposta è costante rispetto a una data variabile di interesse Y, allora la mancata risposta si definisce 'ignorabile' rispetto a Y e l'analisi riferita ai soli rispondenti risulterà non distorta. Dato che questa misura non si può ottenere in modo diretto perché Y nella pratica non è nota sui non rispondenti, si può ripiegare su un vettore X di variabili ausiliarie e/o altri paradati, noto per tutte le unità del campione, condizionatamente al quale la variabile Y sia indipendente dal tasso di risposta. In questo caso si dirà che la mancata risposta è ignorabile rispetto a Y, condizionatamente al vettore X (Little and Rubin, 2002).

Mentre l'associazione tra il vettore X e il tasso di risposta è studiata sui dati in essere, l'indipendenza condizionata tra il tasso di risposta e la variabile Y viene postulata sulla base di informazioni e conoscenze esterne. L'analisi del tasso di risposta (o propensione alla risposta) condizionatamente al profilo dei rispondenti aiuta a identificare i gruppi più a rischio di mancata risposta ed è utile sia ai fini di calibrazione delle stime, sia per interventi correttivi da effettuare in corso d'indagine, sia, infine, per la revisione del disegno di indagine nel suo complesso.

Analogamente alla propensione alla risposta è possibile considerare altri paradati, come:

- eleggibilità, che misura la quantità e la tipologia dei casi non eleggibili erroneamente inclusi nelle liste di indagine;
- tecnica di indagine (per le indagini *mixed-mode*), per studiare se la modalità di risposta influenza il risultato della misura;
- numero di solleciti, per studiare quanti farne e su quali sottogruppi;
- profili di risposta critici, per tenere sotto controllo particolari tendenze come, ad esempio, segnalare sempre la prima modalità di risposta o privilegiare traiettorie che abbreviano il tempo di compilazione del questionario.

La variabilità dei paradati può anche essere messa in relazione a gruppi significativi, come le unità statistiche assegnate ad uno stesso rilevatore o a un ufficio territoriale, per evidenziare criticità locali sulle quali è opportuno intervenire.

Parte di queste analisi mira a creare correttori delle stime tramite modelli che tengono conto delle diverse fonti di errori campionari e non campionari. Tale approccio è complicato dalla necessità di disporre di informazione ridondante, di ricorrere a ipotesi spesso difficili da verificare in pratica, considerando per giunta solo una o poche fonti di errore alla volta. In questa sede si considerano, piuttosto, le determinanti della qualità per migliorare le future occasioni di indagine o per modificare in maniera flessibile l'indagine in corso, raccogliendo all'origine dati più completi ed affidabili a parità di costo.

Come esempio dell'approccio proposto, viene descritta un'analisi della propensione a rispondere oppure a utilizzare una delle due tecniche previste (CAWI

[1] o CAPI-PAP[2]) per l'indagine sugli 'Aspetti della vita quotidiana' (AVQ) condotta nel 2019. I paradati in questione sono messi in relazione a profili significativi delle unità di rilevazione per suggerire adattamenti nella strategia di contatto e sollecito.

Nel paragrafo 2 si descrivono l'indagine e i dati utilizzati, mentre nel terzo viene illustrato il modello logistico adottato per l'analisi e si presentano i principali risultati. Infine, nel paragrafo 4 sono delineati i possibili metodi e strumenti per il miglioramento della qualità della rilevazione nell'ambito di una strategia che coinvolge i diversi attori che ruotano intorno al complesso processo di indagine.

## 2. Dati

L'analisi presentata nel paragrafo 3 è stata effettuata sfruttando le informazioni disponibili relativamente alla fase di raccolta dei dati dell'Indagine Multiscopo sulle Famiglie: Aspetti della vita quotidiana (AVQ), con riferimento all'anno 2019. Questa indagine rientra nel Programma statistico nazionale ed ha lo scopo di rilevare annualmente le informazioni necessarie per conoscere le abitudini dei cittadini e i problemi che essi affrontano ogni giorno e se sono soddisfatti del funzionamento di quei servizi di pubblica utilità che dovrebbero contribuire al miglioramento della qualità della vita.

Le informazioni vengono raccolte attraverso una tecnica mista, che si avvale di un questionario online da auto-compilare da parte dei rispondenti (tecnica CAWI) oppure di un'intervista diretta con questionario elettronico, somministrato da un intervistatore, e contestuale consegna di un questionario cartaceo da auto-compilare da parte dei componenti della famiglia (tecnica CAPI-PAP). In particolare, le due tecniche sono applicate in sequenza: per l'indagine dell'anno 2019, infatti, le famiglie hanno potuto compilare il questionario online (tecnica CAWI) nel periodo dal 28 febbraio al 31 marzo 2019, utilizzando le credenziali riportate nella lettera di invito. Qualora una famiglia non avesse avuto la possibilità di rispondere all'indagine tramite Internet, al termine del periodo previsto un intervistatore comunale si è recato presso l'abitazione della famiglia stessa (tecnica CAPI-PAP), per rivolgere le stesse domande del questionario online a tutti i suoi componenti. Sono stati fatti due solleciti postali durante il periodo previsto per la compilazione con tecnica CAWI; per le famiglie che avevano già fatto accesso al questionario web senza concludere l'intervista, i solleciti sono stati effettuati tramite e-mail.

---

[1] Computer-Assisted Web Interviewing.

[2] Computer-Assisted Personal Interviewing e Paper and Pencil.

Il campione teorico del 2019 è formato da 25.177 famiglie residenti in 783 comuni. In particolare, rispetto al totale delle famiglie del campione, il 35% circa ha fatto accesso al portale web e, fra queste, quasi il 90% ha portato a termine l'intervista con tecnica CAWI. Alla fine del periodo previsto per la rilevazione con tecnica CAWI, le famiglie non rispondenti – ad esclusione di quelle che, pur avendo contattato il *contact center*, sono risultate essere non eleggibili e di quelle che avevano iniziato la compilazione del questionario senza concluderla – sono state oggetto di intervista con tecnica CAPI-PAP; delle 16.348 famiglie complessive coinvolte, il 71,5% ha terminato l'intervista.

In questa prima sperimentazione sono stati utilizzati i soli dati disponibili e/o raccolti nell'ambito della fase di conduzione d'indagine: per ciascuna famiglia del campione teorico sono disponibili informazioni ausiliarie relative alla famiglia anagrafica (ad esempio, il numero di componenti, la dimensione demografica del comune di residenza della famiglia, la provincia di residenza) o relative al singolo membro della famiglia (il sesso, l'età, la cittadinanza, il paese e la provincia di nascita, il ruolo nella famiglia[3]). Le informazioni raccolte attraverso il questionario non sono al momento disponibili.

La rilevazione è stata condotta avvalendosi di due differenti sistemi di gestione di indagine, uno per la parte CAWI e uno combinato per la parte CAPI-PAP. Il sistema di gestione della tecnica CAPI-PAP prevede la raccolta delle informazioni relative ai tentativi di contatto, al numero di visite effettuate dall'intervistatore presso la famiglia ma, non essendo l'intervistatore obbligato alla registrazione, questi dati possono essere incompleti o parziali: la loro disponibilità potrebbe essere invece utile per integrare e approfondire le analisi sulla propensione alla risposta da parte delle famiglie.

