VOLUME LXXIX – N. 2 APRILE – GIUGNO 2025

RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA

Volume stampato grazie al contributo liberale della Banca d'Italia (anno 2024)

SIEDS
SOCIETÀ ITALIANA
DI ECONOMIA DEMOGRAFIA E STATISTICA

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

info@sieds.it

rivista@sieds.it

VOLUME FUORI COMMERCIO

INDICE

EQUIVALENCE SCALES FOR THE DISTRIBUTION OF CONSUMPTION EXPENDITURES IN ITALY: PROPOSED METHODS TO UPDATE THE CARBONARO SCALE

Andrea Cutillo, Alessandro Romeo, Livia Celardo

Abstract. Comparing the well-being of different households requires knowledge not only of the resources available to them, but also of their needs. This seems to be possible, albeit with some effort and approximation, by defining an equivalence coefficient that indicates, using a reference household, how much a family with different characteristics needs to enjoy the same standard of living. Equivalence scales are typically used to make the expenditure of households of different size and composition comparable. In this paper, using data from the 2017-2019 Italian Household Budget Survey (IHBS), we propose an equivalence scale calculated using a simple method based on recent improvements in absolute poverty methodology. To validate our analysis, we compared the results with the original Carbonaro scale calculated in 1985, a version of the Carbonaro scale updated using the 2017-2019 HBS data, and scales calculated using complete household demand systems. Finally, we present a simple and intuitive method to assess which of these scales performs better in the Italian economic context. Preliminary results suggest that the Carbonaro scale, as originally constructed, is now outdated and needs to be revised. On the contrary, the proposed approach of using absolute poverty thresholds as a proxy for essential expenditure seems to provide encouraging results, especially in the light of the evolution of the Italian socio-economic context and the relationship that must exist between absolute and relative poverty.

1. Introduction

Comparing household welfare necessitates knowledge of their resources and needs. This is approximated by defining an equivalence coefficient indicating how much a family with different characteristics requires to enjoying the same living standard as a reference family. Equivalence scales offer parameters for comparing expenditure or income levels, accounting for economies of scale deriving from sharing expenditures. They are a key tool in welfare analysis, addressing income and consumption distributions, inequality and poverty (Buhmann et al., 1988; Deaton and Zaidi, 2002).

Deaton and Zaidi (2002) identify three main approaches for calculating equivalence scales: the subjective approach, which relies on household surveys; the normative approach, where institutions set scales based on objective assumptions; and the behavioural or utility-based approach, which analyses household consumption expenditure patterns. The subjective approach is generally seen as

unsuitable for welfare analysis due to difficulties in designing effective surveys. The normative approach, such as the OECD modified scale, facilitates international comparisons but may not fully adapt to different socio-economic contexts. The behavioural approach appears more appropriate, expecially in country-specific studies, though it relies on assumptions that can be hard to verify and presents econometric challenges (Blundell and Lewbel 1991; Dudel et al., 2021). Importantly, inequality measures are highly sensitive to the equivalence scale used (Ferreira and Ravallion, 2011).

In Italy, the National Institute of Statistics (Istat) uses two equivalence scales: the modified OECD scale for calculating equivalent incomes and the At Risk of Poverty indicator, and the Carbonaro scale (Carbonaro, 1985), used for equivalent consumption expenditure and the Relative Poverty indicator. The Carbonaro scale, based on Engel's behavioural approach, adjusts expenditures according to household size. After the 2022 revision of Italy's absolute poverty methodology, Istat is considering updating its relative poverty methodology, as the Carbonaro scale may no longer reflect the current socio-economic environment.

Preliminary findings suggest that the Carbonaro scale's validity is declining. Even when updated with current data, the method - using food expenses as a proxy for essential needs - seems increasingly unsuitable in modern Italy. Methods based on complete demand systems have also shown unsatisfactory results, raising questions about whether well-being should be measured using only essential expenses or all types of expenditures. These methods also fail to clarify the best point in the expenditure distribution for calculating scale coefficients. In contrast, a modified Engelian approach, which builds on but diverges from the Carbonaro model, shows promising results. This approach uses absolute poverty thresholds as proxies for essential expenses, as they represent exactly the minimum expenditure required to avoid severe social exclusion in Italy today.

The paper proceeds as follows: Section 2 defines equivalence scales and their plausibility criteria. Section 3 outlines the methods compared. Section 4 presents results and discusses scale plausibility. Section 5 introduces a simple empirical method for selecting the most suitable scale. Section 6 concludes.

2. General definition of equivalence scale and plausibility of the scales

Let $\mathbf{z} = (z_1, \ldots, z_k)$ denote k household characteristics. All households can choose between m goods with prices captured in $\mathbf{p} = (p_1, \ldots, p_m)$. Household demand is then given by the demand function $D(p, y, z) = \mathbf{q} = (q_1, \ldots, q_m)$, where q_i is the demand for good i and y is household income. Household utility is given by $U(\mathbf{q}, z)$. The expenditure function can be defined by $E(u, \boldsymbol{p}, \boldsymbol{z}) = \min_q [\mathbf{p}'\mathbf{q} | U(\mathbf{q}, z) = u]$. Then, household equivalence scales are defined as:

$$S(u, \boldsymbol{p}, \boldsymbol{z}_h, \boldsymbol{z}_r) = \frac{E(u,\boldsymbol{p},\boldsymbol{z}_h)}{E(u,\boldsymbol{p},\boldsymbol{z}_r)} \quad (1)$$

where $\boldsymbol{z}_h$ and $\boldsymbol{z}_r$ are the household characteristics of two different households *h* and *r*. Therefore, an equivalence scale is a function that calculates the ratio of expenditures between two households, with different compositions but the same level of utility and facing identical prices. According to the literature, several criteria have been proposed for assessing approaches for equivalence scale estimation by the resulting scale values, mainly based on empirical regularities.

We assume that the equivalence scales only depend on household size *n*, such that they can be written as $S(u, \boldsymbol{p}, n)$. Following Dudel et al. (2021) the criteria are:

$$S(u, \boldsymbol{p}, n+1) > S(u, \boldsymbol{p}, n), \quad (2)$$

$$S(u, \boldsymbol{p}, n+1) \leq S(u, \boldsymbol{p}, n) + 1, \quad (3)$$

$$S(u, \boldsymbol{p}, n+i+1) - S(u, \boldsymbol{p}, n+i) \leq S(u, \boldsymbol{p}, n+i) - S(u, \boldsymbol{p}, n+i-1). \quad (4)$$

Criterion (2) states that equivalence scales must be strictly increasing functions of household size, based on the assumption that each additional household member incurs costs. Criterion (3) posits that the effect of the household size must be no greater than one, due to economies of scale. Criterion (4) states that the increase in scale should diminish or at least remain constant with household size.

3. The scales under different approaches

- MODEL 1 - Engel's approach: the Carbonaro scale

The Carbonaro scale is a behavioural approach that reflects consumer behavior in Italy, following Engel's (1895) method of using household expenditure to assess welfare. Engel's approach suggests that the share of household expenditure on essential goods varies by household type and decreases as income rises. According to this theory, two households have comparable well-being if they spend the same proportion of their total expenditure on basic goods. As household size increases, expenditure on essentials grows, requiring a higher total expenditure to maintain the same ratio. Equivalence scales are derived by comparing the expenditures of different household types that allocate the same share of their budget to essential goods. Starting from the beginning, essential goods were associated with food expenditures, so much so that Engel's approach was often referred to as the Engel food ratio method. So did Carbonaro, who employed a double logarithmic function to model the Engel curve for food expenditure:

$$logA = a + blogY + clogN + u\,, \quad (5)$$

where *A* is the food expenditure of the generic household *i*, *Y* is the total consumption and *N* is the household size, while *b* and *c* are, respectively, the regression coefficients of total expenditure and household size in logarithmic form. Deriving with respect to the number of members, and equating the elasticity with respect to the number of members of total expenditure to the one of food expenditure (in order to keep the food ratio constant), the elasticity is found:

$$e = \frac{\delta(logY)}{\delta(logn)} = \frac{c}{1-b}. \quad (6)$$

Using data for the period 1981-1983, the Carbonaro scale was calculated on the basis of the elasticity obtained (e = 0.67). Moving from the constant elasticity to the equivalence coefficients, setting the scale coefficient for a household with a size equal to one as the reference (k_1=1), the scale for households with a larger size is:

$$K_{N+1} = K_N(1 + \frac{e}{N}). \quad (7)$$

- MODEL 2 - Engel's approach: the absolute poverty threshold as a proxy of essential expenditures

The basic idea of Engel's method is that essential goods are necessities, where consumer choice is very limited. While food expenses were initially identified as essential expenditures, today's socio-economic context. Carbonaro himself addressed this issue in 1985. However, even if other essential expenses like clothing and housing are considered alongside food, another challenge remains. The cost of these essentials can vary based on household preferences and available resources; for instance, fulfilling the need for clothing may involve purchasing more items and/or opting for higher-cost clothing, involving discretionary spending.

The methodology of absolute poverty thresholds (Istat, 2009; Cutillo et al., 2022), calculates poverty thresholds as the monetary value of a basket of goods and services deemed essential. This basket includes not only food but also housing costs, and a residual category covering clothing, education, health, mobility, information, and communication needs. Absolute poverty thresholds are determined based on the minimum costs required to meet essential needs, both in terms of necessities and the lowest market prices available. Therefore, these thresholds accurately represent the expenditure necessary for households to meet their basic needs. In other words, they are exactly what was previously measured by the proxy variable food expenditure.

This approach aligns with Engel's theory, where households allocate their income first to meet minimum essential needs, and after this threshold consider discretionary spending. For example, once the threshold is met, households might choose higher-quality food, greater heating in the home, purchasing branded clothing, and so forth.

Thus, we estimate the same equation as in (1), substituting food expenditure with the absolute poverty thresholds, which are specific to each type of household. Indeed, the thresholds vary by region, type of municipality, and the number and age of household members. Since the threshold is exogenous to survey data, its share relative to household expenditures can exceed 1 for all households in absolute poverty. This contrasts with food expenditure, which inherently does not pose this issue. It would be appropriate to exclude absolute poor households, as they do not even meet a minimal standard of well-being. However, empirical results are largely consistent, and we include the entire sample in this initial analysis to ensure comparability under the same conditions. This is also why we employ the original formulation of Carbonaro as presented in equation (1), despite other functional forms (e.g., Carbonaro, 1991; Deaton and Muellbauer, 1986; De Santis, 1996):

$$logT_i = a + blogY_i + clogN_i + u_i \ , \tag{8}$$

where *T* represents the absolute poverty line for the generic household *i* and the other letters have the same meaning as in equation (5). It can be seen that, with the exception of the dependent variable, the model is the same as equation (5) and therefore equations (6) and (7) are still valid for calculating elasticity and scale.

- MODEL 3 - Engel's approach: absolute poverty threshold as a proxy of essential expenditures with territorial controls and non-constant elasticity

The model in equation (4) is modified in order to take into account two aspects related to the number of components. First, a number of territorial variables in dichotomous form are added in order to control for the different distribution of households by size between different regions and different types of municipality:

$$logT_i = a + blogY_i + clogN_i + \sum_j d_j K_{ji} + u_i \ , \tag{9}$$

where K_{ji} is equal to 1 if the generic household *i* resides in the generic territory *j* (given by the interaction between region and type of municipality) and 0 otherwise.

Secondly, we allow the elasticity not to be constant as the number of components changes, but to be differentiated as the number of components varies[1].

- MODEL 4 - A complete demand system: the Stone-Gary approach

The Linear Expenditure System (LES) proposed by Stone (1954) is the first expenditure system based not on a single equation, but on a system of equations,

[1] We run the same regression in pairs of numbers of components (e.g. one and two components; two and three components; and so on). Other forms could be used (e.g. De Santis, 1996, p. 46) but for the sake of comparison in this step of the analysis, we prefer to use equations as similar as possible to the one in Carbonaro (1985).

each of which covers expenditures for one of the *m* goods (in our case, the Coicop divisions of expenditures). Starting from the Stone-Geary utility function:

$$X_i = p_i a_i + b_i\left(X - \sum_{j=1}^{n} p_j a_j\right) \tag{10}$$

with *X* denoting total expenditures, X_i expenditure on good *i;* $p_i a_i$ being interpreted as the minimum expenditure on good *i* (prices by quantities); and b_i being the marginal budget share of good *i* , with the restriction that $\sum b_i = 1$. The system of equations can be estimated for each household type (Deaton 1975), whereby we set prices equal to one as a common practices in the literature (e.g. Dudel et al., 2021). The unit values of the scale *S* can then be estimated as follows:

$$S = \frac{\sum_{i=1}^{m} a_i^h}{\sum_{i=1}^{m} a_i^r} \tag{11}$$

where a_i^r is the reference household's minimum expenditure on good *i* facing prices *p* (set to one in our case) for good *i*; a_i^h is thus the comparison household's minimum expenditure on good *i.* The core idea of the model is that household expenditure on the good *i* depends on a minimum expenditure component (the constant term a_i) plus a fixed proportion of the supernumerary expenditure (the term in brackets in equation (10)). In this context, the scale is determined by the ratio of the sum of minimum expenditures between the household type *h* and the reference household *r*. Last, we estimated the parameters of interest non-linearly and we further added to equation (10) a quadratic term of the supernumerary expenditure to count for non-linearity of the Engel curves (Howe et al., 1979; Dudel et al., 2021).

- MODEL 5 - A complete demand system: the quadratic almost ideal demand system

The Almost Ideal (AI) demand system was first developed by Deaton and Muellbauer (1980) and it became very popular in the economic literature to measure household consumption behavior relative to change in commodity prices. Starting from the price-independent generalized logarithmic (PigLog) class of preferences and formalizing as in Ray (1983) in order to estimate equivalence scale:

$$w_i = a_i + \sum_{j=1}^{m} \gamma_{ij} \, log p_j + b_i^* \log\left(\frac{X}{SP}\right) \tag{12}$$

where w_i is the expenditure share for good *i*; a_i is a constant term. γ_{ij} represents the effect of a change in the price of commodity *j* on the share of expenditures on commodity *i*; $b_i^* = b_i + \eta_i d_h$, where b_i is the marginal effect of log expenditure and d_h is a dummy variable indicating the respective household type relative to the

reference household; P is a price deflator for expenditure[2]; S is the unit value of the equivalence scale for each household type relative to the reference household ($S = 1 + \rho d_h$, where ρ measures the needs of the comparison household relative to the needs of the reference household). The sum of η_i and b_i gives the expenditure elasticity of each household type relative the comparison household. The parameters of the model can be estimated non-linearly also adding a quadratic term (Banks et al., 1997), thus coming up with a Quadratic Almost Ideal (QAI) demand system.

4. The results

The findings presented in this section are derived from pooling data from three consecutive years: 2017, 2018, and 2019[3]. Initially, we compute an updated Carbonaro scale using the same methodology (MODEL 1). Compared to the original model, we observe a substantial increase in the elasticity of consumption relative to household size (from 0.67 to 0.87), resulting in equivalence scale values detailed in Table 1 (with a two-person household serving as the reference) which entail a decrease of the economies of scale over the years for larger households.

However, these outcomes starkly contrast with observed trends in Italian consumer spending over recent decades, as also noted by the Inter Institution Scientific Commission on Absolute Poverty. Specifically, there has been an increase in economies of scale for larger families, particularly in three areas. Food expenditure (facilitated by the widespread availability of large-scale distribution, which allows large families to buy large packages at lower unit prices). Rent and imputed rent costs (with a significant increase in the cost per square metre of smaller dwellings compared to larger ones). Energy expenditure (due to system charges on bills, which have shifted from variable consumption-based to fixed costs).

The obtained results suggest that the original Carbonaro scale is no longer suitable for the contemporary Italian context. First, under consistent methodology, the results differ significantly. Second, these divergent outcomes cannot be justified based on prior findings. It appears plausible that food expenditure, particularly in an advanced socio-economic context such as Italy nowadays, may no longer serve as the most appropriate proxy for defining household essential needs.

[2] The deflator P is obtained by setting a constant term and adding-up price products of different commodities. A formalized analysis of the deflator can be found in Deaton and Muellbauer (1980).

[3] As in the previous methodology, we used a pooled sample of three years (2017, 2018 and 2019) to obtain more robust estimates. Expenditure in 2020 and 2021 is distorted by the health emergency, which caused all expenditure categories except food and housing to fall, and 2022 was a year of very high inflation due to the economic recovery and the Russia-Ukraine war.

Table 1 – *Equivalence scales by different methods*.