È importante inoltre ricordare che circa il 97% (24.465) delle famiglie del campione di AVQ era già presente nel campione dell'Indagine da lista (L) della rilevazione censuaria 2019 sulla popolazione; le restanti famiglie sono state estratte dalle Liste anagrafiche comunali (Lac) del comune di appartenenza, in quanto il comune nel Censimento era incluso nella sola indagine areale. La presenza della quasi totalità delle famiglie nelle due indagini campionarie permetterà, in una fase successiva di analisi, l'accesso ad un maggior numero di variabili ausiliarie e quindi l'applicazione di altri strumenti di analisi e modelli della non-risposta.

[3] Intestatario della Scheda di Famiglia (ISF), coniuge dell'ISF o altro membro.

## 3. Analisi della propensione alla risposta

### *3.1. Modello utilizzato*

L'analisi dei dati riguarda le 23.093 famiglie eleggibili[4] dell'indagine AVQ per l'anno 2019 per le quali si studia l'associazione tra il comportamento all'intervista – "Risponde con tecnica CAWI" (CW), "Risponde con tecnica CAPI-PAP" (CP), "Non risponde" (NR) – e alcune caratteristiche disponibili nella lista di partenza. A questo scopo viene utilizzato un modello di regressione logistica multinomiale (o politomica) ad effetti fissi.

Le variabili ausiliarie inserite nel modello sono le seguenti[5]:

- *Ripartizione geografica* ("Nord-ovest"; "Nord-est"; "Centro"; "Mezzogiorno");
- *Classe di popolazione del comune* ("<=5.000 abitanti"; "5.001-50.000 abitanti"; "50.001-150.000 abitanti"; ">=150.001 abitanti");
- *Cittadinanza dell'ISF* ("Italiana"; "Straniera")
- *Età dell'ISF* ("<50; ">=50");
- *Numero di componenti della famiglia anagrafica* ("<=2"; ">2").

Un'altra caratteristica riferita all'ISF, come il *Sesso*, non figura nel modello finale in quanto è risultata avere un impatto debolmente significativo sul fenomeno in esame.

Il modello utilizzato – privo di effetti interattivi – si articola in due equazioni, cioè quante sono le categorie di comportamento previste, al netto della categoria scelta come riferimento (CP):

$$\begin{aligned}\ln\left(\frac{\Pr(Y_i = \mathrm{CW}|\mathbf{x}_i)}{\Pr(Y_i = \mathrm{CP}|\mathbf{x}_i)}\right) &= \alpha_{\mathrm{CW}} + \beta_{\mathrm{CW}}^{11}x_i^{11} + \beta_{\mathrm{CW}}^{21}x_i^{21} + \cdots + \beta_{\mathrm{CW}}^{53}x_i^{53} \\ \ln\left(\frac{\Pr(Y_i = \mathrm{NR}|\mathbf{x}_i)}{\Pr(Y_i = \mathrm{CP}|\mathbf{x}_i)}\right) &= \alpha_{\mathrm{NR}} + \beta_{\mathrm{NR}}^{11}x_i^{11} + \beta_{\mathrm{NR}}^{21}x_i^{21} + \cdots + \beta_{\mathrm{NR}}^{53}x_i^{53}\end{aligned} \tag{1}$$

avendo indicato con $\mathbf{x}_i = (x_i^{11}, x_i^{21}, x_i^{31}, x_i^{41}, x_i^{42}, x_i^{43}, x_i^{51}, x_i^{52}, x_\mathrm{i}^{53})$ il vettore dei regressori indicatori[6] osservati sulla famiglia $i$ – relativi, nell'ordine, alle variabili

[4] Sono state escluse dall'analisi anche le famiglie che non è stato possibile contattare per errori di lista.

[5] La classificazione dell'*Età dell'ISF* e del *Numero dei componenti della famiglia anagrafica* è stata scelta sulle base di analisi preliminari dell'associazione con la variabile dipendente.

[6] Ogni variabile esplicativa viene qui rappresentata utilizzando tanti regressori indicatori quante sono le modalità della variabile meno una, scelta come categoria di riferimento. Ad esempio, la variabile *Ripartizione geografica,* con quattro modalità, viene rappresentata nel modello utilizzando i tre regressori indicatori seguenti: $X^{51}$ (= 1 se Centro, 0 altrimenti), $X^{52}$ (= 1 se Nord-est, 0 altrimenti), $X^{53}$ (= 1 se Nord-ovest, 0 altrimenti), avendo assunto "Mezzogiorno" come categoria di riferimento.

*Numero di componenti della famiglia anagrafica*, *Età dell'ISF*, *Cittadinanza dell'ISF*, *Classe di popolazione del comune*, *Ripartizione geografica* – e con le lettere greche i parametri da stimare.

### 3.2. Risultati

Nella Tabella 1 è riportata la variazione dell'adattamento del modello ai dati ottenuta tralasciando una variabile esplicativa alla volta e mantenendo invece le altre (statistica del Chi-quadrato), con il *p-value* associato.

I risultati mostrano che ciascuna delle variabili ha un effetto significativo sulla propensione di una famiglia a un certo comportamento (colonna 3); in particolare, il livello della statistica del Chi quadrato (colonna 2) evidenzia che la *Ripartizione geografica* è il fattore più rilevante tra quelli considerati. Seguono, staccati per importanza, la *Classe di popolazione del comune* e la *Cittadinanza dell'ISF*. Infine, l'*Età dell'ISF* e il *Numero di componenti della famiglia anagrafica*, sembrano avere un impatto più limitato sul fenomeno in esame.

**Tabella 1 –** *Variabili esplicative per statistica del Chi-quadrato di Wald e p-value associato.*

| Variabile | Chi quadrato di Wald | Pr > ChiQuad |
|---|---|---|
| Ripartizione geografica | 1.461,81 | <,0001 |
| Classe di popolazione del comune | 489,30 | <,0001 |
| Cittadinanza dell'ISF | 393,45 | <,0001 |
| Età dell'ISF | 47,29 | <,0001 |
| Numero di componenti della famiglia anagrafica | 47,22 | <,0001 |

*Fonte:* nostre elaborazioni su dati Istat

Quando tutte le variabili coinvolte nel modello sono categoriche, come nel caso in questione, l'interpretazione dei risultati può risultare più immediata esaminando direttamente gli *Odds Ratio*[7] (OR), anziché i parametri in (1).

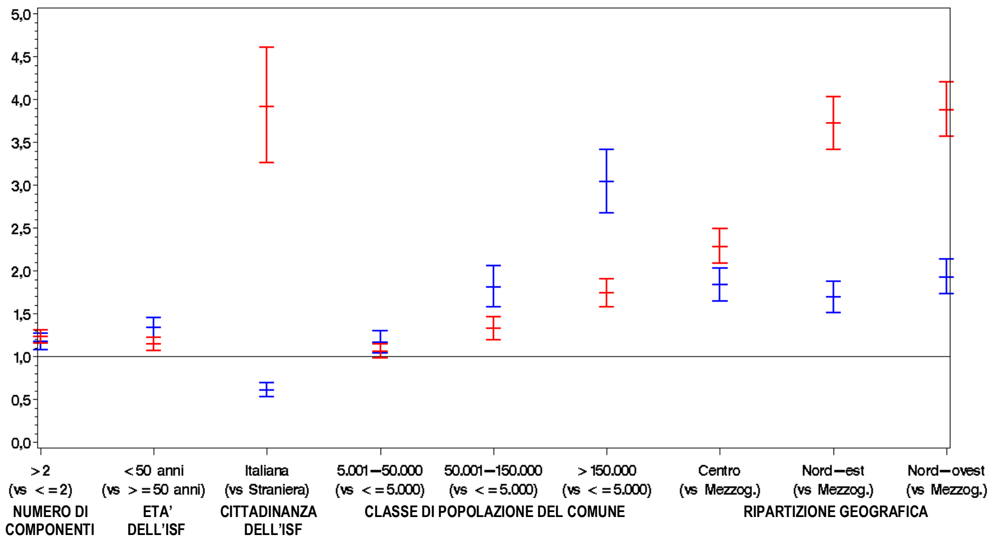
Nella Figura 1 per ciascuna variabile esplicativa sono riportate le stime di massima verosimiglianza degli OR e dei rispettivi intervalli di confidenza al 95%, relative alle due equazioni del modello. Tali stime permettono una caratterizzazione

[7] L'OR non è altro che il rapporto tra gli *Odds* relativi al confronto tra due categorie della variabile risposta – ad esempio, CW e CP – in due situazioni alternative di una delle variabili esplicative (ad esempio, "Centro" e "Mezzogiorno", se si considera la variabile *Ripartizione geografica*). In altre parole, l'OR permette di valutare immediatamente di quanto cresce o decresce il rischio che la variabile di risposta assuma una certa categoria (CW), anziché un'altra scelta come riferimento (CP), a seguito della variazione del valore assunto dalla variabile esplicativa – ad esempio, da "Mezzogiorno" a "Centro", se si sceglie "Mezzogiorno" come modalità di riferimento – e al netto degli effetti di tutte le altre variabili.

delle famiglie secondo la propensione alle varie tipologie di comportamento prese in considerazione.

Nel contesto in esame, un interesse specifico è rivestito dalla categoria CW, che gioca un ruolo particolarmente importante nel contenimento del tasso di mancata risposta e dell'errore di misurazione e, più in generale, nel miglioramento del processo di raccolta dei dati.

**Figura 1 –** *Odds Ratio e intervalli di confidenza al 95% per i regressori indicatori relativi al modello (1)*




*Nota:* *la modalità di riferimento per ogni variabile ausiliaria è indicata tra parentesi
*Fonte:* nostre elaborazioni su dati Istat

Specificamente, la propensione a rispondere con tecnica CAWI, piuttosto che con modalità CAPI-PAP, risulta maggiore per le famiglie:

- del Nord (OR superiori a 3,70) e, in misura più contenuta, per quelle del Centro (OR = 2,28);
- con ISF italiano (OR = 3,83), soprattutto se ha un'età inferiore ai 50 anni (OR = 1,15);
- di almeno tre componenti (OR = 1,23).

Per di più, la propensione alla tecnica CAWI tende a crescere all'aumentare della dimensione comunale (OR = 1,73 per la classe di popolazione più ampia).

Tuttavia, all'aumentare della dimensione comunale, si riscontra una tendenza ancora più marcata a non rispondere all'indagine (OR pari a 1,16 e 1,80 per le classi di popolazione intermedie; superiore a 3 per le famiglie che vivono in comuni con più di 150.000 abitanti). Inoltre, il rischio di mancata risposta risulta più elevato per le famiglie:

- con ISF straniero (OR = 0,61), soprattutto se ha un'età inferiore ai 50 anni (OR = 1,34);
- di almeno tre componenti (OR = 1,17).

Infine, vale la pena evidenziare che le famiglie del Mezzogiorno mostrano un comportamento più incline alla risposta, con una maggiore propensione ad utilizzare la tecnica CAPI-PAP.

Alla luce dei risultati ottenuti, per la prossima edizione dell'indagine, potrebbe essere proposta una modifica dell'impianto di rilevazione, passando da un approccio puramente sequenziale del *mixed-mode* a uno concorrente in cui le tecniche sono in campo contemporaneamente e invitando al CAWI o al CAPI-PAP i profili di popolazione più propensi a ciascuna tecnica. Inoltre, per favorire l'utilizzo della modalità CAWI, si potrebbe valutare l'introduzione di una strategia di sollecito (alternativa alla lettera cartacea o all'e-mail) basata su solleciti telefonici ad opera di un *contact center* esterno o direttamente da parte dei rilevatori comunali, eventualmente mirata a specifici segmenti di popolazione.

## 4. Discussione e sviluppi futuri

L'esempio presentato nel paragrafo precedente è una semplice ma efficace descrizione di come le scelte da effettuare in fase di raccolta dati possano essere supportate da risultati quantitativi derivanti da modelli statistici che interpretano il processo di risposta in funzione di variabili ausiliarie note per tutta la popolazione di riferimento. Lo stesso tipo di analisi potrebbe ovviamente essere sfruttata a monte, ossia nella fase di progettazione di una nuova edizione di indagine, per migliorare quegli aspetti del processo che hanno influito negativamente sulla partecipazione all'indagine o per sperimentare l'uso di disegni adattivi, che sono finalizzati a ottenere una risposta più bilanciata individuando i sotto-processi della raccolta dati e i sottogruppi di unità con un impatto maggiore sulla rappresentatività della risposta (Bethlehem *et al.*, 2011).

È possibile usare vari modelli statistici che, operando sotto condizioni diverse (Cobben, 2009), offrono differenti spunti di osservazione del processo di raccolta dati e, di conseguenza, suggeriscono l'adozione di azioni diverse, ma tutte mirate a effettuare quegli interventi correttivi che, nel migliorare il tasso di risposta in termini di entità e composizione, cercano anche di ottimizzare i costi e l'impegno di risorse umane dedicate alla conduzione e al monitoraggio della rilevazione. Rientrano in queste azioni anche i *responsive design* che hanno la stessa finalità dei disegni adattivi, ma, a differenza di questi, individuano i sotto-processi di raccolta dati e i sottogruppi di unità in base alle informazioni raccolte via via durante la rilevazione stessa. Modelli logistici semplici o innestati, modelli a classi latenti, modelli *multilevel* o modelli di sopravvivenza sono alcuni esempi di applicazioni possibili finalizzate a identificare effetti di autocorrelazione nella risposta indotti dagli enti territoriali e/o dagli organi intermedi di rilevazione (comuni e rilevatori) o la presenza di una variabilità residua spiegabile da informazioni di cui non si sta tenendo conto.

Requisito fondamentale per la costruzione di qualsivoglia modello che vada nell'ottica descritta è, come già detto nel paragrafo 1, la disponibilità di paradati e variabili ausiliarie, ossia di quegli 'ingredienti' imprescindibili per l'analisi del processo di indagine. È quindi importante che l'organizzazione della raccolta dati sia tale da garantire la memorizzazione tempestiva e sistematica di queste informazioni e che queste vengano definite durante la progettazione dell'indagine stessa. Si pensi, ad esempio a tutte le informazioni relative alla conduzione di una rilevazione CAPI, come numero, giorno ed esito delle visite, oppure, in un'indagine CAWI, ai dati riguardanti la tempistica dei solleciti postali o via e-mail. Accanto a queste informazioni, vanno considerate anche tutte quelle presenti nei registri statistici disponibili in Istituto, quali ad esempio le informazioni demografiche, strutturali, sociali e territoriali, che rappresentano variabili note a monte del processo di raccolta dati e disponibili per tutta la popolazione di riferimento dell'indagine.