Household size	Method					
	Carbonaro	Carbonaro updated	Absolute threshold without territorial controls and with constant elasticity	Absolute threshold with territorial controls and without constant elasticity	Stone-Geary	Quaids
1	0.60	0.53	0.67	0.69	0.70	0.87
2	1.00	1.00	1.00	1.00	1.00	1.00
3	1.33	1.44	1.25	1.29	1.17	1.15
4	1.63	1.86	1.45	1.55	1.34	1.21
5	1.91	2.27	1.63	1.78		
6	2.16	2.67	1.79	2.00	1.45	1.27
7+	2.40	3.07	1.94	2.24		
Relative poverty (%)	11.5	13.1	10.7	11.9	14.1	12.1

Before showing the results obtained through the Engel's approach based on absolute poverty thresholds (MODEL 2), some checks are necessary. First, given the same family structure, the amount of essential expenditure must increase as total expenditure increases, for all family sizes. The regression coefficient b in the equation $logT = a + bY$ is positive and statistically significant for all the different family sizes[4]. Secondly, given the same family structure, the share of essential expenses should decrease as total expenditure increases, for all family sizes. Through the equation $s_t = a + blogY$, where s_t is the share of essential expenses out of total expenses, the coefficient b is always negative and significantly different from 0.

Through this approach, the elasticity is 0.49, implying greater economies of scale for larger families, in line with the evolution of Italian society as described above[5]. Compared to the original Carbonaro scale, and using a two-person household as the reference, the coefficients now range from 0.67 for single-person to 1.94 for households with 7+ members (in the Carbonaro scale, 0.60 and 2.40 respectively).

When territorial controls are included and different elasticity with respect to component growth is allowed (MODEL 3), the scale adjusts slightly, with coefficients for larger families increasing. These values indicate that to maintain the same level of well-being, expenditure need to increase by 44.2% when moving from

[4] As the thresholds are exogenous and equal for identical households, this effect is due to territorial differentiation. The cost of living, and thus both the thresholds and household expenditure are higher in wealthier areas. From a mathematical point of view, whatever the motivation, it is important that this empirical rule is respected.

[5] Moreover, the R^2 is equal to 0.79 in respect of 0.42 when using the food expenditures.

one to two members, by 28% from two to three, by 20.4% from three to four, by 15% from four to five, by 12.4% from five to six, and by 12.2% from six to seven.

Regarding complete demand systems, estimated separately by household size, households with five or more members are grouped together. While MODEL 4 (the Stone-Gary approach) meets all the plausibility criteria outlined in Section 2, MODEL 5 (the quadratic almost ideal demand system) does not, particularly failing criterion (4) in section 2. Both scales exhibit a narrow range of values, indicating significant economies of scale, meaning that expenditure needs to increase minimally with household size to maintain the same level of well-being. Despite their theoretical validity, we argue that complete demand systems face three critical issues. First, it remains unclear whether to consider all expense types or focus solely on essential expenses when assessing a specific well-being level. Second, these systems are highly sensitive to the categorization of goods, which can vary significantly across different aggregations. Third, they fail to address a fundamental question about where to standardize welfare levels across households.

5. An empirical way to evaluate the scales

All the scales presented in the previous section, except for the Quaids (Table 1), meet the plausibility requirements outlined in Section 2. The results, however, vary significantly between scales. This variability is well-documented in the literature, as well as the fact that complete demand systems yield lower scale values than Engelian models. In evaluating the different approaches, in the previous section we considered their credibility within the socio-economic context. This section introduces a simple and empirical method to confirm or refute our choices, based on the relationship between absolute and relative poverty (based on equivalent consumption). In fact, relative poverty is a measure of inequality (Darvas, 2017), and it relies on the same rationale as the At Risk of Poverty Rate indicator on incomes. That is, also relative poverty indicates risk of poverty. It is thus expected that a significant percentage of absolute poor households are also relative poor[6]. In the 2017-2019 period, the incidence of relative poverty through the original Carbonaro scale is 11.5%, with 88.5% of absolute poor households also identified as relative poor. Using the updated Carbonaro scale (MODEL 1), the overlap decreases to 82.9%, though relative poverty rises to 13.1%, further indicating that food-ratio method is no longer valid.

Using the Engelian method based on absolute poverty thresholds (without differential elasticity and territorial controls – MODEL 2), the incidence of relative poverty is slightly lower (10.7% vs. 11.5%), and the overlap between absolute and relative poverty rises to 92.3%. When we allow the elasticity to vary with the number

[6] The overlap cannot be perfect (i.e. 100% of the absolute poor households are also relative poor), because absolute poverty thresholds differ for the cost of living in the territory, while the relative poverty threshold is single.

of household members and add territorial controls (MODEL 3), the incidence of relative poverty (11.7%) is nearly the same as the reference, and the percentage of absolute poor households who are also relative poor increases to 94.3%.

MODEL 4 shows a relative poverty incidence of 14.1%. Given the high number of households at risk of poverty, the percentage of absolute poor who are also relative poor is 92.7%, slightly lower than MODEL 3 despite the last has fewer households at risk of poverty. MODEL 5 yields a relative poverty incidence of 12.1%, but the percentage of absolute poor households at risk of poverty is notably low at 88.4%.

These results further support our choice to use an Engelian model with absolute poverty thresholds as a proxy for essential expenditure.

6. Discussion and conclusions

After the 2022 revision of the methodology for calculating absolute poverty in Italy, Istat is considering refining relative poverty measures by updating the Carbonaro scale (1985), which dates back to the 1980s and may no longer be in line with the current Italian economic system. In this paper, we have presented preliminary results by comparing different scales. In our considerations, we use both theoretical issues and empirical evaluations, also supported by a simple and intuitive method. Equivalence scales are used to calculate the incidence of relative poverty, an inequality indicator that identifies households at risk of poverty. It is expected that households that are poor in an absolute sense are also poor in a relative sense.

Our results suggest that Carbonaro scale is less and less suitable for today's Italy for two main reasons. First, new estimation with recent data produces different results. Second, the new estimates indicate a decrease in economies of scale over time for larger households, contradicting the trends observed in Italians' consumption expenditure in recent decades, as noted by the Inter Institution Scientific Commission on Absolute Poverty. In particular, economies of scale have increased for large households, especially for food, energy and rent expenditure.

Following an Engelian approach, we calculated equivalence scales using the absolute poverty methodology and its recent improvements. Absolute poverty thresholds represent the minimum monetary value of essential goods and services needed to avoid severe social exclusion, which goes beyond food and includes housing, clothing, education, health, mobility and communication. Calculated at minimum cost, these thresholds accurately reflect the expenditure required by households to meet essential needs, i.e. what was previously measured by the proxy variable food expenditure. The use of poverty thresholds has a double advantage. Firstly, the goods and services considered do not only include food, and thus seem to be better suited to the basic needs in an advanced socio-economic context. Furthermore, being calculated at the lowest affordable cost eliminates potential bias resulting from household preferences even in the purchase of essential goods.

The results of the Engelian scale based on absolute poverty thresholds outperform those of the updated Carbonaro scale, showing greater economies of scale for large households. Our model adjustments address regional disparities and different household composition, improving accuracy. In particular, the differentiated elasticity by household size maximises the overlap between households in absolute poverty and those at risk of poverty.

Scales based on complete demand systems have performed poorly for several reasons. Theoretical concerns include uncertainties about the inclusion of all or only essential expenditure and the inability to standardise welfare levels across households. Empirically, these systems do not meet the plausibility criteria of the equivalence scale, are sensitive to the choice of expenditure categories considered and show a low overlap between absolute and relative poverty.

Future research will prioritise the Engelian model, which ensures a consistent welfare level assumption based on equal allocation of expenditure to essential goods across different households. Ongoing considerations include refining the age scale, determining the inclusion or exclusion of households that are too poor (that do not even reach a minimum level of welfare) or too rich (that cannot be considered representative of the majority of the distribution) and exploring quadratic forms of the Engel curve. The latter point is relevant since the quadratic form implies that the elasticity depends on the level of expenditure itself, an issue that violates the income-independence assumption of Engel's model, but which can be resolved with some assumptions and simplifications.

Acknowledgements

We are grateful to A. Brandolini, S. Gerosa, M. Raitano and R. Zelli for the fruitful exchange of ideas and for their valuable suggestions.

References

BANKS J., BLUNDELL R., LEWBEL, A. 1997. Quadratic Engel curves and consumer demand. *Rev Econ Stat,* Vol, 79, No, 4, pp. 527–39.

BLUNDELL R., LEWBEL A., 1991. The information content of equivalence scales. *J Econ,* Vol. 50, pp. 49–68.

BUHMANN B., RAINWATER L., SCHMAUS G., SMEEDING T., 1988. Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (LIS) Database, *Review of Income and Wealth*, Vol. 94, pp.115-42.

CARBONARO, G. 1985. Nota sulle scale di equivalenza, in Commissione di indagine sulla povertà e sull'emarginazione, *La povertà in Italia,* Roma, Istituto Poligrafico e Zecca dello Stato.

CARBONARO G., 1991. La distribuzione quantitativa del reddito, in MARBACH G. (Ed.) *Statistica economica*, Torino, UTET, pp. 205-228.

CUTILLO A., RAITANO, M., SICILIANI, I., 2022. Income-Based and Consumption-Based Measurement of Absolute Poverty: Insights from Italy. *Social Indicators Research,* Vol. 161, pp. 689–710.

DARVAS Z. 2017. Why is it so hard to reach the EU's poverty target? *Policy Contribution Issue No. 1*, Bruegel, Brussels.

DEATON A., 1975. *Models and projections of demand in Post-War Britain*. Cambridge Studies in Applied Econometrics. Springer, Cambridge.

DEATON A., MUELLBAUER J., 1980. An almost ideal demand system. *American Economic Review* Vol.70, No.3:7, pp. 312-326.

DEATON A., MUELLBAUER J., 1986. On measuring child costs: with applications to poor countries. *Journal of Political Economy* Vol. 94, No. 4:7, pp. 20–744.

Deaton A., Zaidi S., 2002. Guidelines for Constructing Consumption Aggregates for Welfare Analysis, *World Bank Working Paper*, No. 135.

DE SANTIS G., 1996. *La misura della povertà in Italia. Scale di equivalenza e aspetti demografici*, Roma, Istituto Poligrafico e Zecca dello Stato.

DUDEL C., GARBUSZUS J.M., SCHMIED J., 2021. Assessing differences in household needs: a comparison of approaches for the estimation of equivalence scales using German expenditure data. *Empirical Economics.* Vol. 60: pp. 1629-59.

ENGEL E., 1895. Die Productions- und Consumtionsverhältnisse des Königreichs Sachsen. In: ANLAGE I., HEINRICH C. (Eds) *Die Lebenskosten Belgischer Arbeiter-Familien früher und jetzt*, Dresden.

FERREIRA F., RAVALLION M. 2011. Poverty and Inequality: The Global Context. In NOLAN, B., W. SLAVERDA, AND T. M. SMEEDING (eds.) *Oxford Handbook of Economic Inequality*. Oxford: Oxford University Press.

HOWE H., POLLAK R.A., WALES T.J. 1979 Theory and time series estimation of the quadratic expenditure system. *Econometrica*, Vol 47, No. 5, pp.1231–47.

ISTITUTO NAZIONALE DI STATISTICA, 2009. La misura della povertà assoluta. *Metodi e norme*, No. 39. Roma.

RAY, R. 1983. Measuring the costs of children: an alternative approach. *Journal of Public Economics*, Vol 22, No 1, pp.89–102.

STONE, R. 1954. Linear expenditure systems and demand analysis: an application to the pattern of British demand. *Economic Journal,* Vol. 64, No. 255, pp, 511–527.

Andrea CUTILLO, Istat, cutillo@istat.it
Alessandro ROMEO, Istat, alessandro.romeo@istat.it
Livia CELARDO, Istat, livia.celardo@istat.it

REVENUES AND EXPENDITURES IN ITALIAN MUNICIPALITIES: A MULTILEVEL LATENT CLASS APPROACH[1]

Roberta Varriale, Domenico Passante, Cecilia Vitiello

Abstract. The statistical register for Public Administrations (REPA) contains information on structural and economic variables on a subset of the Italian PA Institutions. REPA includes different subpopulations, we focus on the subpopulation of municipalities, that are naturally grouped in regions. The economic variables of REPA include accrual and cash values, for both revenues and expenditures. This paper aims to analyse the economic management behaviour of Italian municipalities. To this aim, we adopt a multilevel latent class approach, which allows to simultaneously cluster municipalities and regions in groups with some typical profile starting from the observed economic variables.

1. Introduction

The statistical Register of Public Administrations (REPA) is a new product of the Italian national institute of statistics (Istat). REPA is integrated into a single logical environment, the Italian Integrated System of Statistical Registers (ISSR) (Luzi *et al.*, 2019), comprising a series of statistical registers (basic, thematic and extended) that centralise and integrate data from administrative sources, statistical surveys carried out by the institute and new and emerging sources for the ongoing production of official statistics. REPA contains information on structural and economic variables for the subset of Italian Public Administrations in the so-called "S13 list" produced by Istat[2], classified in different subpopulations.

In this paper, we consider the subpopulation of Italian municipalities included in the S13 list, naturally grouped in regions, and analyse their economic management behaviour, both in terms of revenue and expenditure. We use a multilevel latent class analysis approach introduced by Vermunt (2003), which allows us to simultaneously classify municipalities and regions into groups with a typical profile based on the information provided by some observed continuous indicators. The structure of the paper is as follows. Section 2 describes the REPA informative contents. In Section

[1] Any opinions and conclusion expressed are those of the authors and do not necessarily respect the views of the Italian national institute of statistics.

[2] https://www.istat.it/it/archivio/190748.

3, the multilevel latent class analysis is introduced and Section 4 reports the obtained results. Section 5 contains some concluding remarks.

2. The statistical register for public administrations

As introduced, REPA contains information on structural and economic variables for the subset of Italian public administrations in the S13 list. It includes different subpopulations, such as local governments, regions and autonomous provinces, ministries, constitutional bodies, social security funds, sanitary districts, etc. Each subpopulation has a specific structure and classification for its economic data. The REPA production process involves data collection, harmonization and integration, as well as data review, editing, and imputation, each tailored to the individual characteristics of each subpopulation. For more details on the contents of REPA and its production process, see (Varriale *et al.*, 2024). REPA is still under development, but the design and implementation of the register are well-advanced for the subpopulation of local governments, including municipalities, unions of municipalities, provinces, mountain communities, and metropolitan cities (Varriale *et al.*, 2021). Since our analysis focuses on the subpopulation of Italian municipalities, we will use the term REPA in the following to refer to this subpopulation and not to the whole register.

The structural variables included in REPA are identifiers and register variables, territorial and stratification variables, status of activity, date of inclusion and possible exclusion from sector S13, transformation events, number of employees. The economic variables are the result of the integration and processing (i.e. imputation) of data from administrative sources. The primary source of information is the Public Administration Database (BDAP), while the auxiliary source is the Information System on the Operations of Public Bodies (SIOPE). Economic variables include accrual and cash values, for both revenues and expenditures. The accrual data for the revenue are the assessments (E1) while the cash data are the collections in accrual (E2) and the residual accounts (E3). For expenditures, the accrual data are the commitments (S1), the cash data are the payments on accrual (S2) and the residual accounts (S3). The information for both revenues and expenditures is organized into several hierarchical levels, following the structure of the certified balance sheet that all local governments are required to publish on an annual basis to certify their primary accounting data for the previous fiscal year (Guandalini *et al.*, 2021). For each statistical unit (local government), there are 148 items for revenues - classified in titles, categories and types - and 1431 items for expenditures - classified in titles, macroaggregates, missions and programs. The highest level of aggregation of the items identified in the certified balance sheet is Titles, which is the level of

aggregation used in our analyses. Table 1 shows an example of a certified balance sheet (revenue) for a single municipality. The first digit of the balance sheet code represents the title, and to derive the information at title level for each economic variable E1, E2 and E3, it is necessary to sum the economic values (symbol "xxx") identified by the same title.

Table 1 – *Example of a certified balance sheet for a single municipality. Revenue.*

Item	Balance sheet code	Title	E1	E2	E3
1	1010101	1	xxx	xxx	xxx
2	1010102	1	xxx	xxx	xxx
…	…	…	xxx	xxx	xxx
…	3059900	3	xxx	xxx	xxx
…	…	…	xxx	xxx	xxx
148	9029900	9	xxx	xxx	xxx

Concerning the revenue, in our analysis we use per capita values of: E1 - Assessments (accrual data), and E2 - Collections in accrual (cash data). We only use the information from the first 3 Titles, which sum is the current revenue: Current revenue tax based, contributory and equated (T1), Current transfers (T2), and Non-tributary revenue (T3). Concerning the expenditures, we use per capita values of: S1 - Commitments (accrual data), and S2 - Payments on accrual (cash data). We only use the information from the first 2 Titles: Current expenditures (T1), and Capital accounts expenses (T2). We consider the subpopulation of Italian municipalities included in the S13 list in 2021 and analyse their economic management behaviour, both in terms of revenue and expenditure, by applying a multilevel latent class analysis. The objective is to cluster both the municipalities and their regions to highlight some typical profiles of economic behaviour.

3. Multilevel latent class approach for REPA

Latent class (LC) analysis (Goodman, 1974) is a well know approach in social science research usually applied to clustering or constructing typologies with observed variables. With a hierarchical data structure, the usual assumption of independence of observations is violated, and a multilevel approach allows correct inferences treating the units of analysis as dependent observations (Snijders and Bosker, 2012). Multilevel data occur when there are data nested in several hierarchical levels, e.g. students in institutes, institutes in school districts, or patients in hospitals. In the case of a two-level data structure, (that is the case of our data) level 1 units are also known as lower-level or individual units, while level 2 units are higher-level or group units. The basic idea of a multilevel LC model, introduced by

Vermunt (2003) is that some of the model parameters are allowed to differ across level-2 units.

In order to describe our data we adopt a three index notation for responses: y_{kji} is the observed value on the *i-th* variable $(i = 1, \dots, I)$, on *j-th* lower unit $(j = 1, \dots, n_k)$ in the *k-th* $(k = 1, \dots K)$ higher level unit. In our analysis, according to the hierarchical nature of the data, municipalities represent lower-lever units *j*, and administrative regions higher-level units *k*. The total number of regions is *K=20*, while the number of municipalities, n_k, differs for each region. For each municipality, we observe economic variables from REPA: *6* revenues and 4 expenditures. The vectors $\boldsymbol{y}_{kj} = (y_{kj1}, .., y_{kji}, \dots y_{kjI})$ and $\boldsymbol{y}_k = (\boldsymbol{y}_{k1}, .., \boldsymbol{y}_{kj}, \dots \boldsymbol{y}_{k\boldsymbol{n}_k})$ contain the *I*-variate responses of municipality *j* from region *k*, respectively.

Applications of multilevel LC models results in simultaneously clustering individuals and groups in unobserved cluster, known as latent classes or mixtures: lower-level units are assumed to belong to one of *L* LCs differing in the distribution of the observed responses; higher-level units are assumed to belong to one of *H* higher-level LCs differing in the distribution of the lower-level LCs. The unobservable variables representing the lower- and higher-level classes membership are denoted by $x_{kj} = l \ \ (l = 1, \dots, L)$ and $w_k = h \ (h = 1, \dots; H)$, respectively.

The multilevel LC model (Vermut 2003) can be expressed in two basic equations. The first define the (mixtures) model for $f(\boldsymbol{y}_k)$, the marginal density for the full response vector of group *k*; that is:

$$f(\boldsymbol{y}_k) = \sum_{h=1}^{H} P(w_k = h) \prod_{j=1}^{n_k} f(\boldsymbol{y}_{kj} | w_k = h) \quad (1)$$

where $P(w_k = h)$ is the probability that group *k* belongs to higher-level LC *h*, while $f(\boldsymbol{y}_{kj}|w_k = h)$ is the conditional density for the response vector of individual *j* in group *k*. The second equation is:

$$f(\boldsymbol{y}_{kj}|w_k = k) = \sum_{l=1}^{L} P\big(x_{kj} = l|w_k = h\big) \prod_{i=1}^{I} f(y_{kji} | x_{kj} = l\,, w_k = h) \quad (2)$$

where $P\big(x_{kj} = l|w_k = h\big)$, is the probability that the individual *j* of group *k* belongs to LC *l* given that group *k* belongs to LC *h* and $f(y_{kji}|x_{kj} = l\,, w_k = h)$ is the conditional density for response variable *i* of individual *j* in group *k* given the membership to lower-level LC *l* and higher-level latent class *h*. In the first equation conditional independence between units is assumed inside each higher-level class. i.e. between municipalities. In the second conditional independence is assumed between the variables measured on the first-level units.

Finally, in order to complete the definition of the LC model, conditional densities $f(y_{kji}|x_{kj} = l\,, w_k = h)$ has to be specified, usually the choice is in the exponential family, in our case we assumed $f(y_{kji}|x_{kj} = l\,, w_k = h)$ normally distributed.

In our analysis, we applied two separate multilevel LC models for income and expenditure, the first using six variables, the second four variables. In both cases, we performed a logarithmic transformation of the original data. This choice was prompted by the observation of a strong right skewness of the distribution of all variables considered. This transformation makes it possible to emphasise the differences between the lowest values, which are the majority in both income and expenditure. The model is estimated with a maximum likelihood approach, with the program Latent GOLD, version 6.0 (Vermunt and Magidson, 2016).

A major issue in the application of the multilevel LC is the choice of the number of LCs. This issue is more complicated than in standard (single-level) LC analysis because it involves multiple, non-independent decisions (Lukociene *et al.*, 2010). To choose the number of higher- and lower-level classes we used both the information criteria (BIC, AIC), and substantive considerations related to the size and interpretation of the latent classes. In particular, following Lukociene et al. (2010), we used BIC with number of groups as the sample size, BIC(Ng), to decide on the number of higher-level classes. To contrast the tendency of information criteria to suggest a very high number of LCs, we compared the percentage difference between the information criteria of two successive models. Table 2 shows the results for income.

Table 2 – *Probability by municipal/regional cluster. Current revenue. Assessments and collections in accrual. Year 2021.*

h – l	BIC	AIC	BIC(Ng)	BIC % diff	AIC % diff	BIC(Ng) % diff
1 – 1	105549.7	105466.0	105477.9			
1 – 2	88893.8	88719.5	88744.4	-15.78	-15.9	
1 – 3	84056.0	83791.0	83828.8	-5.44	-5.6	
1 – 4	79829.8	79474.0	79524.8	-5.03	-5.2	
1 – 5	77522.9	77076.5	77140.2	-2.89	-3.0	
2 – 4	76421.2	76037.5	76092.3		-4.3	-4.32
3 – 4	75552.6	75141.1	75199.8		-1.2	-1.17
4 – 4	75352.3	74912.9	74975.7		-0.3	-0.30

4. Main results

In 2021, the number of municipalities was equal to 7,904 (S13 list). On the revenue side, we chose the solution with *H=3* latent classes (or clusters) at level-1 and *L=4* latent classes at level-2. For the expenditure, on the other hand, *H=4* and

L=2. In the following, we will use the term municipal cluster to denote the level-1 latent class and the regional cluster for the level-2 latent class. Results are presented in terms of median per capita values.

Table 3 shows the probabilities that region have of belonging to one of the 4 regional clusters, and the probability that municipalities have of belonging to one of the 3 municipal clusters, conditional to the regional cluster membership. Level-2 membership probabilities are evenly distributed, while the conditional level-1 probabilities show extreme values. Cluster 2 at regional level is characterised by a strong presence of cluster 1 at municipal level (0.78), while regional cluster 3 is characterised by a strong presence of municipal cluster 2 (0.87). In contrast, regional clusters 1 and 4 have a similar characterisation in terms of conditional probability of municipal cluster 1 (0.56 vs 0.51), but quite different for the other two municipal clusters.

Table 3 – *Probability by municipal/regional cluster. Current revenue. Assessments and collections in accrual. Year 2021.*

h	1	2	3	4
$P(w_k=h)$	0.34	0.25	0.25	0.15
$P(x_{kj}=l/w_k=h)$				
$l=1$	0.56	0.78	0.00	0.51
$l=2$	0.29	0.03	0.87	0.05
$l=3$	0.15	0.19	0.13	0.44

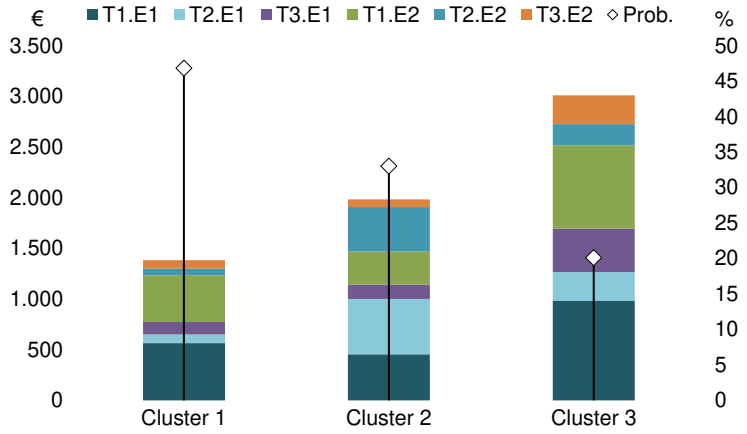
Figure 1 shows the estimated values per capita of municipal cluster. As introduced, per capita values refer to the first 3 revenue titles of municipalities (T1, T2 and T3) and refer to accrual (E1) and cash data (E2). As regards the estimated value of current revenue tax based, contributory and equated, collections (T1.E1) the maximum value was reached in cluster 3 (984 euros per capita), while the minimum values in cluster 2 (455 euros per capita). Regarding collections in accrual (T1.E2) the maximum estimated valued was equal to 823 euros per capita (cluster 3), while the minimum value was equal to 327 euros per capita (cluster 2).

The joint analysis of results in Table 1 and Figure 1 give us the possibility of describing the typical profiles of clusters, both at the municipal and regional level. In particular, for each budget item, we used the per capita value in order to make comparisons that take into account the different population of each municipality. It is possible to check whether the value of budget items for a municipality deviates excessively from the estimated values in the clusters with the highest probability and to analyse the reasons for this in order to verify the efficient use of resources.

Figure 1 – *Current revenue by municipal/regional cluster. Assessments and collections in accrual. Year 2021, per capita values in euros.*

The analysis of current revenue by municipal cluster (Figure 2) shows that the probability that a municipality belongs to cluster 1 was the highest ($P(x_{kj}=1)$ = 0.47). Cluster 1 has the per capita values of T1.E1 equal to 566 euros per capita, T2.E1 was 86 euros per capita and T3.E1 was 123 euros per capita.

Figure 2 – *Current revenue by municipal cluster and probability. Assessments and collections in accrual. Year 2021, per capita values in euros.*



It is also particularly important to be able to compare collections in accrual (E2) and Assessments (E1) by calculating the collection capacity[3] which measures the

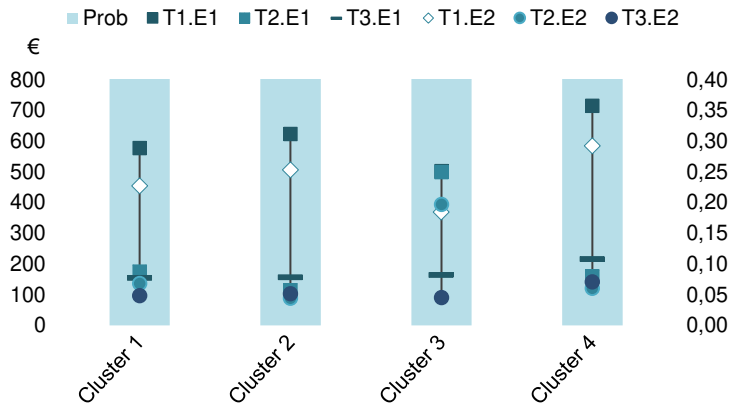
[3] The collection capacity is the percentage ratio between the collections in accrual accounts and assessments.

ability of municipalities to collect assessed revenue. As regards the titles making up current revenue, collection capacity was equal to 81% for current revenue tax based, contributory and equated, 79% for current transfers and 66% for non-tributary revenue. Comparing the values, the collection capacity for non-tributary revenue was the lowest, which means a lower efficiency in collecting that type of revenue.

Cluster analysis also allows us to calculate a number of relevant indicators used to assess the ability of municipalities to use their own revenues (tributary and non-tributary) to finance their institutional activities, in accordance with the principles laid down in the doctrine of fiscal federalism. In cluster 1 the degree of taxation autonomy[4], which measures the ability of the authority to levy resources by exercising its taxing power, was equal to 73%. The degree of financial autonomy[5], which measures the degree of autonomy of the municipality, i.e. the ratio of own revenues to current revenues, was equal to 84%.

With regard to regional clusters, the highest probability (0.34) is to belong to the cluster 1, while the lowest probability belongs to cluster 4 (Figure 3). The per capita values of cluster 1 were 578 euros (T1.E1) as regards the assessments and 455 euros (T1.E2) as regards collections in accrual. In cluster 1 the degree of taxation autonomy was equal to 64% and the degree of financial autonomy was equal to 83%.

Figure 3 – *Current revenue by regional cluster and probability. Assessments and collections in accrual. Year 2021, per capita values in euros.*



[4] The degree of taxation autonomy is the percentage ratio between current revenue tax based, contributory and equated and current revenue.

[5] The degree of financial autonomy is the percentage ratio between current revenue tax based, contributory and equated + non-tributary revenue and current revenue.

Table 4 shows estimates values of conditional probability for the current expenditures (T1) and capital accounts expenses (T2). Also for expenditure, level-2 membership probabilities - $P(w_k=h)$ - are evenly distributed, while the conditional level-1 probabilities - $P(x_{kj}=l/w_j = h)$ - show extreme values. Cluster 2 and 4 at regional level are characterised by a strong presence of cluster 1 and cluster 2 at municipal level, respectively (0.83 and 0.93), while regional cluster 1 and 3 are less extreme.

Table 4 – *Probability by municipal/regional cluster. Current expenditures and capital accounts expenses. Commitments and payments in accrual. Year 2021.*

h	1	2	3	4
$P(w_k=h)$	0.35	0.29	0.20	0.15
$P(x_{kj}=l/w_j = h)$				
$l=1$	0.66	0.83	0.46	0.08
$l=2$	0.34	0.17	0.54	0.93

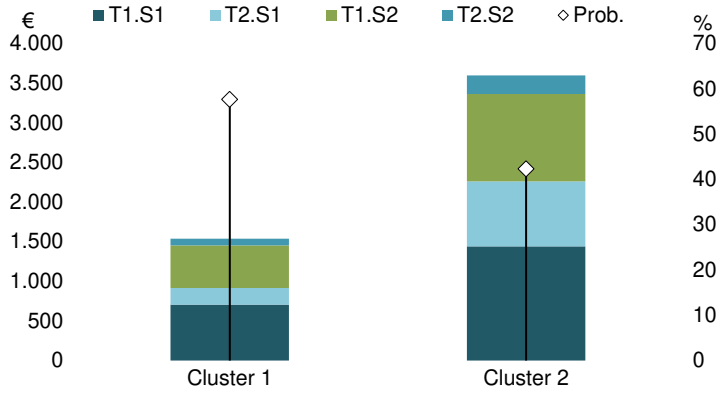
The per capita values of current expenditure (T1), both commitments and payments in accrual (S1 and S2), were higher in cluster 2, respectively 1,441 and 1,100 euros (Figure 4). Also capital accounts expenses (T2), commitments and payments in accrual, were higher in cluster 2 (respectively 819 and 237 euros).

Figure 4 – *Current expenditures and capital accounts expenses by municipal/regional cluster. Commitments and payments in accrual. Year 2021, per capita values in euros.*

Figure 5 shows estimated values and probability ($P(x_{kj}=1)$) referring to municipal cluster for expenditure. The probability for a municipality to belong to

cluster 1 was the highest (0.58) but with lower per capita values. In cluster 1, which has the highest probability, the spending capacity[6], which measures the ability of municipalities to pay the amounts committed, was equal to 76% for current expenditures (T1) and 41% for capital accounts expenses (T2). In cluster 2 the spending capacity was lower for capital accounts expenses (29%), while was equal to cluster 1 for current expenditures. This means, firstly, that for capital accounts expenses the percentage of expenditure realised, compared to committed sums, was 35 percentage points lower than for current expenditure (cluster 1). Secondly, the municipalities included in cluster 1 are able to spend more of the sums allocated for long-term investments than those included in cluster 2.

Figure 5 – *Current expenditures and capital accounts expenses by municipal cluster and probability. Commitments and payments in accrual. Year 2021, per capita values in euros.*

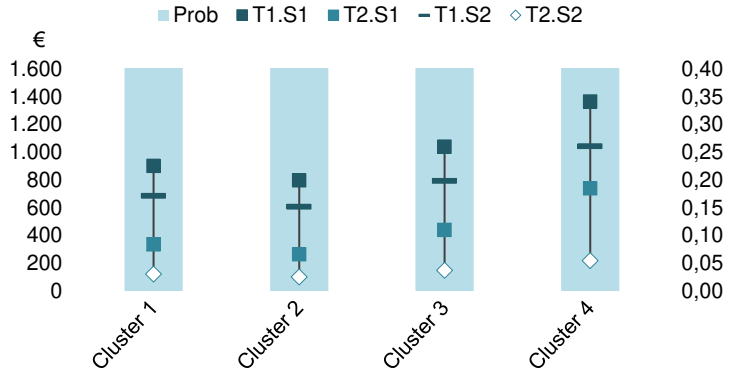


In Figure 6 the analyses of estimated values and probabilities refer to regional cluster shows that cluster 1 has the highest probability (0.35).

As regards the commitments, current expenditures was equal to 903 euros per capita and capital accounts expenses was equal to 335 euros per capita. The spending capacity was equal to 76% for current expenditures, while as regards the capital accounts expenses the indicator was equal to 36%. The same considerations made for municipal clusters regarding the ability to spend all budgeted resources also apply to regional clusters.

[6] Spending capacity is calculated as the percentage ratio between payments on accrual accounts and commitments.