Se l'applicabilità dei modelli di analisi dipende dalla disponibilità di paradati e variabili ausiliarie, la loro usabilità dipende dal modo in cui vengono presentati i risultati. È, quindi, di cruciale importanza costruire strumenti di monitoraggio che permettano ai supervisori di indagine di interpretare in modo semplice e diretto i risultati dell'analisi, così da individuare facilmente quali modifiche apportare alla raccolta dati e quale sia l'impatto delle stesse sull'andamento della rilevazione.

Il grafico riportato in Figura 1 è una dimostrazione della semplicità auspicata, come potrebbero esserlo altri strumenti e indicatori di monitoraggio, come, ad esempio, le carte di controllo e/o gli R-*indicators* (RISQ, 2009), che permettono di supervisionare più aspetti del processo di raccolta dati e, di conseguenza, consentono di intervenire in modo rapido sul campo con azioni mirate e specifiche per il problema incontrato. Ad esempio, le carte di controllo potrebbero essere usate per monitorare il comportamento del rilevatore CATI durante la fase di contatto per capire se viene registrato un numero di rifiuti troppo alto rispetto alla media (Murgia e Simeoni, 2005), posizionandosi costantemente e non in modo sporadico al di sopra del limite superiore della carta di controllo. In tal caso, si potrebbe pensare di intervenire con un *debriefing* sul come motivare i rispondenti alla partecipazione. Gli R-*indicators*, a loro volta, potrebbero essere usati per individuare i segmenti di popolazione sottorappresentati in modo da agire su di essi con solleciti mirati o assegnandoli a rilevatori più esperti.

L'uso di modelli di analisi della non-risposta e di strumenti di monitoraggio deve essere inquadrato all'interno di una strategia per la qualità dell'intero processo di indagine e non limitata alla sola fase di conduzione della raccolta dati. Tale strategia dovrà essere finalizzata a: i) definire obiettivi di qualità misurabili, ii) scegliere i modelli di analisi più adatti a individuare, per poi correggere, le cause distorsive della non-risposta; iii) rendere disponibili i paradati e le variabili ausiliarie di interesse; iv) costruire strumenti di monitoraggio semplici e applicabili a tutte le indagini, a prescindere dalla popolazione e dalla tecnica di raccolta dati usata.

La strategia dovrà essere articolata nei seguenti passi:

1) Scelta del modello e degli strumenti. Per individuare i modelli e gli strumenti di monitoraggio più adatti agli scopi prefissati, è necessario partire dai dati di indagini già concluse e per le quali si hanno informazioni sulle azioni messe in campo durante la fase di raccolta dati. Da questa prima fase della strategia si potrà non solo valutare l'efficacia e l'applicabilità dei metodi, ma anche stabilire se le azioni di *field* messe effettivamente in atto abbiano contribuito a migliorare la rappresentatività della risposta. Per questo ultimo aspetto ci si potrebbe basare sulla definizione di scenari derivanti da modelli che usano dati simulati.

2) Sperimentazione dei modelli e degli strumenti. Una volta individuati i metodi e gli strumenti, questi potrebbero essere usati in via sperimentale su indagini che devono andare in *field* e per le quali si dispone di variabili ausiliarie. La sperimentazione potrebbe essere fatta in corso di indagine oppure durante un'indagine pilota. Nel primo caso, dal campione di indagine potrebbe essere estratto un sotto-campione casuale di unità che a sua volta potrebbe essere diviso in due gruppi: uno di test e uno di controllo (RISQ, 2009). Al sottogruppo di test si applicano i nuovi strumenti di monitoraggio e le azioni correttive da questi suggerite, mentre sul sottogruppo di controllo si opera come di consueto. Al termine della rilevazione si potrà così osservare se il campione di test ha ottenuto risultati migliori in termini di bilanciamento della mancata risposta.

3) Definizione, raccolta e memorizzazione di paradati e variabili ausiliarie. A monte delle sperimentazioni, e in modo sistematico per tutte le indagini, dovrà essere fatto uno studio per capire quali informazioni ausiliarie e paradati sono già disponibili e quali sono quelle da aggiungere tramite collegamento a registri e/o progettando la raccolta di paradati attraverso i questionari elettronici, per le indagini che fanno uso di tecniche assiste da computer, e mediante l'osservazione del rilevatore per le indagini dirette. In caso di rilevazione affidata a ditte esterne, sarà importante prevedere la raccolta, memorizzazione e invio di queste informazioni in modo sistematico e tempestivo durante la fase di raccolta dati. A latere si potrebbe perseguire un obiettivo di medio-lungo periodo che è quello di semplificare la quantità di reportistica di monitoraggio sviluppata dalle ditte

esterne che genera sovente un carico di lavoro eccessivo per entrambi le parti. Per le indagini il cui questionario elettronico è sviluppato in Istituto, dovranno essere date indicazioni sul tipo di paradati da memorizzare.

4) Centralizzazione dei sistemi di monitoraggio. Dovranno essere potenziate e sfruttate al massimo le funzionalità offerte dall'esistenza di un sistema centralizzato per la gestione della fase di raccolta dati in modo da standardizzare la rilevazione delle informazioni per tipologia di indagine e facilitare così l'implementazione e l'utilizzo dei modelli di analisi della non-risposta.

A monte della strategia sopra decritta è fondamentale tenere conto del disegno del questionario di indagine che deve essere tale da contenere l'errore di misura dovuto alla tecnica (*mode measurement effect*). Ciò è vero soprattutto nelle indagini *mixed-mode*, per le quali questo errore può verificarsi quando uno stesso rispondente risponde diversamente ad uno stesso quesito al variare della tecnica di rilevazione utilizzata, e inoltre perché non tutte le tecniche sono 'adatte' a somministrare qualsiasi tipologia di domanda (Istat, 2018). In questi casi è importante che la fase di raccolta dati sia preceduta da una fase di test del questionario mirata a scegliere il disegno che limiti l'impatto distorsivo del *mode measurement effect*, nel caso valutando anche se la misurazione di una o più variabili target possa essere influenzata dalla tecnica, indipendentemente dalla diversa propensione ad essa di particolari sottoinsiemi di rispondenti.

**Riferimenti bibliografici**

COBBEN F. 2009. Nonresponse in household surveys. Methods for analysis and adjustment. PhD thesis. University of Amsterdam, Statistics Netherlands, Amsterdam.

BETHLEHEM J.G., COBBEN F., SCHOUTEN B. 2011. *Handbook of nonresponse in household surveys*. Wiley Handbooks in Survey Methodology.

ISTAT (2018). L'effetto tecnica nelle indagini mixed-mode – Aspetti teorici e sperimentazioni su indagini sociali che utilizzano il web. Collana Letture statistiche: Metodi. Disponibile al link: https://www4.istat.it/it/archivio/211135.

KREUTER F. (Ed.). 2013. *Improving surveys with paradata: Analytic uses of process information* (Vol. 581). Hoboken: John Wiley & Sons.

LITTLE R.J.A., RUBIN D.B. 2002. *Statistical analysis with missing data.* Wiley series in Probability and Statistics. New York: John Wiley & Sons, Inc., NY, USA.

LYBERG L.E., STUKEL D.M. 2017. The roots and evolution of the total survey error concept. In BIEMER P.P et al. (Eds.), *Total Survey Error in Practice*, John Wiley & Sons, pp. 1-22.