Figure 6 – *Current expenditures and capital accounts expenses by regional cluster and probability. Commitments and payments in accrual. Year 2021, per capita values in euros.*



5. Concluding remarks and further developments

The statistical register for Public Administrations REPA is an object of the Italian Integrated System of Statistical Registers (ISSR), with information on structural and economic variables on a subset of the Italian PA. This paper aims to analyse the economic management behaviour of Italian municipalities. We adopt the multilevel latent class analysis which allows to simultaneously cluster municipalities and regions in groups with some typical profile. The information used for the analysis are the revenues and the expenditures. For each budget item, we used the per capita value in order to make comparisons that take into account the different population of each municipality.

The obtained results shows the characteristics of the municipal and regional cluster in terms of cluster size and value of each budget item analysed. It is also possible to calculate a number of relevant indicators used to assess the ability of municipalities to finance their institutional activities, in accordance with the principles laid down in the doctrine of fiscal federalism and to check whether for a municipality the value of the indicators and the budget items making up the clusters deviates excessively from the estimated value and to analyse the causes in order to verify the efficient use of resources. The result obtained can also help politicians and policymakers by providing useful information on the most common budget structures and the main characteristics of the main budget items.

We chose to analyse pro-capita data in order to focus on the economic structure of lower and higher units (municipalities and region) and partially absorb the effect

of the different magnitude of the municipalities. On the other side using median rather than means allows a description of clusters not affected by extreme values. Further analysis could take into account other variables as dimension of municipalities, geographical collocation, administrative characteristics, both as a support to cluster description or to carry out specific conditional analysis. Moreover, comparing the obtained results with those from an analysis that considers the two levels (municipalities and region) separately could provide insights into the substantive findings offered by multilevel techniques.

References

GOODMAN L.A. 1974. The analysis of systems of qualitative variables when some of the variables are unobservable: Part I -A modified latent structure approach, *American Journal of Sociology*, Vol.79, pp. 1179-1259.

GUANDALINI A., PASSANTE D., VARRIALE R. 2022. The revenues of local governments in the statistical register for public administrations: inequality decomposition by sources, *RIEDS*, Vol. 76, No.3, pp. 17-28.

LUKOCIENE O., VARRIALE R., VERMUNT J.K. 2010 The Simultaneous Decision(s) about the Number of Lower- and Higher-Level Classes in Multilevel Latent Class Analysis, *Sociological Methodology*, Vol. 40, pp. 247-283.

LUZI O., ALLEVA G., SCANNAPIECO M., FALORSI P.D. 2019. Building the Italian Integrated System of Statistical Registers: Methodological and Architectural Solutions. *ESS Workshop on the use of administrative data and social statistics*, Valencia, 4-5 June2019.

SNIJDERS T.A.B., BOSKER R. 2012. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling, second edition*. London etc.: Sage Publishers.

VARRIALE R., LORI M., MANTEGAZZA F. 2021. Stato di avanzamento dei lavori Focus: Frame PA enti territoriali. Nota tecnica Istat, Dicembre 2021.

VARRIALE R., ALBO N., CASAGRANDE C., OLIVIERI V. 2024. The statistical register for public administrations: the reference framework and some methodological aspects. *RIEDS*, Vol. 78, No.1, pp. 173-184.

VERMUNT J.K. 2003. Multilevel Latent Class Models, *Sociological Methodology*, Vol. 33, No. 1, pp. 213-239.

VERMUNT J.K., MAGIDSON J. 2016. *Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax*. Belmont, MA: Statistical Innovations Inc.

Roberta VARRIALE, Sapienza University of Rome, roberta.varriale@uniroma1.it
Domenico PASSANTE, Istat, passante@istat.it
Cecilia VITIELLO, Sapienza University of Rome, cecilia.vitiello@uniroma1.it

THE CITY AS A MEASURE OF SUSTAINABILITY. A MULTIDIMENSIONAL ANALYSIS[1]

Giuseppe Lecardane, Monica Carbonara, Adriano Pareto

Abstract. Cities are at a the turning point in the global comparison of sustainable development, as they are the centre of a growing majority of the world's population. They are the engine of local and national economies and represent the hub of well-being; more than 80% of global economic activities are concentrated in urban centres. The climate crisis and the need to protect the environment have pushed all the countries of the world to reorganise their urban centres, with the intention of creating real "sustainable cities". Goals 11 of the 2030 Agenda of the United Nations calls for making cities and human settlements more inclusive, safe, resilient and sustainable. Cities must meet specific environmental, social, economic criteria and be redesigned in their spatial, social and economic organisation. They must become a laboratory of sustainability and inclusion, able to forge a strong alliance with its citizens and the environment. The aim of this paper is to analyse the requirements for a new urban centre model through exploratory methods of multivariate analysis and the comparison of characteristic indicators that in a common vision can bring out significant peculiarities and dynamics in the urban context. The multidimensional complexity of the study required the identification, selection and measurement of a set of indicators relating to the macro-areas of a demographic, social, economic and environmental nature and a multivariate synthesis analysis for comparisons in terms of urban sustainability. A study that offers insights to understand the logic and dynamics of our cities as the keystone for the interpretation and regulation of urban, social and economic development processes.

1. Introduction

Currently, more than half of the world's population lives in urban areas, a percentage that is expected to increase to 68% by 2050. Cities are the engine of local and national economies, but besides the opportunities, urbanisation also brings considerable challenges.

[1] The paper is the result of the common work of the authors. In particular: sections are attributed as follows: M. Carbonara paragraphs 1 and 2.1 and 4, A. Pareto paragraph 2.2 and G. Lecardane paragraph 3.

The climate crisis and the need to protect our environmental heritage have prompted countries around the world to reorganise their urban centres with the aim of creating 'sustainable cities'. The concept of sustainable cities is closely linked to the Sustainable Development Goals, set in 2015 by the 193 UN member states. In Goals 11 of the UN 2030 Agenda, the common goal is to make cities and human settlements more inclusive, safe, resilient and sustainable. To achieve this goal, cities must meet specific environmental, social and economic criteria, integrate innovative technologies, have an efficient and accessible transport system, expand public spaces and green areas making them inclusive and safe, implement careful planning of human settlements and, finally, better manage their energy resources for a lower impact on the environment.

The aim of this work is to determine the conditions necessary to identify an ideal model of a 'sustainable city' by establishing a set of individual indicators for macro-areas of a demographic, social, economic and environmental nature. Through a multivariate synthesis analysis, Italian provincial capitals are compared in terms of urban sustainability.

The study also aims to offer food for thought on the logic and dynamics of our cities and the related urban, social and economic development processes.

2. Data and method

2.1 Data

A set of indicators for 9 macro-areas (education, work, economic well-being, politics and institutions, culture, social services, territory and environment, established economy, infrastructure and mobility) (Tab. 1) was identified on the basis of "A misura di comune", a multi-source system, in which sources of an experimental nature are valorised alongside other, more consolidated ones.

The objective of the system is to provide an increasingly detailed integrated information framework of indicators available at municipal level, useful for the planning, programming and management tasks of local authorities.

Table 1 – *Macro-areas and individual indicators.*

Macro-area	Indicator
Education	a1. Alphabetical proficiency of students
	a2. Numerical proficiency of students
Work	a3. Employment rate
	a4. Inactivity rate
Economic well-being	a5. Irpef taxpayers with income of less than 10,000 euros - Incidence on total taxpayers
Politics and institutions	a6. Women and political representation at local level (Municipal Councils) - Impact on total elected
	a7. Women in Municipal Councils – Impact on total Council members
Culture	a8. Libraries registered in the National Library Registry per 100 thousand inhabitants
Social services	a9. Expenditure on social interventions and service for municipalities by type of user
Territory and environment	a10. Urban air quality – PM10
	a11. Total density of green areas
	a12. Differentiated collection of urban waste (Incidence of differentiated collection on total waste)
	a13. Cars in circulation with emission standards lower than Euro 4 (Incidence on total cars)
Established economy	a14. Entrepreneurship rate
Infrastructure and mobility	a15. Road accident rate a16. Density of bike paths a17. Seat-km offered by local pubblic transport a18. Availability of pedestrian areas

Source: Istat

2.2 Composite index construction

The 18 individual indicators of sustainability have different units of measurement and ranges; some have positive polarity[2] (e.g., employment rate), while others have negative polarity (e.g., urban air quality – PM10). Therefore, they were normalised by transformation into z-scores and the signs of the indicators with negative polarity were reversed. Assuming that the indicators of each macro-area are substitutable (i.e., a deficit in one component may be compensated by a surplus in another and

[2] The polarity of an individual indicator is the sign of the relation between the indicator and the phenomenon to be measured (+ if the individual indicator represents a dimension considered positive and - if it represents a dimension considered negative).

vice versa), a set of 9 full compensatory composite indices (one for each macro-area) was obtained by arithmetic mean of individual indicators. In the case of macro-areas with only one individual indicator (i.e., economic well-being, culture, social services and established economy) no aggregation was done. Finally, the Wroclaw taxonomic method was applied for constructing a ranking of the cities according to their sustainability (Mazziotta and Pareto, 2017). The method rests on the concept of 'ideal unit': a hypothetical city that has, for each indicator, the most desirable value among all the cities (optimal score). The Euclidean distance from each city to the 'ideal unit' is then calculated as follows:

$$\mathrm{D}_i = \sqrt{\sum_{j=1}^{m} (z_{ij} - z_{0j})^2}$$

where z_{ij} is the standardised value[3] of the index j for the city i and z_{0j} is equal to $\max_i (z_{ij})$, as all the 9 indices have positive polarity. The composite index for the unit i is given by:

$$\mathrm{WTM}_i = \frac{\mathrm{D}_i}{\overline{\mathrm{D}}_0 + 2\sigma_0}$$

where $\overline{\mathrm{D}}_0$ and σ_0 are the mean and the standard deviation of the distances D_i.

WTM is a partially compensatory composite index, since we assume that a deficit in one area may be only partially compensated by a surplus in another and viceversa. The index is equal to zero when the distance between a given city and the 'ideal unit' is null (all the values coincide). The higher is the index, the greater is the difference between the two units.

3. Results

Through a multivariate synthesis analysis, the Wroclaw taxonomic method was applied to the Italian provincial capitals to build a ranking based on urban sustainability and comparability with respect to the ideal city. A hypothetical city that has, for each indicator, the most desirable value among all the cities (optimal score).

With WTM method, a weighting of the elementary indicators is implicitly implemented, which are more influential on the synthetic index, the greater the distances recorded with respect to the ideal situation.

[3] Each indicator is transformed into a standardised variable with mean 0 and variance 1.

Table 2 – *Summary statistics – 2021.*

Indicator	Mean	Median	Min	Max	Std. dev.	CV
a1	199.7	200.8	181.4	213.0	6.8	3.4
a2	45.7	46.7	35.0	54.2	5.0	10.9
a3	196.2	197.6	173.2	215.8	9.9	5.0
a4	49.4	48.3	41.4	58.5	4.1	8.3
a5	26.1	25.3	18.8	40.8	4.9	18.9
a6	30.1	28.1	9.4	47.5	8.1	27.1
a7	40.7	40.0	10.0	66.7	7.6	18.6
a8	30.5	27.3	3.1	141.3	19.2	63.0
a9	160.7	142.9	13.4	618.8	87.6	54.5
a10	24.1	18.0	0.0	75.0	20.2	84.0
a11	18.9	13.9	0.3	71.2	16.1	84.9
a12	62.1	66.5	11.3	87.5	16.1	25.9
a13	26.9	26.0	2.7	50.9	7.9	29.5
a14	90.0	89.0	52.1	145.7	16.9	18.8
a15	3.6	3.6	1.2	7.1	1.1	31.3
a16	40.4	20.0	0.0	197.8	49.1	121.4
a17	2.321.2	1.688.0	158.0	16.827.2	2.202.0	94.9
a18	43.4	23.4	0.0	684.0	82.6	190.2

Source: Istat

Starting from the comparison of the distributions of the 18 indicators through the main statistical measures of location and statistical dispersion (Table 2), the outcome of the descriptive analysis outlines almost symmetrical distributions with mean and median similar to each other but heterogeneous, with considerable levels of dispersion (Std. dev. and CV) for several indicators (a9, a16, a17 and a18).

Suitable characteristics for the application of the WTM synthesis method and the purposes of the analysis.

Subsequently, assuming the principle of substitutability of the indicators of each macro-area, the synthesis analysis was concentrated on a more limited number of complete compensatory composite indices from 18 to 9 (one for each macro-area) by means of the arithmetic mean of the individual indicators (Table 3). Furthermore, the polarity (positive or negative) of the relationship between indicator and phenomenon was specified.

Table 3 - *Indicators for Wroclaw Taxonomic Method (WTM) and polarity (+/-) – 2021.*

Indicator	Macro-area	Polarity
V1	Education	+
V2	Work	+
V3	Economic Well-Being	+
V4	Politics and Institutions	+
V5	Cukture	+
V6	Social Services	+
V7	Territory and Environment	+
V8	Established Economy	+
V9	Infrastructures and Mobility	+

Source: Istat

The indicators have been normalized and standardized to obtain data purified from the units of measurement and the comparison process. This approach is of absolute importance when dealing with the multidimensional phenomenon; combination of domains that must be as homogeneous as possible.

Figure 1 shows the cartogram of the WTM index developed for the 109 Italian provincial capitals. The values assumed by the synthetic indicator highlight the positioning of Italian cities in terms of sustainability, which decreases as one proceeds towards the highest positions.

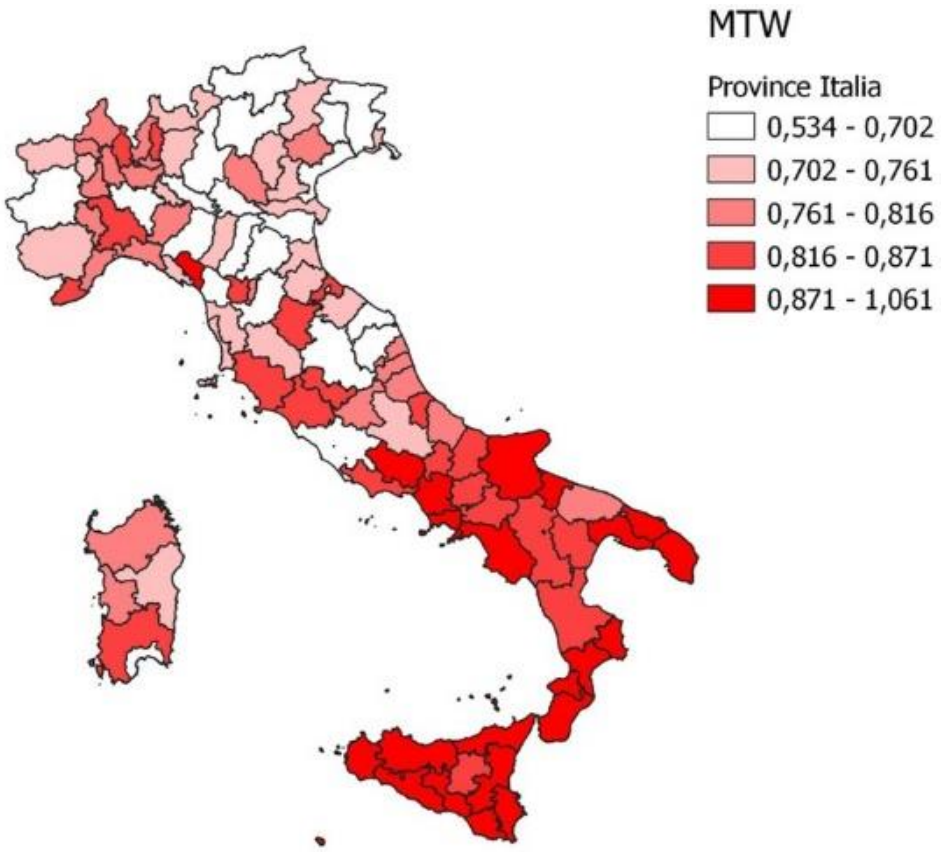
The outcome of the analysis returns the classic subdivision of the decreasing territorial dualism North and South.

From the ranking (Table 4) we can observe the positioning of Italian municipalities according to the degree of sustainability that decreases moving towards the highest ranks. The positioning distinguishes Bologna and Trento as the cities with the best performances of urban sustainability, differently Catania as the city with the highest negative impact.

In the ranking of the top five most sustainable cities, Macerata, Trieste and Venice stand out, cities of small and medium demographic size in central-northern Italy. It is also interesting to note the ninth place occupied by Cagliari among the most sustainable, smart and inclusive Italian cities.

At the bottom of the list, with the worst livability, are the cities of southern Italy and in particular: Trapani, Barletta, Crotone and Agrigento. These are the cities that struggle the most to respond to urban emergencies and to guarantee an acceptable quality of life for their inhabitants.

Figure 1 - *Map of the WTM index – 2021.*



The characteristics of a sustainable city are intrinsic to those of a circular city, embracing the use of renewable energy sources, virtuous waste management, the adoption of practices for sustainable mobility and the reuse of the material and energy resources used. The goal is to organize more efficient, livable, green and digital city spaces, perfectly integrated with each other and it is clear how urgent it is to make cities more inclusive, safe, resilient and sustainable.

In Italy there is still a long way to go but let's see in detail how far our cities are from the ideal and sustainable one using the WTM method. In the analysis of the 9 sustainability indicators, the "ideal" city has a WTM index equal to zero with optimal performance. Furthermore, the index is equal to zero when the distance between a given city and the ideal unit is zero (all values coincide). The higher the index, the greater the difference between the two units.

Table 4 - *WTM ranking - Provincial capital cities – 2021.*

Rank	Provincial capitals	Euclidean distances	WTM	Rank	Provincial capitals	Euclidean distances	WTM
1	Bologna	7,319	0,534	56	Verona	10,857	0,792
2	Trento	7,904	0,577	57	Teramo	10,873	0,793
3	Macerata	8,406	0,613	58	Treviso	10,962	0,800
4	Trieste	8,501	0,620	59	Ascoli Piceno	11,012	0,803
5	Venezia	8,543	0,623	60	Prato	11,068	0,807
6	Pavia	8,659	0,632	61	Chieti	11,071	0,808
7	Bolzano	8,692	0,634	62	Rieti	11,084	0,809
8	Pordenone	8,709	0,635	63	Monza	11,150	0,813
9	Cagliari	8,750	0,638	64	Asti	11,174	0,815
10	Firenze	8,860	0,646	65	Terni	11,206	0,817
11	Parma	9,097	0,664	66	Varese	11,236	0,820
12	Udine	9,103	0,664	67	Campobasso	11,274	0,822
13	Roma	9,245	0,674	68	Matera	11,366	0,829
14	Ferrara	9,291	0,678	69	Grosseto	11,367	0,829
15	Mantova	9,291	0,678	70	Rimini	11,380	0,830
16	Perugia	9,355	0,682	71	Benevento	11,428	0,834
17	Lucca	9,372	0,684	72	Pistoia	11,435	0,834
18	Ancona	9,383	0,684	73	Viterbo	11,572	0,844
19	Modena	9,436	0,688	74	Carbonia	11,610	0,847
20	Torino	9,478	0,691	75	Cosenza	11,654	0,850
21	Brescia	9,589	0,699	76	Imperia	11,758	0,858
22	Cremona	9,626	0,702	77	Lecco	11,777	0,859
23	Padova	9,632	0,703	78	Potenza	11,786	0,860
24	Biella	9,687	0,707	79	Alessandria	11,786	0,860
25	Ravenna	9,807	0,715	80	Avellino	11,804	0,861
26	Reggio Emilia	9,812	0,716	81	Isernia	11,818	0,862
27	Belluno	9,815	0,716	82	Enna	11,836	0,863
28	Livorno	9,828	0,717	83	Latina	11,843	0,864
29	Pisa	9,841	0,718	84	Arezzo	11,850	0,864
30	Siena	9,844	0,718	85	Pescara	11,928	0,870
31	Cuneo	9,869	0,720	86	Lecce	11,935	0,871
32	Gorizia	9,885	0,721	87	Salerno	11,952	0,872
33	L'Aquila	10,050	0,733	88	Massa	12,231	0,892
34	Bergamo	10,061	0,734	89	Brindisi	12,365	0,902
35	Pesaro	10,095	0,736	90	Catanzaro	12,370	0,902
36	Nuoro	10,096	0,736	91	Caltanissetta	12,409	0,905
37	Lodi	10,108	0,737	92	Frosinone	12,478	0,910
38	Sondrio	10,252	0,748	93	Caserta	12,488	0,911
39	Aosta	10,272	0,749	94	Ragusa	12,508	0,912
40	La Spezia	10,286	0,750	95	Napoli	12,640	0,922
41	Rovigo	10,335	0,754	96	Messina	12,751	0,930
42	Forlì	10,376	0,757	97	Siracusa	12,752	0,930
43	Vicenza	10,399	0,759	98	Reggio	12,807	0,934
44	Novara	10,484	0,765	99	Taranto	12,828	0,936
45	Fermo	10,498	0,766	100	Vibo Valentia	12,843	0,937
46	Savona	10,527	0,768	101	Foggia	13,007	0,949
47	Genova	10,538	0,769	102	Andria	13,030	0,950
48	Bari	10,614	0,774	103	Palermo	13,090	0,955
49	Verbania	10,641	0,776	104	Trani	13,101	0,956
50	Milano	10,669	0,778	105	Agrigento	13,262	0,967
51	Vercelli	10,679	0,779	106	Crotone	13,394	0,977
52	Como	10,728	0,783	107	Barletta	13,435	0,980
53	Piacenza	10,792	0,787	108	Trapani	13,624	0,994
54	Sassari	10,814	0,789	109	Catania	14,541	1,061
55	Oristano	10,841	0,791				

In figure 2, the Euclidean distances of the provincial capitals belonging to the main territorial divisions (North-East, North-West, Centre, South Italy and in the Major Islands) have been calculated with respect to the ideal value.

Figure 2 - *Distances of Italian cities from the "ideal sustainable city" – 2021.*

The results show better performances and those closer to the "ideal" value in the cities of Bologna and Trento in the North-East (7.319 and 7.904), Pavia in the North-West (8.659), Macerata and Florence in the Centre (8.406 and 8.860), L'Aquila and Bari in the South (10.050 and 10.614), Cagliari in the Major Islands (8.750).

On the contrary, the cities that stand out for worse performances and furthest from the ideal value are Lecco and Alessandria in the North-West (11.777 and 11.786), Treviso and Rimini in the North-East (10.962 and 11.380), Massa and Frosinone in the Centre (12.231 and 12.478), Crotone and Barletta in the South (13.394 and 13.435), Trapani and Catania in the Major Islands (13.625 and 14.541)

The challenge for more sustainable and liveable cities in Italy is still a distant goal, despite the fact that there are realities and good practices in the territories that go in the right direction.

Considering the constant increase in the urban population, it is now essential to focus on sustainable city models, real laboratories to guide new strategies for the transformation of our societies.

4. Conclusions

The analysis of the Italian provincial capitals proposed in this paper has made it possible to identify an ideal model of a 'sustainable city'. We carried out a multivariate analysis using a set of demographic, social, economic and environmental indicators in order to compare Italian provincial capitals in terms of urban sustainability. The results confirmed the North-South dualism. Southern cities, in fact, struggle to respond to urban emergencies and guarantee an acceptable quality of life. The study also offered food for thought on the logic and dynamics of our Italian cities and their urban, social and economic development processes.

References

DEL VECCHIO F. 2010. *Statistica per l'analisi di dati multidimensionali*. Padova: CLEUP.

LECARDANE G., FULLONE F., CARBONARA, M. 2024. A Multidimensional analysis for monitoring environmental health in Italy, *RIEDS. Rivista Italiana di Economia Demografia e Statistica*, Vol. LXXVIII, No. 2, pp. 259-270.

LECARDANE G. 2022. Environmental monitoring in Italian cities. A multidimensional analysis with composite indices. *RIEDS. Rivista Italiana di Economia Demografia e Statistica*, Vol. LXXVI, No. 4, pp. 109-119.

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, Vol. 70, Cham: Springer, pp. 159-191.

Giuseppe LECARDANE, Istat, lecardan@istat.it
Monica CARBONARA, Istat, mocarbon@istat.it
Adriano PARETO, Istat, pareto@istat.it

A QUANTITATIVE ANALYSIS OF THE RELATIONSHIPS BETWEEN THE SUSTAINABLE DEVELOPMENT GOALS[1]

Viviana Amati, Marta Nai Ruscone

Abstract. The 2030 Agenda for Sustainable Development has the seventeen Sustainable Development Goals (SDGs) at its heart and describes the path to ensure well-being, reduce inequality, encourage economic growth, and preserve the environment. Goals and targets are interconnected by construction. A standard approach to quantify synergies and trade-offs relies on the computation of positive and negative correlation coefficients between the indicators. In this work, we propose a method based on copulas to analyze SDGs interactions, offering a practical and innovative solution. Copulas allow studying more specific types of dependence beyond the linear association described by the correlation coefficients. We illustrate the practical application of this approach by analyzing the interactions between the goals' targets related to health and well-being (Goal 3) and those related to biodiversity (Goals 6,13,14,15).

1. Introduction

The 2030 Agenda for Sustainable Development was adopted by all the members of the United Nations in 2015. Its origin dates to the United Nations Conference on Sustainable Development (Rio+20) that took place in Rio de Janeiro (Brazil) in June 2012, when the nations recognized that they must act in collaborative partnership to achieve sustainable development. The 2030 Agenda is a pivotal document with the seventeen Sustainable Development Goals (SDGs) at its core. Those goals build upon and replace the Millennium Development Goals2, milestones that the countries must have achieved in 2015 to eradicate poverty and improve the life on Earth.

The concept of sustainable development is multidimensional, and its dimensions concern the social, economic and environment domains (Redclift, 1991). It is well-established that countries should deploy strategies to eradicate poverty and deprivations. Those strategies are successful only if the countries simultaneously adopt strategies fostering economic growth, addressing social needs — such as

[1] The Authors contributed equally to this work
[2] https://research.un.org/en/docs/dev/2000-2015

health, education, and inequality — and preserving the environment (United Nations, 2015).

The seventeen SDGs align with the areas defining the concept of development. They are classified into three groups (Boar *et al.*, 2020). One group relates to the social area and includes the Goals 1 (no poverty), 2 (zero hunger), 3 (good health and well-being), 4 (quality education), 5 (gender equality), 7 (affordable and clean energy), 11 (sustainable cities and communities) and 16 (peace, justice and strong institutions). Another group concerns the economic area and comprises Goals 8 (decent work and economic growth), 9 (industry, innovation and infrastructure), 10 (reduced inequalities) and 12 (responsible consumption and production). The last group relates to the environment domain and collects Goals 6 (clean water and sanitation), 13 (climate action), 14 (life below water) and 15 (life on land). Goal 17 does not belong to any group since it regards the governance and the countries' commitment to act to monitor and achieve the Goals.

Each Goal includes targets whose progress is measured by a set of indicators created by the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs). Goals and targets are interconnected by construction (Griggs et al., 2014; Le Blanc, 2015) since the sociological, economic and environment area are interdependent. Several approaches have been used to describe the interactions among the goals, such as synergies — progress in one Goal triggers the advancement in another — and trade-offs — improvement in one Goal determines a worsening in another. Understanding this pattern of relationships is vital for policymakers to prepare efficient strategies to move and reach, given the urgency of the 2030 Agenda for Sustainable Development.

A first approach relies on the computation of positive and negative correlation coefficients between the indicators collected for a subset of countries for which data are available (e.g., Kroll *et al.*, 2019; Kostetckaia *et al.*, 2022; Zhao *et al.*, 2023). The sign of the correlation coefficient indicates whether two goals move in the same direction and foster each other (positive coefficient) or move in the opposite direction and progress in one goal might lead to a regression of the other (negative coefficient). Another class of studies (e.g., Pham-Truffert *et al.*, 2020; Allen *et al.*, 2019; Ehrensperger *et al.*, 2019; Griggs *et al.*, 2017) aims at understanding the causality between the goals using the information provided by experts. The experts have to evaluate the influence that progress in a target has on the progress of the other targets. The answers are then aggregated in a point scale (Nilsson et al., 2017) from negative to positive values, where positive values indicate synergies and negative values indicate trade-offs.

In this paper, we align with the studies based on correlation and propose a method that employs copulas to analyze the interactions between SDGs. While correlation and Kendall's tau measure a global association between variables, copulas allow

studying more specific types of dependence beyond linear association. To illustrate the method, we consider the SDGs related to health and well-being (Goal 3) and those related to biodiversity (Goals 6,13,14,15), whose links are well described theoretically. We select the corresponding SDG indicators and study the dependence among the targets of these goals, thereby deepening the insights of most of the current studies looking at the goal level. In line with other studies (e.g., Lusseau and Mancini, 2020), we study the dependence by distinguishing the countries into two groups based on their economies (lower and middle-lower vs middle-upper and high) to determine heterogeneities among countries.

The remainder of the paper is organized as follows: We describe the data in Section 2 and briefly introduce copulas in Section 3. In Section 4, we provide an illustrative example of using copulas to study relationships among the Goals. We conclude by discussing the results, limitations of the analysis, and future steps.

2. Data

We consider the Tier 1 class indicators as of 6 March 2024, defined by the IAEG-SDGs group.[3] An indicator belongs to the Tier I group if it "is conceptually clear, has an internationally established methodology and standards are available, and data are regularly produced by countries for at least 50 per cent of countries and of the population in every region where the indicator is relevant." (IAEG-SDGs).

The indicators are supposed to be collected yearly and made available for each country by the UN Statistic Divisions and the World Bank. We considered the most recent values dated to 2022. If the values in 2022 were not available, we imputed the missing data using the most recent value before 2022. For Goal 3, the proportion of missing data was 49.4% and reduced to 9.2% when imputing data using the value observed in 2019 and 2020 and to 5.2% when considering the value observed in 2015. For Goal 6, the proportion of missing data was 59.4% and reduced to 1.7% when imputing data using the value observed in 2020. For Goal 13, the 39.4% of missing data in 2022 was reduced to 19.1% and to 9.7% when considering the values in 2020 and 2017, respectively. For Goal 15, the 5.7% of missing data was reduced to 3.5% and to 1.6% when considering the values in 2020 and 2017, respectively.

The total number of indicators is 61, of which 20 were selected because they are available for at least 75% of the countries. All the indicators for the targets of Goal 14 did not meet the selection criterion; therefore, there are no indicators related to the use of oceans, seas, and marine resources. Table 1 describes the full list of indicators considered in the study.

[3] https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/

Table 1 – *Goals and indicators used in the study.*

Goal 3: Ensure healthy lives and promote well-being for all at all ages
3.1.1 Maternal mortality ratio
3.2.1 Under-5 mortality rate
3.2.2 Neonatal mortality rate
3.3.4 Hepatitis B incidence per 100,000 population
3.3.5 Number of people requiring interventions against neglected tropical diseases
3.4.1 Mortality rate attributed to cardiovascular disease, cancer, diabetes or chronic respiratory disease
3.4.2 Suicide mortality rate
3.9.3 Mortality rate attributed to unintentional poisoning
Goal 6: Ensure availability and sustainable management of water and sanitation for all
6.4.1 Change in water-use efficiency over time
6.4.2 Level of water stress: freshwater withdrawal as a proportion of available freshwater resources
6.5.1 Degree of integrated water resources management
6.6.1 Change in the extent of water-related ecosystems over time
Goal 13: Take urgent action to combat climate change and its impacts
13.1.1 Number of deaths, missing persons and directly affected persons attributed to disasters per 100,000 population
Goal 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainable manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss
15.1.1 Forest area as a proportion of total land area
15.1.2 Proportion of important sites for terrestrial and freshwater biodiversity that are covered by protected areas, by ecosystem type
15.2.1 Progress towards sustainable forest management
15.4.1 Coverage by protected areas of important sites for mountain biodiversity
15.5.1 Red List Index
15.6.1 Number of countries that have adopted legislative, administrative and policy frameworks to ensure fair and equitable sharing of benefits
15.8.1 Proportion of countries adopting relevant national legislation and adequately resourcing the prevention or control of invasive alien species

The indicators are based on a different scale, and the interpretation of the values sometimes goes in different directions. For most indicators, high values indicate a small development, while low values indicate a high development with respect to a target. We reverse the scale of the indicators whose interpretation did not adhere to

this direction by using the complement to 100 when the indicators represent percentages and the difference between the maximum value observed and the indicator's value when the unit of measure has no upper boundary. Finally, we normalized the indicators by using the min-max normalization so that we control for the range of values.

The following analysis focuses on the 111 countries for which the values for all the indicators was available. Among the selected countries, 43 are classified as lower or middle-lower and 58 as medium-upper or high economies countries according to the World Bank country classification by income level.[4]

3. Methods

We study the dependence between target indicators using bivariate copulas. A bivariate copula is a function that joins the bivariate distribution function to their marginal distributions. It describes the dependence structure existing across pairwise marginal random variables. Sklar's theorem (see Nelsen, 2013) shows that every bivariate/multivariate distribution can be defined via copula representation.

Let (X_1, X_2) be a bivariate random variable with marginal cumulative distribution functions $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$ and joint cumulative distribution function $F_{X_1,X_2}(x_1, x_2; \theta)$. Sklar's theorem affirms that it exists a copula function $C\left(F_{X_1}(x_1), F_{X_2}(x_2); \theta\right)$ with $C: I^2 \rightarrow I$ such that

$$F_{X_1,X_2}(x_1, x_2; \theta) = C\left(F_{X_1}(x_1), F_{X_2}(x_2); \theta\right), \;\; x_1 , x_2 \in \mathrm{I} \subseteq R \; .$$

Copula functions are helpful tools for handling multivariate continuous distributions with given univariate marginals (Nelsen, 2013). They are applied to describe the dependence structure between the marginal distributions of an arbitrary joint distribution. As proven by Sklar's theorem (Sklar, 1959), we can factorize an arbitrary joint distribution in the product of its marginal distributions and dependence structure captured by the copula distribution. Thus, applying copulas allows for separately modeling the marginals and the dependence structure.

By changing the copula function, we can construct new bivariate distributions with different dependence structures. The association parameter of the copula function indicates the strength of the dependence, which may also be different from the linear one that characterizes the multivariate normal distribution.