MURGIA M., SIMEONI G. 2005. Improving the quality of assisted coding of occupation in CATI surveys through control charts. CLADAG - Classification and Data Analysis Group 2005, 6-8 giugno, Parma.

RISQ – Representative Indicators for Survey Quality (2009). Disponibile al link: https://www.cmi.manchester.ac.uk/research/projects/representative-indicators-for-survey-quality/publications/.

## SUMMARY

### Quality in data capturing processes: the case of the survey on the aspects of daily life

The quality of the statistical information depends on several factors such as definitions and concepts adopted, sampling design, data capturing techniques, tools and methods for the analysis and representations of data. The Total Survey Error (TSE) approach represents a theoretical framework for the evaluation of the quality of survey data that takes into account the different types of error that might arise in each step of the survey process, from the design phase to the analysis of data, in order to compare costs and benefits. In this framework, the survey error is defined as the deviation of a survey response from its underlying true value. For every phase of the survey it is necessary to evaluate the efficiency of the actions taken. This work focuses on the assessment of the data collection phase through indicators based on data that are collected during the process itself, the so-called paradata.

The analysis of paradata and their association with other information -auxiliary data as well as survey data- is useful in constructing performance indicators of the data collection phases to guide the necessary adjustment process.

This paper reports an example of the approach followed, based on the data of the Istat survey ‘Aspects of daily life, 2019’. Specifically, it describes the analysis of the propensity to respond and the data collection techniques used (CAWI vs CAPI-PAP) to detect those respondents profiles that need an adaptation of either contact or reminder strategies to improve their propensity to respond. In this example the analysis is conducted using a logistic model. The possible methods and tools for improving the quality of the survey are drew as part of a strategy involving the different actors of the data collection process.

Claudio CECCARELLI, Istat, clceccar@istat.it
Marco FORTINI, Istat, fortini@istat.it
Manuela MURGIA, Istat, murgia@istat.it
Alessandra NUCCITELLI, Istat, nuccitel@istat.it
Rita RANALDI, Istat, ranaldi@istat.it
Francesca ROSSETTI, Istat, frrosset@istat.it

# FILTERED CLUSTERING FOR EXCHANGE TRADED FUND

Gloria Polinesi, Maria Cristina Recchioni

## 1. Introduction

Time series are one of many instruments used to represent data present in a variety of fields, from brain activity to finance. Researchers apply clustering techniques to time series data for many reasons. Zhang et al. (2011) mention three main objectives when detecting similarities between time series: time, shape, and change. Similarity in time means that time series are grouped together when they move similarly in time; similarity in shape occurs when time series share common trends or sub patterns. Finally, similarity in change means that time series show similarity in fitted parameters referring to underlying models, which may be different.

Due to the nature of this type of data, cluster analysis of time series requires particular techniques. Mantegna (1999) and other authors use a raw data approach for stock return time series. A single observation (day, week or month) of the time series represents a characteristic of the element and stocks are grouped together when they are correlated: the Pearson correlation coefficient quantifies the degree of interdependence between pairs of financial assets.

Clustering algorithms allow leading information about structural organization aspects to be extracted from a correlation matrix of return time series whereas correlation matrices can be represented as complete graphs lacking a notion of hierarchy (De Prado, 2016). Clustering tools, spectral methods (random matrix theory), and correlation-based graphs are all algorithms used to extract information from complex systems of correlation matrices. Indeed, correlation matrices are subject to non-stationary market conditions and 'measurement noise' due to the finite length of the time series, which makes the analysis difficult without applying these filtering tools.

We contribute to the existing literature applying the work of Miceli and Susinno (2004) outlined above to obtain a cluster of Exchange Traded Fund (ETF) returns that reflects the classification per investment class provided by the Italian Stock Exchange (commonly known as the Borsa Italiana). In the following sections, we describe hierarchical algorithms and alternative methods to filter correlation matrices and we conclude with some results.

## 2. Hierarchical clustering algorithm

This section shows the use of a hierarchical clustering algorithm to filter correlation matrices of stock return time series to reduce the number of parameters. In fact, filtering procedure permits to extract the structure hidden in large correlation matrices keeping significant links and removing the noisy ones. The analysis considers both a static financial market and a complex system that evolves over time.

In his first work in 1999, Mantegna investigated the correlation matrix to detect the hierarchical organization of stocks in a financial market. In an ultrametric space, the minimal spanning tree (MST) between stocks reveals the topological layout of the financial market that holds important meaning from the economic point of view. A MST, the minimum structure in terms of sum of distances between nodes, groups stocks with respect to the economic sector of the underlying companies. As a consequence, the MST is the tree associated with the single linkage clustering algorithm so playing the same role as a dendrogram.

Tumminello *et al.* (2010) also confirm that elements (or nodes) share information according to the communities they belong to and that communities are organized in a nested structure. Hierarchical clustering algorithms enable this complex structure to be detected. Furthermore, Spelta and Araújo (2012) describe the minimal spanning tree as the corresponding representation of a fully-connected system (network) where sparseness replaces completeness in a suitable way.

The steps necessary to draw an MST can be summarized as follows.

We start with the correlation matrix of the time series of N stock returns, computed as the difference of the logarithm of stock prices in the time horizon T[1]

$$r_i(t) = logP_i(t) - logP_i(t-1) \tag{1}$$

The elements of the correlation matrix for each pair of stocks,

$$c_{ij} = \frac{E(r_i r_j) - E(r_i)E(r_j)}{\sigma_i \sigma_j} \tag{2}$$

are converted into distance elements:

[1] The prices and returns of stocks, the financial assets in general, can be daily, weekly, monthly, or yearly.

$$d_{ij} = \sqrt{2 - 2c_{ij}} \tag{3}$$

MSTs are based on the distance matrix computed thus. This tree graph allows the number of links connecting stocks to be reduced from (N(N-1))/2 (total number of parameters in the distance matrix) to N-1. In general, minimal spanning trees allow hierarchical organization to be detected in sectors and subsectors of stocks, but the literature shows that the result changes if the frequency of data changes. For more details about algorithms to derive an MST, see Moret and Shapiro (1991).

MSTs are associated with the dendrogram of the single linkage clustering algorithm (SLCA); however, MSTs retain some information that the single linkage dendrogram disregards (Raffinot, 2017).

This author tests some Single Linkage (SL) variants: complete linkage (CL), average linkage (AL) and Ward's method (WM), associated with different dendrograms or hierarchical trees. We recall that:

- at each step, SL combines two clusters that contain the closest (minimum distance) pair of elements
- CL works opposite to SL: At each step, it combines two clusters that hold the farthest (maximum distance) pair of elements
- AL considers the distance between two clusters as the average distance between pairs of elements belonging to those clusters
- at each stage, WM merges two clusters if they provide the smallest increase in squared error.

Other authors concentrate on MSTs as characteristic tree graphs to describe the correlation matrices. For example, Onnela *et al.* (2003b) emphasize the aspects already presented by previous authors, but also criticize the fact that the minimal spanning tree (or simply "asset tree") only represents the static average of an evolving complex system. These authors explore the dynamics of the asset tree by computing the correlation matrix for each rolling window of width T and draw the MST for each period to see how the structure of the minimal spanning tree changes over time. They demonstrate that the basic structure of MSTs is very robust with respect to time, but it shrinks during market crises due to the strong global correlation, which makes the behavior of the assets very homogeneous.