[4] https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-world-bank-classify-countries

Most of the measures of association characterizing the relationship between two variables X_1 and X_2 can be computed using copula functions. For example, the coefficient Kendall's τ is defined as

$$\tau(X_1, X_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 .$$

The equation above shows that Kendall's τ depends only on the underlying copula since it is invariant with respect to marginal distributions.

Similarly, tail dependence coefficients can also be defined using copulas. Tail dependence coefficients concern the level of dependence between more extreme values in the upper, lower, or both quadrant tails of a bivariate distribution.

Considering the probability of the joint occurrence of extremely small or large values, the upper and lower tail dependence coefficients are defined as

$$\lambda_L = \lim_{t \to 0^+} P\left(X_2 \leq F_2^{-1}(t) \middle| X_1 \leq F_1^{-1}(t)\right) = \lim_{t \to 0^+} \frac{C(t,t)}{t}$$

$$\lambda_U = \lim_{t \to 1^-} P\left(X_2 > F_2^{-1}(t) \middle| X_1 > F_1^{-1}(t)\right) = \lim_{t \to 1^-} \frac{1 - 2t + C(t,t)}{1 - t}$$

in case the limits exist.

The coefficients λ_L and λ_U take different values according to the copula chosen. For instance, the Gaussian and Frank copulas do not exhibit tail dependence, i.e., $\lambda_U = \lambda_L = 0$. The Student's T copula has symmetric tail dependence, i.e., $\lambda_U = \lambda_L$. The Clayton and Gumbel copulas are characterized only by lower λ_L or upper λ_U tail dependence.

These parameters λ_L and λ_U measure the dependence in the tails of the joint distribution, i.e. low/high values of one variable are associated with high/low values of the other one. They represent the probability that one variable is extreme given that the other is extreme too. The tail dependence parameters are directly associated to the parameters of some copula families. Further examination of copulas and measures of dependence can be found in Nelsen (2013) and Joe (1997).

In the following, we use the Clayton copula since we are interested in the left tail of the joint distribution of two indicators. The choice of the copula is based on the expectations that biodiversity and health (Griggs et al., 2017) are synergetic, meaning they work together in a way that an improvement in one goal implies an improvement in the other goal. This relationship manifests itself in low values of the indicators.

We represent the dependence structure implied by the Clayton Copula as a dependence graph, a graph in which the nodes are the indicators and the ties denote the presence and strength of the dependence as measured via a suitable copula-based tail coefficient describing the extreme of the co-movements of the dependence in the lower tail.

Finally, we utilize the dependence structure to cluster the indicators by applying the Louvain clustering method for community detection (Blondel *et al.*, 2008). This method allows us to extract non-overlapping sets of synergetic indicators. The Louvain method clusters the nodes in groups that are internally well-connected and weakly connected to other sets of nodes. The chosen node partition is generated by an algorithm that iteratively maximizes the modularity coefficient (Clauset *et al.*, 2004), and the best partition is the one with the maximum modularity.

4. Results

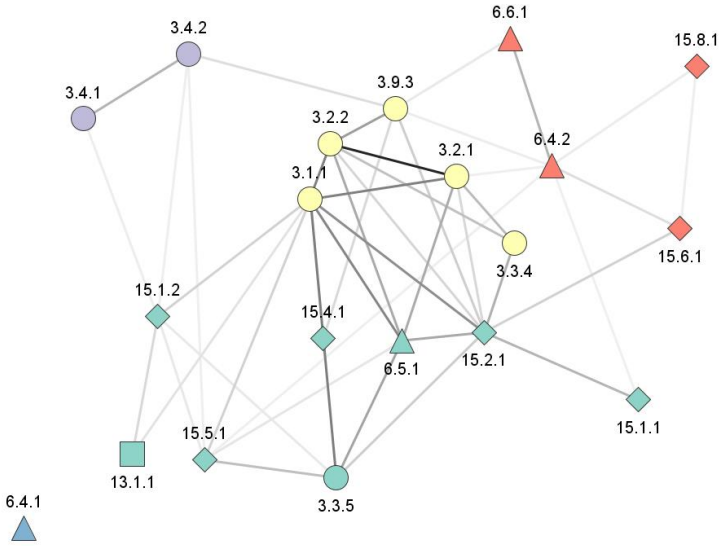
Figures 1 to 3 represent the dependence graphs. Nodes are the indicators. The node shape depicts the goal, and the color indicates the membership to the clusters obtained with the Louvain method.

We start by commenting on the dependence graph computed on all the 111 countries (Figure 1). The clustering identified 5 groups, each with its unique characteristics and significant implications for global development and sustainability.

The first group contains the indicators of target 6.4.1, which is independent from all the other targets. The second group (violet) includes the indicators of target 3.4, which is weakly connected to the other indicators. The third group (yellow) contains 5 indicators concerning health, specifically maternal and child mortality, hepatitis B incidence, and mortality due to unintentional poisoning. The fourth group (orange) has 4 indicators and relates to the legislative, administrative, and policy framework for sharing benefits, controlling invasive alien species, and managing water resources. The last group refers to forest preservation, management of water resources, and interventions against neglected tropical diseases.

Figure 2 shows the dependence graph for 58 countries classified as high and middle-upper economies. The clustering identified 6 groups. Among those, two groups include only one indicator, the one for targets 6.4.1 and 15.8.1, indicating weak dependences on all the other targets. The third group (orange) contains the indicators of target 3.4, which are also dependent on mountain biodiversity and interventions against neglected tropical diseases. The fourth group (yellow) includes

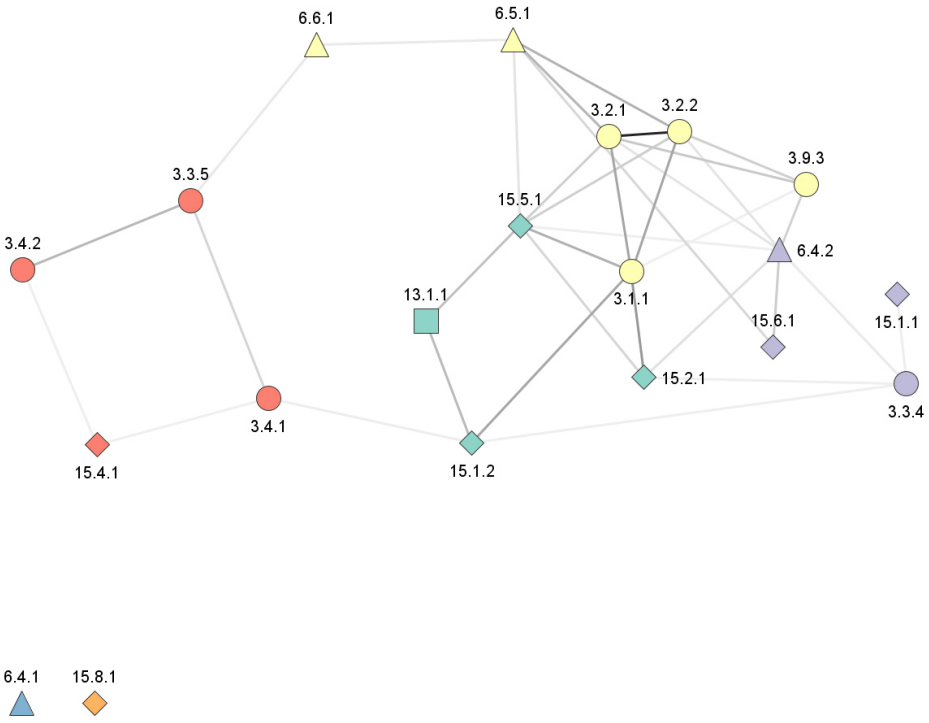
Figure 1 – *Dependence graphs for Goals 3, 6, 13, and 15 for the 111 countries for which the data are available. The node colour indicates the membership in the group, the shape the goal, the intensity of the link the dependence strength as measured by the Clayton copula.*



six indicators concerning health (specifically maternal and child mortality and the mortality due to unintentional poisoning) and the sustainable management of water and sanitation. The fifth group (green) has 4 indicators and relates to the presence of forests and deaths due to natural disasters. The last group (violet) has 4 indicators measuring the legislative, administrative, and policy framework for sharing benefits, forest management, water stress, and hepatitis B incidence.

Finally, Figure 3 reports the results for the 43 countries classified as low and middle-low economies. The first group (red) includes 8 indicators relating to the management of forests, changes in the water-related ecosystems, and the deaths for united poisoning. The second group (green) has 6 indicators all related to mortality, specifically maternal child, cardiovascular disease, cancer, diabetes and chronic respiratory disease, and suicide mortality rate. The third group (violet) contains 3 indicators concerning water resource management, water-use efficiency changes, and biodiversity. The last group (yellow) includes four indicators concerning hepatitis B incidence, interventions against neglected tropical diseases, deaths due to natural disasters, and water stress.

Figure 2 – *Dependence graphs for Goals 3, 6, 13, and 15 for the 58 middle-high and high economies countries for which the data are available. The node colour indicates the membership in the group, the shape the goal, the intensity of the link the dependence strength as measured by the Clayton copula.*

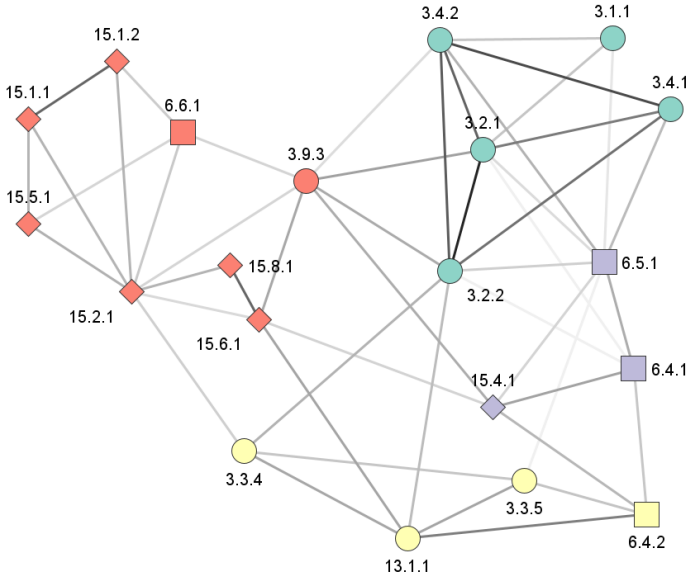


5. Discussion and conclusions

In this paper, we propose using copulas to deepen the analysis of trade-offs and synergies among the sustainable development goals and their targets. We illustrate the method by analyzing the interconnections among the targets of Goals 3, 6, 13, and 14 related to health and biodiversity. Given the documented synergies between the Goals, we delved into the dependencies, describing how an improvement in one goal implies an improvement in the other goal using the Clayton copula.

The clustering analysis indicates that targets related to maternal and child mortality and improvement in sanitation are strongly associated. Similarly, the forest and the water resources management show a strong lower tail dependence. Those results are in line with theoretical expectations (Griggs et al., 2017)

Figure 3 – *Dependence graphs for the Goals 3, 6, 13, and 15 for the 43 countries classified as low and middle-low economies for which the data are available. The node colour indicates the membership in the group, the shape the goal, the intensity of the link the dependence strength as measured by the Clayton copula.*



We performed the analysis on three groups, including all the 111 countries that did not present any missing data, the subset of the 58 countries classified as middle-upper and high economies, and the 43 countries classified as low and middle-low economies for which the data are available. The results are stable, though a few indicators with weak dependencies on the others were classified in different clusters across the three contingents of countries. The main difference is that the indicator of target 6.4.1 did not show any dependence on the other indicators when all the countries and those classified as middle-upper and high economies are considered. However, target 6.4.1 is associated with mountain biodiversity and the management of water resources when focusing on low and middle-low economy countries.

It is important to point out that our study is an illustration and the results are far from being complete and cannot be interpreted in a causal sense. Results are incomplete because of the high number of missing data and the focus on a limited set of targets and goals. Results cannot be interpreted in a causal sense since they only describe the dependencies among indicators and the targets they refer to, and

this dependence describes how improvement in one goal dimension is associated with improvements in other goal dimensions.

The next natural step is to extend the analysis to all seventeen goals, use different copulas to describe various types of dependencies, and complement the results on the correlation coefficients. This future research holds the promise of further enhancing our understanding of the complex interconnections among sustainable development goals and their targets. Results can guide policymakers and practitioners in their efforts toward sustainable development.

References

ALLEN C., METTERNICHT G., WIEDMANN T. 2019. Prioritising SDG targets: Assessing baselines, gaps and interlinkages, *Sustainability Science*, Vol. 14, pp. 421-438.

BOAR A., BASTIDA R., MARIMON F. 2020. A systematic literature review. Relationships between the sharing economy, sustainability and sustainable development goals, *Sustainability*, Vol. 12, No. 17, pp. 1-14.

BLONDEL V. D., GUILLAUME J. L., LAMBIOTTE R., LEFEBVRE E. 2008. Fast unfolding of communities in large networks, *Journal of statistical mechanics: theory and experiment*, Vol. 2008, No. 10, pp. 1-12.

EHRENSPERGER A., DE BREMOND A., PROVIOLI I., MESSERLI P. 2019. Land system science and the 2030 agenda: exploring knowledge that supports sustainability transformation, *Current Opinion in Environmental Sustainability*, Vol. 38, pp. 68-76.

GRIGGS D. J., SMITH M. S., ROCKSTRÖM J., ÖHMAN M. C., GAFFNEY, O., GLASER G., KANIE N., NOBLE I., STEFFEN W., SHYAMSUNDAR P. 2014. An integrated framework for sustainable development goals, *Ecology and society*, Vol. 19, No. 4, pp. 1-24.

GRIGGS D. J., NILSSON M., STEVANCE A., MCCOLLUM D. 2017. *A guide to SDG interactions: from science to implementation*. International Council for Science, Paris.

JOE H. 1997. *Multivariate models and multivariate dependence concepts*. CRC press.

KOSTETCKAIA M., HAMENTNER M. 2022. How sustainable development goals interlinkages influence European Union countries' progress towards the 2030 agenda, *Sustainable Development*, Vol. 30, No. 5, pp. 916-926.

KROLL C., WARCHOLD A., PRADHAN P. 2019. Sustainable Development Goals (SDGs): Are we successful in turning trade-offs into synergies? *Palgrave Communications*, Vol. 5, No. 1, pp. 1-11.

LE BLANC D. 2015. Towards integration at last? The sustainable development goals as a network of targets, *Sustainable Development*, Vol. 23, pp. 176–187.

LUSSEAU D., MANCINI F. 2019. Income-based variation in Sustainable Development Goal interaction networks, *Nature Sustainability*, Vol. 2, No. 3, pp. 242-247.

NELSEN R. B. 2013. *An introduction to copulas*. Springer.

NILSSON M, GRIGGS D, VISBECK M, RINGLER C. 2017. A framework for understanding sustainable development goal interactions. In GRIGGS D. J., NILSSON M., STEVANCE A., MCCOLLUM D. (Eds) *A Guide to SDG Interactions: From Science to Implementation*, International Council for Science: Paris, pp. 18-31.

PHAM-TRUFFERT M., METZ F., FISHER M., RUEFF H., MESSERLI P. 2020. Interactions among Sustainable Development Goals: Knowledge for identifying multipliers and virtuous cycles, *Sustainable development*, Vol. 28, No. 5, pp. 1236-1250.

REDCLIFT M. 1991. The multiple dimensions of sustainable development. *Geography*, Vol. 76, No. 1, pp. 36-42.

SKLAR A. 1959. Fonctions de Répartition à n Dimensions et Leurs Marges, *Annales de l'Institut Statistique de l'Université de Paris*, Vol. 8, No. 3, pp. 229-231.

UNITED NATIONS. 2015. Transforming our world: the 2030 Agenda for Sustainable Development. *United Nations: New York, NY, USA*.

ZHAO M., MENG Z., XU B., YU D., WU S., LI R. 2023. Prior goal setting of SDGs based on the interaction correlation, *Frontiers in Sustainability*, Vol. 4, pp. 1-8.

Viviana AMATI, Department of Statistics and Qualitative Methods, University of Milano- Bicocca, viviana.amati@unimib.it

Marta NAI RUSCONE, Department of Economics, University of Genoa, marta.nairuscone@unige.it

ASSESSING ECONOMIC RESILIENCE IN ITALY: A COMPARATIVE ANALYSIS OF SHOCKS AND SPATIAL DYNAMICS

Sheila Chapman, Vito Pipitone[1]

Abstract. This paper aims to assess the resilience of local economic units in Italy during two recent exogenous shocks: the financial crisis of 2008-2013 and the COVID-19 pandemic of 2019-2020. It examines resilience in terms of labour market performance and investigates the role of specialization, location and spillovers effects. It applies spatial analysis techniques to employment dynamics in Local Labour Market Areas to: (i) disentangle the role of specialization from that of location; (ii) identify the presence of homogeneous behavior patterns across units.

1. Introduction

The word "resilience" originates from the Latin verb *resilire*, that denotes the capacity of an agent to rebound following a shock or a disturbance.

Academic interest in the notion soared after the 2008 crisis. Notwithstanding its wide diffusion, there is no consensus on the exact definition of the concept. The most common ones range from engineering to ecological to adaptive resilience. As aptly noted by Compagnucci *et al.* (2022), by considering the ability/speed with which an economy recovers its initial equilibrium, engineering resilience refers to an approach *à la* Solow, according to which there exists a long-run growth path. Ecological resilience, instead, allows agents to reach a new equilibrium, given the magnitude of the shock. It studies a system's ability to adapt/evolve and follows an evolutionary, or Schumpeterian, approach. Within the evolutionary approach Martin (2012) and Martin and Sunley, (2015) develop the adaptive definition of resilience that focuses on the ability of an economy to withstand and/or recover from a shock by adapting its structure to the new conditions. From an empirical view point, resilience is often analysed with reference to two different phases, i.e. an area's ability to: (i) withstand shocks (resistance); and/or (ii) recover from them (recovery) (Martin, 2012).

Within economic geography, resilience focuses on the diverse abilities of territories to react to a downturn; by long, it is acknowledged as an important feature of territorial analysis. Even when a crisis hits all the units of a territory equally, its effects may spread unequally due to differences in local responses. Coming to the determinants of local resilience, the literature focuses on the sectoral structure of production as a key element that shapes the local reaction to a crisis, especially in the long run[2]. Highly specialized areas are generally thought to allow better

[1] Co-first author

[2] See, among others, Compagnucci et al. (2022), Lazzeretti and Oliva (2022), Lagravinese (2015).

resistance as they offer consolidated ties, shared routines, dedicated institutions and the like, while non specialized ones are generally more exposed to market forces, especially if they host new, innovative activities. However, diversified/unspecialized areas are also more sheltered from sector-specific crisis; being more flexible, they may grasp new opportunities and recover well in the longer run. Specialized ones, instead, could remain locked in traditional, often lagging, sectors (Boschma, 2015, Martini and Platania, 2019, Compagnucci *et al.*, 2022). What matters for resilience and long-run recovery is an appropriate mix of differentiation and specialization (Foray *et al.*, 2018), or production complexity (Hausmann *et al.*, 2013). In this framework the presence of technology-led, knowledge-intensive activities nested within the pre-existing structure is often found to be crucial for long run recovery. It allows agents to interact and cooperate, exchange ideas and develop innovative technologies, eventually adapting the local economy to the new external conditions.

This paper analyses the resilience of local economic units in Italy during and after two recent shocks: the 2008-2013 Great Crisis and the 2019-2020 COVID-19 pandemic. The two shocks are very different in nature. The first one, originated by the US credit boom and house price bubble, evolved into the sovereign debt crisis of 2011-13; it had both real and financial effects. The literature still debates its exogenous/endogenous character. In Italy as elsewhere, it caused the longest and deepest recession in peacetime. The second shock instead was entirely exogenous and had essentially short-lived effects. On this background, we investigate the role played by sectorial specialization and by location in local reactions to the crisis, with a focus on the well-known North-South divide. The latter is mirrored in the more diversified and advanced production structure of the Centre-North and in the far less diversified South, largely concentrated in low-technology traditional activities[3]. This suggests higher resistance in the South, but lower recovery post-crisis, and the opposite in the rest of the country. The paper tests this hypothesis. It focuses on the role of specialization and disentangles its effects from those of location. It addresses Local Labor Market Areas (LLMAs) in order to identify homogeneous local growth dynamics and spillover patterns for a rather fine territorial breakdown that differs from administrative boundaries and reflects only socio-economic factors[4].

The paper is organized as follows: paragraph 2 briefly describes the data and the methodology, providing some descriptive statistics; paragraph 3 presents the results while paragraph 4 summarizes and outlines some indication for further research.

2. Data and methodology

As in most of the literature, resilience is analysed with reference to employment[5]. Data are taken from ISTAT's (the Italian National Institute of Statistics) Labour Force Survey and refer

[3] See, among others, Chapman and Pipitone (2023), Chapman and Pipitone (2022), Asso et al. (2021).
[4] LLMAs are often identified as the best territorial breakdown to analyse local growth processes. For a discussion, see Martini and Platania (2019).
[5] The other alternative is value added, which is not currently available for Italian LLMAs.

to 610 LLMAs over 2008-2022[6]. For each LLMA we retrieve data on employment and on 17 specialization groups[7].

On the basis of employment dynamics in Italy we identify years 2008-13 and 2019-20 as crisis periods, in which we measure local units' ability to keep employment growth in line with the national rate (resistance)[8]. The periods between 2013-19 and 2020-22 are instead the years in which we expect recovery, defined as local units' capacity to outperform the country's average employment growth.

We measure resilience with reference to the indexes defined by Martin *et al.* (2016) for regional employment. As in Martini and Platania (2019), both resistance (y_{res}) and recovery (y_{rec}) indexes are modified in order to account for LLMAs.

$$y_{res} = \frac{(\Delta EMP_{llma}/EMP_{llma}) - (\Delta EMP_{nat}/EMP_{nat})}{|\Delta EMP_{nat}/EMP_{nat}|}$$

$$y_{rec} = \frac{(\Delta EMP_{llma}/EMP_{llma})}{(\Delta EMP_{nat}/EMP_{nat})}$$

($\Delta EMP_{llma}/EMP_{llma}$) and ($\Delta EMP_{nat}/EMP_{nat}$) are the percentage changes in employment respectively in the LLMA and in the country over a given period, measured as the difference between the first and last years of each period[9].

A preliminary picture of the territorial pattern of resistance/recovery in Italy's LLMAs may be gained from Figs. 1a and 1b.

Fig.1a maps resistance in the two crises, namely 2008-13 and 2019-20. At first sight, the map confirms a marked divide between the Centre-North and the South. However, behaviour changes a lot over time.

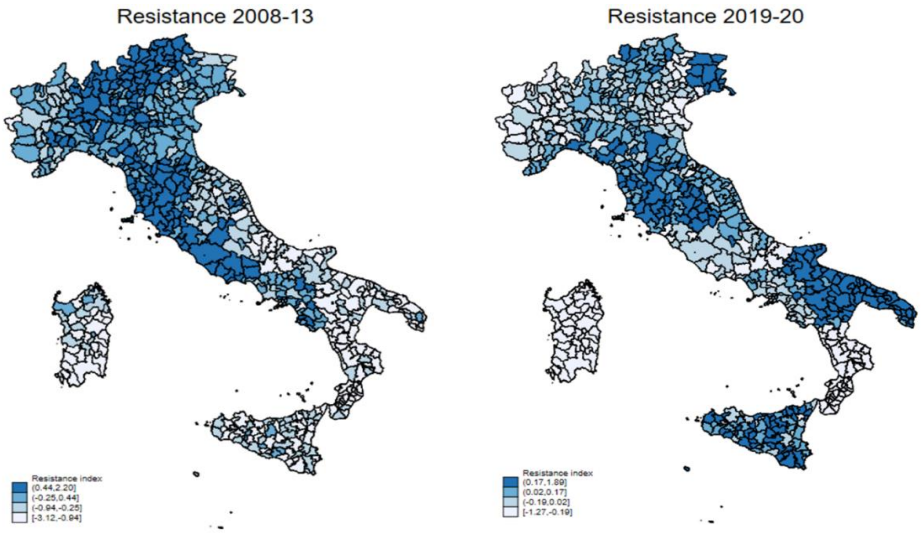
[6] LLMAs are self-contained areas defined on the basis of residents' commuting patterns for work. They are identified by ISTAT with reference to the 15th General Census of Population and Housing, using the new EURO methodology.

[7] These are: non-specialized LLMAs; non-manufacturing LLMAs (distinguished in: highly specialized urban, multi-specialized urban, non-specialized urban, port-oriented urban, Tourism, Agriculture); "Made in Italy" LLMAs (distinguished in Textile/clothing, Leather/hides, Machinery, Wood/furniture, Agro-Food, Jewelry/ eyewear/musical instruments); heavy manufacturing LLMAs (Means of transportation, Metal production/processing, Construction materials, Petrochemical/pharmaceutical).

[8] See, also Lagravinese, 2015, and Iacobucci and Perugini, 2021). Actually, after falling markedly in 2008-10, employment recovered somewhat in 2011, but fell further in 2012-13. It then grew until 2019; fell in 2020 and recovered later on. By 2022 it was slightly above the 2008 level.

[9] That is, respectively 2013 vs. 2008; 2019 vs. 2013; 2020 vs. 2019 and 2022 vs. 2020.

Figure 1a – *Resistance patterns in Italian LLMAs (2008-13 and 2019-20).*



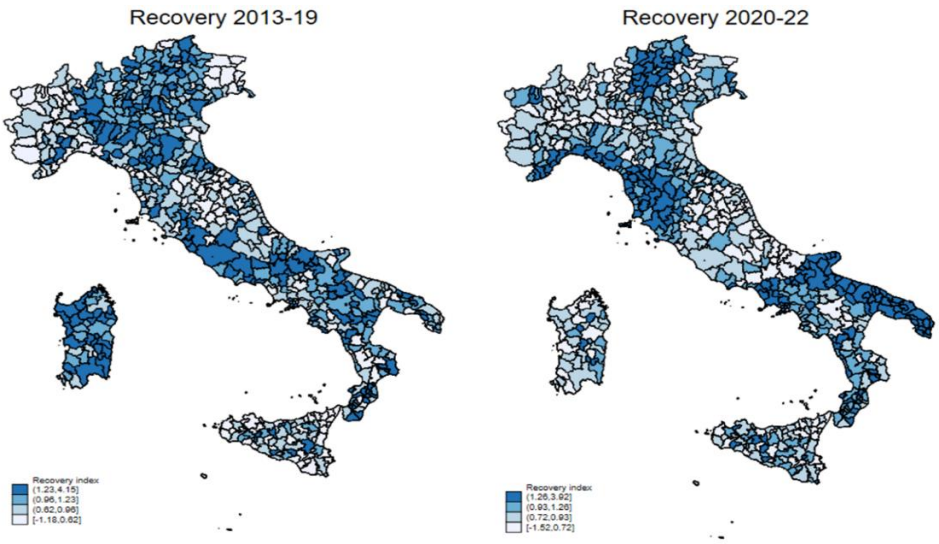
Source: our elaboration

Contrasting the literature, in 2008-13 the North and Centre (especially the Tyrrhenian coast) resist far better than the South[10]. Instead reactions to the 2019-20 shock are closer to the literature's indications, as many units in the North and in the Centre are badly hit, while the South generally resists better[11].

Fig. 1.b focuses on recovery patterns (respectively 2013-19 and 2020-22). In both periods the North-South divide is less evident. In 2013-19 recovery does not show any clear-cut territorial pattern, as many LLMAs throughout the country perform well (or badly). After the pandemic, recovery patterns change a lot: in the first place, response shows an evident regional distribution. Units perform well, or badly, depending on what occurs in the other units in the same (NUTS2) region, suggesting strong spillover effects. In the second place, and quite unexpectedly, while most units of the Centre-North (with some exception) lag behind, many ones in the South show high recovery. This again contrasts the indications of the literature that predict quicker and better recovery for more diversified, advanced territories such as those of the Centre-North[12].

[10] Martini and Platania (2019) reach similar results.
[11] Especially in the areas that refer to the (NUTS2) regions of Abruzzo, Campania, Puglia, Basilicata and Sicily. This does not apply to units in the areas of Molise, Calabria and Sardinia.
[12] As suggested by an anonymous referee, this can be explained by significant post-pandemic price pressure on manufacturing activity strongly affecting energy-intensive LLMAs in the Centre-North.

Figure 1b – *Recovery patterns in Italian LLMAs (2013-19 and 2020-22).*



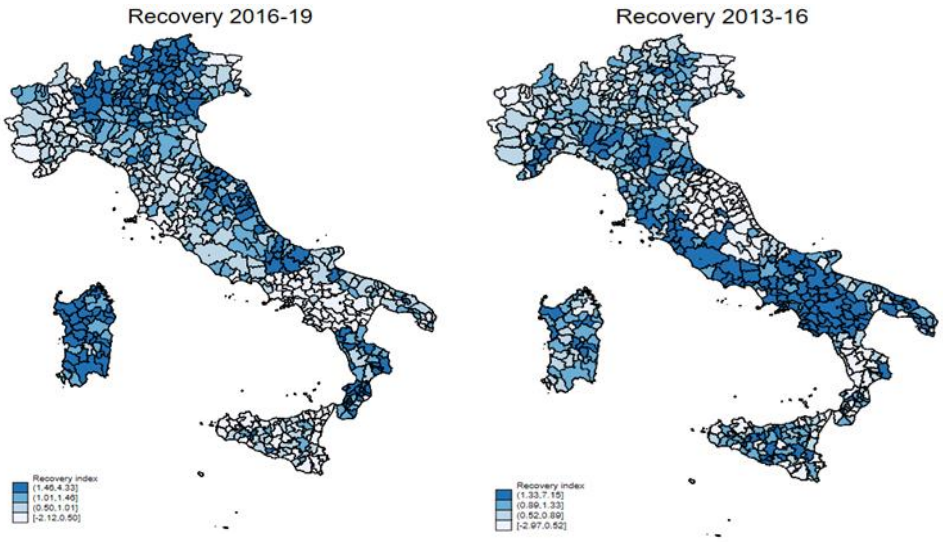
Source: our elaboration

On the other hand, Figs. 1a and 1b also suggest that resistance patterns carry on in the short run and shape the early phases of recovery[13]. This hypothesis is preliminarily investigated by splitting the 2013-19 recovery period into two phases: an initial one running from 2013 to 2016 and a later one starting in 2016, when resistance effects, if any, could be weaker. The breakdown is shown in Figure. 2. In the short run the map on the left (2013-16) broadly confirms similarity between resistance and recovery patterns, while in the longer run the map on the right (2016-19) shows that recovery largely follows different trajectories.

We conclude that the descriptive analysis strongly suggests that resilience patterns of Italian LLMAs confirm the literature's indications only in part. It also provides further hypothesis that are tested by means of spatial estimation techniques.

13 As expected, the correlation between resistance and recovery indexes for 2008-19 is negative, indicating that high resisting units recover weakly while low resisting ones display higher recovery. However, the correlation turns positive for 2019-22, suggesting that in this case recovery is largely dictated by resistance performance.

Figure 2 – *Recovery patterns in Italian LLMAs (2013-16 and 2016-19).*



Source: our elaboration

3. Estimates and results

We proceed to estimate the following simple equation for resistance and recovery indexes, conditional on (i) the specialization of each LLMA; and (ii) resistance/recovery in neighboring units, in order to capture spatial spillovers:

$$y = \beta_0 + \beta_1 X + \rho W y + \varepsilon \quad (1)$$

where $\boldsymbol{y}$ is a 610 x 1 vector of the resistance/recovery index for each LLMA; $\boldsymbol{X}$ is a 610 x 17 matrix of the LLMAs' 17 specialization dummies, $\boldsymbol{Wy}$ is the spatially lagged dependent variable y, $\boldsymbol{\varepsilon}$ is the vector of 610 x 1 vector of (normally distributed) errors. The response parameters are $\boldsymbol{\beta_0}$ (the constant), $\boldsymbol{\beta_1}$ and $\boldsymbol{\rho}$.

Equation (1) is estimated for each sub-period. Following the indications suggested by the descriptive statistics, the 2013-19 recovery is also split into two sub-periods. Results are shown in Tables 1 (2013-19) and 2 (2019-22).

Table 1 shows that during the Great Recession (2008-13) sectorial specialization matters for resistance: most sectors (14 out of 16), both in services and in manufacturing, perform significantly better than non-specialized ones (the reference group, omitted to avoid multicollinearity). Instead, specialization seems much less important for recovery; over the whole post-Recession period (2013-19), only two sectors -tourism and agriculture- significantly outperform non-specialized LLMAs[14].

[14] One –wood and furniture- significantly underperforms them.

The temporal breakdown in two sub-periods adds some detail: in the first place, specialization proves to be ineffective in the short run (2013-16). When it is significant, it determines underperformance. This occurs for four categories, three of which belong to manufacturing. However specialization gains importance in the mid-term (2016-19), when no manufacturing sector is (any more) significantly below the reference group[15] and two non-manufacturing ones (tourism and agriculture) are above it. This also suggests that recovery in manufacturing LLMAs is slower than in urban service-oriented ones.

Results change deeply for the COVID-19 shock (Table 2), when specialization loses importance: it has significant (negative) effects on resistance only for 3 sectors.

In the early recovery (2020-22) it actually appears as a drawback, as 9 out of 10 manufacturing categories and 2 out of 5 urban, service-oriented ones perform significantly worse than non-specialized ones. In other terms, results suggest that while specialization did help LLMAs to resist during the 2008-13 crisis, it was less important for the following recovery, especially in the short run. Specialization patterns were even less important in the 2019-22 crisis: they were mostly irrelevant for resistance and actually had a negative impact on recovery, at least in the short term. This has important implications; given that in Italy non-specialized units are entirely located in the South, this result ultimately helps understand the relatively good performance of many Southern LLMAs in 2020-22 (see Figure 1b)[16].