Spelta and Araújo (2012) also propose a measure called "residuality coefficient" that compares the relative strengths of the connections above and below a threshold distance in the tree in order to assess structural changes in the MST over time.

Matesanz and Ortega (2015) draw an MST for each time window to evaluate temporal changes in the time series of countries' debt-to-GDP ratio. They calculate the agglomerative coefficient (Kaufman and Rousseeuw, 2009) of each temporal tree and cophenetic correlation (Sokal and Rohlf, 1962) between hierarchical trees for different times. An agglomerative coefficient close to 1 implies a highly nested tree structure and the cophenetic correlation instead gives an idea of how similar the grouping structure is between two different hierarchical trees. For example, during the market crises beginning in 2008, the value of the agglomerative coefficient was much less than 1 and hierarchical trees for overlapping windows were not correlated.

Although the work of Matesanz and Ortega (2015) refers to different time periods and type of data considered, the results confirm that the structure of hierarchical trees tends to be flat and different from others during market crises.

A critical aspect in considering dynamic MSTs is represented by the fact that the choice of time windows (number and length) is arbitrary, as asserted by Marti *et al.*, 2017. In fact, a trade off exists between data that is too noisy and too smoothed for small and large window widths, respectively (see Onnela *et al.*, 2003a for details).

### *2.1 Extensions of the MST: different algorithms*

Following the work of Marti *et al.* (2017), this section lists the different algorithms used to replace the minimal spanning tree and its corresponding clusters with the goal of improving upon the seminal work of Mantegna (1999). These algorithms are both hierarchical, with correlated graphs, and non-hierarchical. The latter consider a spectral method based on the study of eigenvalues of correlation matrices.

With respect to hierarchical algorithms, Tumminello *et al.* (2005) introduce a graph to filter correlation matrices that preserves the hierarchical organization of the minimal spanning tree but includes more information. This graph is known as the planar maximally filtered graph (PMFG); it represents an extension of the MST. The basic difference between the two is the number of links considered: the MST contains $N-1$ links, compared to $3(N-2)$ for the PMFG, where N is the number of nodes in the graph[2].

Therefore, PMFG holds the hierarchical skeleton of the minimal spanning tree but is enriched with loops and cliques. As explained in Tumminello *et al.* (2010), a clique of $k$ elements is a complete subgraph that links all $k$ elements. Due to

[2] For planar filtered graphs, the genus is equal to 0. According to the definition in Tumminello et al. (2005), the genus is a topologically invariant property of a surface defined as the largest number of non-isotopic simple closed curves that can be drawn on the surface without separating it, i.e., the number of handles in the surface.

Kuratowski's theorem, PMFGs can only have cliques of 3 or 4 elements. The number of 3-cliques and 4-cliques that can be built is $3(N-8)$ and $N-3$ respectively.

Tumminello *et al.* (2007b) introduce the average linkage minimum spanning tree (ALMST), i.e., the spanning tree associated with the average linkage clustering algorithms (ALCA). These authors show that ALMST is able to detect groups defined in terms of economic sectors and sub-sectors of stock return slightly better than the MST.

Musmeci *et al*. (2015) recently introduced the directed bubble hierarchical tree (DBHT), a novel clustering algorithm based on the topological structure of the PMFG. In contrast to other hierarchical techniques, the DBHT first identifies clusters and then sets the intra- and inter-group hierarchy.

From the non-hierarchical side, random matrix theory (RMT) is the main approach used to investigate the structure of return correlation matrices of financial assets.

Random matrix theory has a long history (Mehta, 2004). The first results in the financial sector can be found in Galluccio *et al.* (1998), Laloux *et al. (*1999) and Plerou *et al.* (1999), Plerou *et al.* (2002).

The basic idea of RMT is to compare ordered eigenvalues, $\lambda_k < \lambda_{k+1}$, of the correlation matrix of returns (Eq. 1) to eigenvalues of a random Wishart matrix $R = \frac{1}{T}AA^T$ of the same size. This is done to understand how different the matrix in question is from a random matrix. Here, $A$ is an $N \times T$ matrix containing N time series of length T whose elements are independent, identically distributed random variables with zero mean and variance $\sigma^2 = 1$.

The random correlation matrix of this set of variables as $T \to \infty$ is the identity matrix; when T is finite, the correlation matrix is generally different from the identity matrix.

RMT proves that in the limit $N \to \infty$ and $T \to \infty$ with a fixed ratio $Q = \frac{T}{N} \geq 1$ and $\sigma^2 = 1$, the eigenvalue spectral density is given by:

$$f(\lambda) = \frac{T}{2\pi}\frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \tag{4}$$

where $\lambda_\pm = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}$ represent the minimum and the maximum eigenvalue of the Wishart matrix.

It can be shown that the eigenvalues deviating from those of a random matrix convey meaningful information stored in the correlation matrix. Indeed, information can be extracted from eigenvalues that are higher than $\lambda_+$ (deviating eigenvalues), which involves correlations between stocks that belong to the same industry or geographical area. The "bulk" of eigenvalues instead agree with RMT, revealing the random correlations.

Onnela *et al.* (2004) remark that random matrix theory offers an interesting comparative perspective with respect to hierarchical clustering techniques.

Other authors use RMT to filter correlation matrices and construct MSTs on this filtered matrix because, in order to extract the structure hidden in large correlation matrices, trees are easier to interpret than inspecting large matrices. With this procedure, Miceli and Susinno (2004) obtain a clusterization per strategies of hedge fund returns. Hedge fund strategies represent the investment styles stated by fund managers (Lhabitant and Learned, 2002). Conlon *et al.* (2007) have also confirmed this result.

## 3. Data and results

We consider a data set composed of 85 ETF return time series traded over the period December 2016-November 2017 (for $NT$ total observations).

According to the classification per investment class provided by the Borsa Italiana, Table 1 shows the ETFs classified into 11 asset classes and the number of ETFs belonging to the class. Summary statistics of returns for the asset classes - mean, variance, kurtosis and skewness- are described in the same Table. The mean value is around 0 for each asset class. The standard deviation instead depends on the asset class considered: emerging equity ETFs are slightly more volatile with respect to other classes considered. The distribution of most ETF returns tends to be non-Gaussian as confirmed by high values of kurtosis and negative values of skewness.

Having filtered the correlation matrix using the RMT approach, we then reconstruct the distance matrix from the filtered correlation matrix. Figure 1 shows the MST extracted from the distance matrix, where the size of the vertex represents the node degree (i.e., the number of edges connected to each node) and the color represents the class it belongs to.

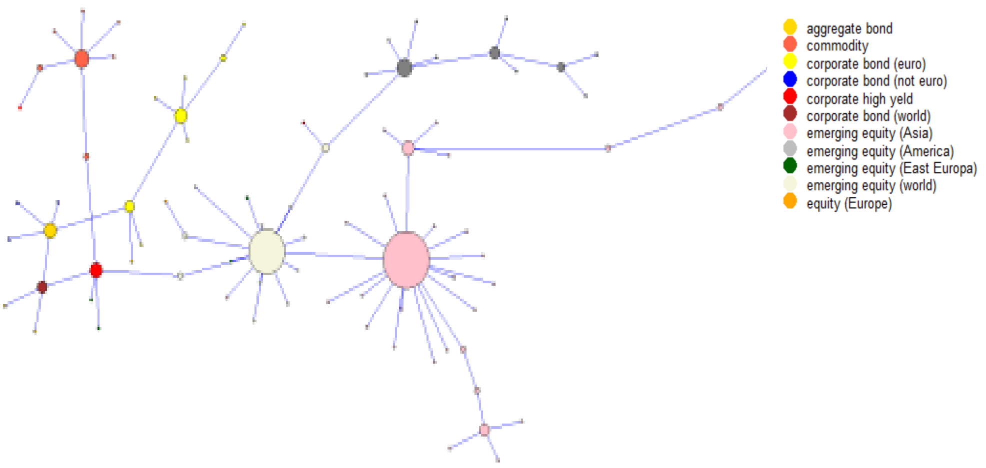
Note that the topological structure of the ETFs in Figure 1 reflects the classification per investment class described in Table 1, where the three main groups are represented by the Equity (emerging and Europe), the Corporate (aggregate, bond and high yield) and Commodity classes, respectively. Indeed, clusters obtained in the MST represent a specific class of ETFs according to the classification per investment classes of Borsa Italiana. Within these groups, specific ETFs act like hubs with higher values of node degree: the MST reveals the importance of the Asian and World Emerging Market classes, which have the highest centralities.

**Table 1 –** *Summary statistics of ETF returns divided into specific classes.*

| ETF Class | # of ETFs per Class | Mean | St. Dev. | Kurtosis (excess) | Skewness |
|---|---|---|---|---|---|
| Aggregate Bond | 4 | 0.0002 | 0.0025 | 5.5703 | -0.4994 |
| Commodity | 8 | 0.0001 | 0.0061 | 0.9043 | -0.0618 |
| Corporate-euro | 8 | 0.0001 | 0.0012 | 0.9405 | -0.4330 |
| Corporate- not euro | 3 | 0.0003 | 0.0030 | 1.4869 | 0.0491 |
| Corporate-high yield | 2 | 0.0004 | 0.0014 | 6.6871 | -0.4438 |
| Corporate-world | 1 | 0.0003 | 0.0024 | 1.3115 | -0.2290 |
| Emerging Equity-Asia | 31 | 0.0010 | 0.0070 | 3.0669 | -0.1522 |
| Emerging Equity-America | 10 | 0.0009 | 0.0129 | 33.3649 | -2.9396 |
| Emerging Equity-East Europe | 4 | 0.0006 | 0.0133 | 4.2257 | 0.2659 |
| merging Equity-world | 13 | 0.0011 | 0.0060 | 1.7502 | -0.1888 |
| Equity-Europe | 1 | 0.0005 | 0.0052 | 0.7576 | 0.1830 |

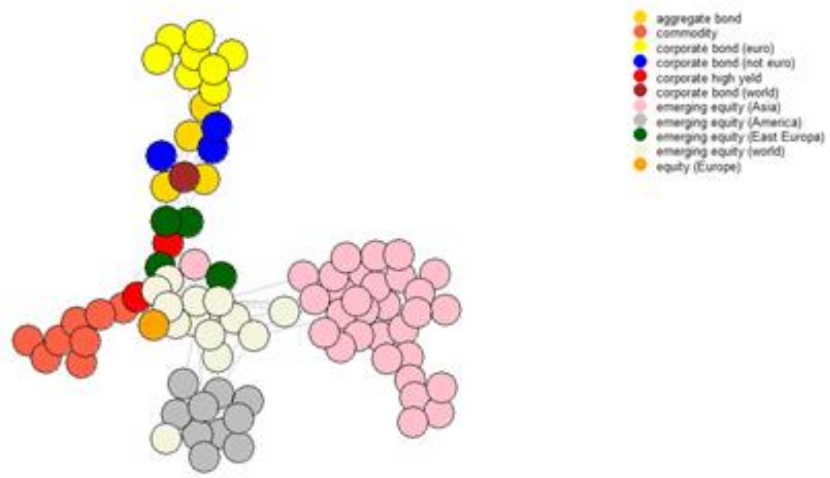
As a robustness exercise of our result, we use the same data to compare the minimum spanning tree with the planar maximally filtered graph (PMFG). Figure 2 shows that correlations of ETFs in the MST are also present in the PMFG, where a classification into investment classes is more evident from the structure of the network**.**

**Figure 1 –** *Minimum spanning tree drawn from the filtered distance matrix.*




*Source: Own elaboration on ETF data.*

**Figure 2 –** *Planar maximally filtered graph drawn from the filtered distance matrix.*




*Source: Own elaboration on ETF data.*

## 4. Conclusion

We have presented hierarchical clustering and spectral methods in order to highlight stronger correlations between time series of financial asset returns. These methods allow information in complex datasets to be filtered by building sparse networks or trees but retaining the relevant edges.

In fact, applying the random matrix approach to the correlation matrix of ETF returns and then drawing a minimal spanning tree as in the work of Miceli and Susinno (2004), allows to obtain clusters of ETFs representing the classification into investment class provided by the Italian Stock Exchange.

We have demonstrated that using RMT to filter a correlation matrix allows trees to be constructed that are easier to interpret with respect to the inspection of large matrices.

## References

CONLON T., RUSKIN H. J., CRANE M. 2007. Random Matrix Theory and Fund of Funds Portfolio Optimisation, *Physica A: Statistical Mechanics and Its Applications,* Vol.382, No.2, pp.565-76.

DE PRADO M. L. 2016. Building Diversified Portfolios that Outperform Out of Sample, *Journal of Portfolio Management*, Vol. 42, No.4, pp. 59-69.

GALLUCCIO S., BOUCHAUD J. P., POTTERS M. 1998. Rational Decisions, Random Matrices and Spin Glasses, *Physica A: Statistical Mechanics and Its Applications*, Vol. 259, No. 3, pp. 449-56.

KAUFMAN L., ROUSSEEUW P. J. 2009. *Finding groups in data: an introduction to cluster analysis*. Hoboken: John Wiley & Sons.

LALOUX L., CIZEAU P., BOUCHAUD J.P., POTTERS M. 1999. Noise Dressing of Financial Correlation Matrices, *Physical Review Letters*, Vol. 83, No.7, pp. 1467.

LHABITANT F. S., LEARNED M. 2002. Hedge Fund Diversification: How Much Is Enough?, *The Journal of Alternative Investments*, Vol. 5, No. 3, pp. 23-49.

MANTEGNA R. N. 1999. Hierarchical Structure in Financial Markets, *The European Physical Journal B-Condensed Matter and Complex Systems*, Vol. 11, No.1, pp. 193-97.

MARTI G., NIELSEN F., BINKOWSKI M., DONNAT P. 2017. A Review of Two Decades of Correlations, Hierarchies, Networks and Clustering in Financial Markets., *arXiv Preprint arXiv:1703.00485*.

MATESANZ D., ORTEGA G. J. 2015. Sovereign Public Debt Crisis in Europe. A Network Analysis, *Physica A: Statistical Mechanics and Its Applications*, Vol. 436, pp. 756-66.

MEHTA M. L. 2004. *Random matrices*. Amsterdam: Elsevier.

MICELI M. A., SUSINNO G. 2004. Ultrametricity in Fund of Funds Diversification, *Physica A: Statistical Mechanics and Its Applications*, Vol. 344, No.1-2, pp. 95-99.

MUSMECI N., ASTE T., DI MATTEO T. 2015. Relation Between Financial Market Structure and the Real Economy: Comparison Between Clustering Methods, *PloS One*, Vol. 10, No.3, pp. e0116201.

MORET B. M., SHAPIRO H. D. 1991. An empirical analysis of algorithms for constructing a minimum spanning tree. In *Workshop on Algorithms and Data Structures*, Vol. 519, Berlin: Springer, p. 400-411.

ONNELA J.P., CHAKRABORTI A., KASKI K., KERTESZ J. 2003a. Dynamic Asset Trees and Black Monday, *Physica A: Statistical Mechanics and Its Applications*, Vol. 324, No.1-2, pp. 247-52.

ONNELA J.P., CHAKRABORTI A., KASKI K., KERTESZ J., KANTO A. 2003b. Dynamics of Market Correlations: Taxonomy and Portfolio Analysis, *Physical Review E, Vol.* 68, No.5, pp. 056110.

ONNELA J.P, KASKI K., KERTÉSZ J. 2004. Clustering and Information in Correlation Based Financial Networks, *The European Physical Journal B*, Vol. 38, No.2, pp. 353-62.

PLEROU V., GOPIKRISHNAN P., ROSENOW B., AMARAL L.A.N., GUHR T., STANLEY H.E. 2002. Random Matrix Approach to Cross Correlations in Financial Data, *Physical Review E*, Vol. 65, No.6, pp. 066126.

PLEROU, V., GOPIKRISHNAN P., ROSENOW B., AMARAL L.A.N., GUHR T., STANLEY H.E. 1999. Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series, *Physical Review Letters*, Vol. 83, No.7, pp. 1471.

RAFFINOT T. 2017. Hierarchical Clustering-Based Asset Allocation, *The Journal of Portfolio Management*, Vol. 44, No.2, pp. 89-99.

SOKAL R. R., ROHLF F. J. 1962. The Comparison of Dendrograms by Objective Methods, *Taxon*, Vol. 11, No.2, pp. 33-40.

SPELTA A., ARAÚJO T. 2012. The Topology of Cross-Border Exposures: Beyond the Minimal Spanning Tree Approach, *Physica A: Statistical Mechanics and Its Applications*, Vol. 391, No. 22, pp. 5572-5583.

TUMMINELLO M., ASTE T., DI MATTEO T., MANTEGNA R. N. 2005. A Tool for Filtering Information in Complex Systems. In *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102, No.30, pp. 10421-10426.

TUMMINELLO M., CORONELLO C., LILLO F., MICCICHÈ S., MANTEGNA R.N. 2007b. Spanning trees and bootstrap reliability estimation in correlation-based networks., *International Journal of Bifurcation and Chaos*, Vol. 17, No. 7, pp. 2319-2329.

TUMMINELLO M., LILLO F., MANTEGNA R. N. 2010. Correlation, Hierarchies, and Networks in Financial Markets, *Journal of Economic Behavior & Organization*, Vol. 75, No.1, pp. 40-5

ZHANG X., LIU J., DU Y., LV T. 2011. A Novel Clustering Method on Time Series Data, *Expert Systems with Applications*, Vol. 38, No.9, pp. 11891-11900.

**SUMMARY**

**Filtered Clustering for Exchange Traded Fund**

In this work, we show how time series of Exchange Traded Funds (i.e., ETF) returns can be clustered by reflecting the classification per investment class provided by the Borsa Italiana. We use the random matrix theory (RMT) filter to "clean" noise from a correlation matrix and we then use the reconstructed filtered correlation matrix to draw the hierarchical tree associated with the single linkage clustering algorithm (minimum spanning tree). The main goal of the paper is to show that RMT as a filter for correlation matrices enables the construction of trees that are easier to interpret with respect to large matrices, even for ETF returns.

Gloria POLINESI, Università Politecnica delle Marche, g.polinesi@univpm.it
Maria Cristina RECCHIONI, Università Politecnica delle Marche, m.c.recchioni@univpm.it

**SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS**
**ISSN ASSEGNATO: 0035-6832**

*Direttore Responsabile:* Prof.ssa CHIARA GIGLIARANO

Iscrizione della Rivista al Tribunale di Roma del 5 dicembre 1950 N. 1864

Associazione all'Unione Stampa Periodica Italiana

TRIMESTRALE

*La copertina è stata ideata e realizzata da Pardini, Apostoli, Maggi* p.a.m.@tin.it – *Roma*

Stampato da CLEUP sc
"Coop. Libraria Editrice Università di Padova"
Via G. Belzoni, 118/3 – Padova (Tel. 049/650261)
www.cleup.it

# ATTIVITÀ DELLA SOCIETÀ

## A) RIUNIONI SCIENTIFICHE

| | |
|---|---|
| XXXVII | La mobilità dei fattori produttivi nell'area del Mediterraneo (Palermo, 15-17 giugno 2000). |
| XXXVIII | Qualità dell'informazione statistica e strategie di programmazione a livello locale (Arcavacata di Rende, 10-12 maggio 2001). |
| XXXIX | L'Europa in trasformazione (Siena, 20-22 maggio 2002). |
| XL | Implicazioni demografiche, economiche e sociali dello sviluppo sostenibile (Bari, 15-17 maggio 2003). |
| XLI | Sviluppo economico e sociale e ulteriori ampliamenti dell'Unione Europea (Torino, 20-22 maggio 2004). |
| XLII | Sistemi urbani e riorganizzazione del territorio (Lucca, 19-21 maggio 2005). |
| XLIII | Mobilità delle risorse nel bacino del Mediterraneo e globalizzazione (Palermo, 25-27 maggio 2006). |
| XLIV | Impresa, lavoro e territorio nel quadro dei processi di localizzazione e trasformazione economica (Teramo 24-26 maggio 2007). |
| XLV | Geopolitica del Mediterraneo (Bari, 29-31 maggio 2008). |
| XLVI | Povertà ed esclusione sociale (Firenze 28-30 maggio 2009). |
| XLVII | Un mondo in movimento: approccio multidisciplinare ai fenomeni migratori (Milano 27-29 maggio 2010). |
| XLVIII | 150 anni di Statistica per lo sviluppo del territorio: 1861-2011. (Roma 26-28 maggio 2011). |
| XLIX | Mobilità e sviluppo: il ruolo del turismo. (San Benedetto del Tronto, 24-26 maggio 2012). |
| L | Trasformazioni economiche e sociali agli inizi del terzo millennio: analisi e prospettive (Università Europea di Roma, 29-31 maggio 2013). |
| LI | Popolazione, sviluppo e ambiente: il caso del Mediterraneo (Università Federico II di Napoli, 29-31 maggio 2014). |
| LII | Le dinamiche economiche e sociali in tempo di crisi (Università Politecnica delle Marche, 28-30 maggio 2015). |
| LIII | Mutamento economico e tendenze socio-demografiche tra sfide e opportunità (Università degli Studi Internazionali di Roma, 26-28 maggio 2016). |
| LIV | Mobilità territoriale, sociale ed economica: modelli e metodi di analisi (Università degli Studi Internazionali di Catania, 25-26 maggio 2017). |
| LV | Coesione sociale, welfare e sviluppo equo e sostenibile (Università degli Studi dell'Insubria, Varese 24-25 maggio 2018). |
| LVI | Benessere e Territorio: Metodi e Strategie (Università Politecnica delle Marche, Ascoli Piceno 23-24 maggio 2019). |