In both episodes spatial effects are positive, significant, and growing over time[17]. This is especially true for resistance. However, a well-known problem of models with spatially lagged endogenous variables as equation (1) is the endogeneity of spatial dependence tha does not allow to interpret coefficients as simple partial derivatives[18]. Hence a change in an explanatory variable in any unit generates global spillovers, implying that it changes the unit's own dependent variable and also the dependent variable of all other units which, in turn, feed back into the initial unit.

[15] Except Means of transportation for which, however, low recovery could be linked to long-standing sectoral issues.
[16] Martini and Platania (2019) reach similar results for the 2008 shock. They find that over 2013-17 the non-specialized LLMAs of the South recover more than the specialized ones of the Centre-North.
[17] Wald's test points to highly significant spatial terms in all periods (including sub-periods).
[18] See, among others, Golgher and Voss (2015).

Table 1 – *Spatial Autoregressive Model Estimates: Resistance (2008-13) and Recovery (2013-19; 2013-16 and 2016-19).*

VARIABLES	Resistance	Recovery	Recovery	Recovery
LLMA specialization	**2008-13**	**2013-19**	**2013-16**	**2016-19**
Urban: highly specialized	**0.540****	0.202	0.139	0.100
	(0.260)	*(0.224)*	*(0.370)*	*(0.250)*
Urban: multi-specialized	**0.386*****	-0.140	**-0.326****	-0.116
	(0.113)	*(0.097)*	*(0.160)*	*(0.109)*
Urban: port-oriented	**0.245***	-0.118	-0.093	-0.166
	(0.141)	*(0.121)*	*(0.201)*	*(0.135)*
Urban: non specialized	**0.188***	-0.0265	-0.052	-0.060
	(0.111)	*(0.096)*	*(0.159)*	*(0.107)*
Tourism	**0.419*****	**0.130***	0.130	**0.140***
	(0.085)	*(0.071)*	*(0.117)*	*(0.079)*
Agriculture	0.0383	**0.146***	0.009	**0.180***
	(0.098)	*(0.084)*	*(0.140)*	*(0.094)*
Textile/clothing	**0.256****	-0.0416	-0.223	0.072
	(0.111)	*(0.095)*	*(0.157)*	*(0.106)*
Leather/hides	**0.253****	-0.078	-0.231	0.017
	(0.126)	*(0.108)*	*(0.179)*	*(0.121)*
Machinery	**0.304*****	-0.0365	-0.247	0.041
	(0.112)	*(0.095)*	*(0.157)*	*(0.106)*
Wood/furniture	**0.209***	**-0.195****	**-0.504****	0.050
	(0.116)	*(0.099)*	*(0.165)*	*(0.111)*
Food	**0.347*****	0.00639	-0.042	-0.009
	(0.095)	*(0.081)*	*(0.135)*	*(0.009)*
Jewel./eyewear/music. instr.	**0.393****	-0.0715	-0.226	-0.027
	(0.188)	*(0.162)*	*(0.267)*	*(0.181)*
Means of transportation	0.0302	-0.151	-0.138	**-0.269***
	(0.156)	*(0.135)*	*(0.223)*	*(0.150)*
Metal production/processing	**0.333*****	-0.0587	**-0.279***	0.032
	(0.120)	*(0.102)*	*(0.169)*	*(0.114)*
Construction materials	**0.467*****	-0.0880	-0.114	-0.139
	(0.149)	*(0.127)*	*(0.211)*	*(0.142)*
Petrolchemical/pharm.	**0.402*****	-0.169	**-0.304***	-0.114
	(0.129)	*(0.110)*	*(0.182)*	*(0.123)*
Wy	**0.821*****	**0.405*****	**0.590*****	**0.680*****
	(0.035)	*(0.049)*	*(0.052)*	*(0.041)*
Constant	**-0.345*****	**0.662*****	**0.416*****	**0.416*****
	(0.0583)	*(0.0596)*	*(0.061)*	*(0.061)*

610 observations; robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2 – *Spatial Autoregressive Model Estimates: Resistance (2019-20) and Recovery (2020-22).*

VARIABLES	Resistance	Recovery
LLMA specialization	**2019-20**	**2020-22**
Urban: highly specialized	0.123	-0.324
	(0.098)	*(0.217)*
Urban: multi-specialized	-0.046	**-0.306*** **
	(0.042)	*(0.094)*
Urban: port-oriented urban	-0.015	0.057
	(0.053)	*(0.118)*
Urban: non specialized	**-0.100** **	**-0.273** **
	(0.042)	*(0.093)*
Tourism	-0.022	-0.074
	(0.031)	*(0.069)*
Agriculture	**-0.146*** **	-0.117
	(0.037)	*(0.082)*
Textile/clothing	-0.041	**-0.252***
	(0.041)	*(0.0917)*
Leather/hides	-0.067	-0.168
	(0.047)	*(0.105)*
Machinery	-0.038	**-0.305*** **
	(0.041)	*(0.092)*
Wood and furniture	-0.016	**-0.207** **
	(0.043)	*(0.096)*
Food	0.048	**-0.132***
	(0.036)	*(0.079)*
Jewelry/eyewear/musical instruments	**-0.140** **	**-0.306***
	(0.071)	*(0.156)*
Means of transportation	-0.046	**-0.553*** **
	(0.059)	*(0.130)*
Metal production/processing	-0.031	**-0.279** **
	(0.045)	*(0.0987)*
Construction materials	-0.069	**-0.229***
	(0.056)	*(0.123)*
Petrolchemical/pharmaceutical	-0.055	**-0.334** **
	(0.048)	*(0.107)*
Wy	**0.871*** **	**0.457*** **
	(0.034)	*(0.060)*
Constant	0.027	**0.827*** **
	(0.020)	*(0.000)*

610 observations; robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

To investigate the spatial dimension of equation (1) we split the total effect of each specialization group into two components: a direct effect, that measures the impact on each unit's resistance/recovery (plus feedbacks) and an indirect one that accounts for spatial spillovers, i.e. it measures the effect on the dependent variables of all other LLMAs. Results are shown in Tabs. 3 and 4.

Table 3 – *Direct and indirect effects of specialization on resistance (2008-13) and recovery (2013-19).*

VARIABLES	**Resistance (2008-13)**		**Recovery (2013-19)**	
LLMA specialization	dir. effect	indir. effect	dir.effect	indir. effect
Urban: highly specialized	0.623**	1.320**	0.178	0.082
Urban: multi-specialized	0.445***	0.944***	-0.161	-0.075
Urban: port-oriented urban	0.283*	0.599*	-0.136	-0.063
Urban: non specialized	0.217*	0.461	-0.036	-0.017
Tourism	0.484***	1.025***	0.106*	0.049*
Agriculture	0.044	0.094	0.147*	0.068
Textile/clothing	0.295**	0.625**	-0.059	-0.027
Leather/hides	0.292**	0.619**	-0.097	-0.045
Machinery	0.351***	0.744***	-0.059	-0.027
Wood/furniture	0.241*	0.510*	-0.216**	-0.099**
Food	0.400***	0.848***	-0.010	-0.005
Jewel./eyew./music.instr.	0.453**	0.961**	-0.096	-0.044
Means of transportation	0.035	0.074	-0.162	-0.075
Metal prod/processing	0.384***	0.815***	-0.084	-0.039
Construction materials	0.538***	1.141***	-0.113	-0.052
Petrolchemical/pharm.	0.463***	0.982***	-0.173	-0.081

Table 4 – *Direct and indirect effects of specialization on resistance (2019-20) and recovery (2020-22).*

VARIABLES	**Resistance (2019-20)**		**Recovery (2020-22)**	
LLMA specialization	dir. effect	indir. effect	dir.effect	indir. effect
Urban: highly specialized	0.147	0.399	-0.379*	-0.184*
Urban: multi-specialized	-0.055	-0.149	-0.290***	-0.141***
Urban: port-oriented urban	-0.018	-0.048	-0.064	-0.031
Urban: non specialized	-0.120**	-0.325**	-0.204**	-0.099**
Tourism	-0.026	-0.070	-0.068	-0.033
Agriculture	-0.174***	-0.474***	0.050	0.024
Textile/clothing	-0.048	-0.131	-0.242***	-0.112**
Leather/hides	-0.078	-0.217	-0.163	-0.079
Machinery	-0.045	-0.122	-0.289***	-0.141***
Wood/furniture	-0.019	-0.052	-0.233**	-0.113**
Food	0.057	0.155	-0.170**	-0.083**
Jewel./eyew./music.instr.	-0.167**	0.454*	-0.238	-0.116
Means of transportation	-0.055	-0.150	-0.527***	-0.256***
Metal prod/processing	-0.037	-0.101	-0.274***	-0.133**
Construction materials	-0.083	-0.224	-0.214*	-0.104*
Petrolchemical/pharm.	-0.066	-0.179	-0.312***	-0.152***

Starting from the Great Recession, Table 3 shows that specialization –which, as argued above, matters a lot for resistance- produces average indirect effects, i.e. spatial spillovers, double in size with respect to own effects. In other terms, it favours resistance in neighbouring LLMAs more than it does in the initial one. Instead specialization has very weak spatial effects

for recovery. In the few cases in which it is significant -tourism, agriculture and (with a negative impact) wood and furniture- the own effect counts more than the indirect one.

As said, the results for the COVID episode differ. Table 4 shows that specialization determines significant (lower) resistance in only a few cases, namely in unspecialized urban LLMAs and in LLMAs specialized in agriculture and in jewelry, eyewear and musical instruments[19]. Indirect effects are again bigger in value than own ones. In recovery (2020-22) specialization has both negative direct and indirect effects, but own effects are far bigger in value (on average more or less double the direct ones).

In other terms, Tables 3 and 4 show that specialization creates big territorial spillovers only during resistance. When recovery occurs own effects prevail. This is especially evident for manufacturing in 2019-22.

4. Conclusion

This paper analyses the resilience patterns of Italian LLMAs over two recent, wide-sweeping, exogenous shocks: the Great Recession and the one tied to the COVID-19 pandemic. It highlights a number of points. First, response patterns change significantly over time; most units do not perform uniformly over the two episodes and in their aftermaths. Moreover, we find that, at least in the short run, recovery patterns largely overlap with resistance ones and that only over time a different trajectory takes shape. This suggests that resistance performance may have an impact on the early years that follow the crisis. Coming to the spatial dimension of the analysis, our results confirm the presence of a divide between the Centre-North and the South. Over time, a rather clear tendency towards the "regionalization" of response may be traced, as LLMAs located within a same region tend to perform uniformly. Also, spatial effects are strong and significant, especially for resistance. In other terms, resistance spreads spatially and has an impact on territories' performance. Instead recovery, if any, is mostly an individual matter.

Finally, and more important, we find that resilience patterns in Italy's LLMAs reflect the literature's indications only in the Great Recession, when specialized units resist well and recover with difficulty. This confirms that in that case traditional, less diversified structures offered a better shelter in a crisis but provided lower opportunities for recovery and often led to a lock-in. However, during the pandemic crisis, the pattern changed radically: it was the non-specialized units –all located in the South- that performed better with respect to the specialized ones of the Centre-North, suggesting that in this case non-specialized/diversified economic systems, such as those prevailing in the South, provided the flexibility/adaptability that allowed resistance and early recovery. This is an interesting result that needs to be investigated; we leave it to further research.

[19] The sector shows a significant positive indirect effect alongside a negative direct one, implying that this type of specialization favours resistance in neighbouring LLMAs, but not at home.

References

ASSO P.F., NEROZZI S., PIPITONE V., 2021. Italy's North-South Divide Yet Again: Some New Evidence from Firms, Sectors and Territories. *Rivista Internazionale di Scienze Sociali*, Vol. 2, pp. 91-124.

BOSCHMA R., 2015. Proximity and Innovation: A Critical Assessment, *Regional Studies*, Vol. 39, No. 19, pp. 61-74.

CHAPMAN S., PIPITONE V., 2023. The productivity of Italian firms: a spatial analysis. *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXXVII, No. 4, pp. 145-56.

CHAPMAN S., PIPITONE V., 2022. Total Factor Productivity in Italian Manufacturing: Does Location Matter? *Rivista Italiana di Economia, Statistica ed Economia*, Vol. LXXVI, No. 3, pp. 155-66.

COMPAGNUCCI F., GENTILI A., VALENTINI E., GALLEGATI M., 2022. Asymmetric response to shocks: the role of structural change on resilience of the Euro area regions, *Applied Economics*, Vol. 54, No. 37, pp. 4324-55.

FORAY D., MORGAN K., RADOSEVIC S., 2018. *The role of smart specialization in the EU research and innovation policy landscape*, European Commission.

GOLGHER A.B., VOSS P.R., 2016. How to Interpret the Coefficients of Spatial Models: Spillovers, Direct and Indirect Effects, Spatial Demography, Vol. 4, pp.175-205.

HAUSMANN R., HIDALGO C.A., BUSTOS S., COSCIA M., SIMEOS A., YILDIRIM M.A., 2013. *The Atlas of Economic Complexity: Mapping Paths to Property*, MIT Press.

IACOBUCCI D., PERUGINI F., 2021. Entrepreneurial ecosystems and economic resilience at local level, *Entrepreneurship and Regional Development*, pp. 1-28.

LAGRAVINESE R., 2015. Economic crisis and rising gaps North-South: Evidence from the Italian regions, *Cambridge Journal of Regions, Economy and Society*, Vol. 8, No. 2, pp. 331-42.

LAZZERETTI L., OLIVA S., INNOCENTI N., 2022. Unfolding Smart Specialisation for Regional Economic Resilience: the role of Industrial Structure, *Investigaciones Regionales*, Vol. 2022/3, No. 54, pp. 5-25.

MARTIN R., 2012. Regional economic Resilience, Hysteresis and Recessionary Shocks, *Journal of Economic Geography*, Vol. 12, No.1, pp.1-32.

MARTIN R., SUNLEY P. 2014. On the Notion of Regional Economic Resilience: Conceptualization and Explanation, *Journal of Economic Geography*, Vol. 15, No.1, pp.1-42.

MARTINI B., PLATANIA M., 2019. Analyzing Resilience in Local Labor Market Areas: A Spatial Analysis for the Case of Italy, *Applied Economics Quarterly*, Vol. 65, No. 3, pp. 189-208.

Sheila A. CHAPMAN, LUMSA University, chapman@lumsa.it

Vito PIPITONE, CNR-ISMed Palermo & LUMSA University, vito.pipitone@cnr.it

SKILLS IN THE ITALIAN LABOUR MARKET[1]

Barbara Boschetto, Elisa Marzilli, Silvia Montecolle, Alessia Sabbatini

Abstract. The skills of the workforce play a strategic role in the economic and social progress of a country. In this context, Italy faces a disadvantage compared to other developed countries. Italian workers, in particular, are noted to have lower levels of education and qualification (Istat, 2022). Additionally, Italy performs weakly in international rankings with regard to the share of human capital employed in scientific or technical fields (European HRST indicator; Istat, 2023). There is also an evident challenge in aligning skills supply with demand, leading to a significant proportion of workers being either underqualified or overqualified for their roles.
In this context, the data of the 2022 ad hoc module of the Labour Force Survey conducted by Istat, which focuses on the skills used in their job by individuals, can be analysed. Respondents were asked to indicate, based on their perception, the time allocated to specific activities, serving as a proxy measure for possessing particular skills.
The investigated skills are related to digital activities, reading, calculate, hard physical work, finger dexterity, internal and external communication and training. In addition, the study explores the way work is performed, observing the degree of autonomy in influencing the way work is carried out, as well as the prevalence of repetitive tasks or tasks ruled by rigid procedures.
The proposed contribution aims to analyse data from the module for employed people, using also a multiple correspondence analysis to explore the connections between the main skill-related variables and other employment characteristics.

1 Introduction

In the era of knowledge, the full utilization of qualified skills becomes a central factor in the economic and social development of a country. Monitoring the skills possessed by the workforce and matching supply and demand of skills represent useful indicators to assess the health of the labour market.

On this front, Italy finds itself at a disadvantage compared to other developed countries. Specifically, Italian workers are reported to be less educated (Istat, 2022)

[1] This article is a joint effort by the authors. In particular, Barbara Boschetto edited paragraphs 5 and 6, Elisa Marzilli paragraphs 1 and 2, Silvia Montecolle paragraph 4, Alessia Sabbatini paragraph 3.

and less qualified. Besides Italy ranks among the lowest in the international standings for indicator HRST (Human resources in Science & Technology), the proportion of human capital employed in scientific or technical professions – those falling under major groups 2 and 3 of the International Standard Classification of Occupations – ISCO-08 (Istat, 2023).

Furthermore, there is often a mismatch between supply and demand of skills, meaning a significant portion of workers is either under-qualified or over-qualified for the specific skills required in their jobs (Galletti F., Gualdi F., 2017). At the same time on the business front, demand for qualified skills is low, indicating a lack of investment in human capital (Brunello & Wruuck, 2021; OCSE, 2018).

Investing in skill development and aligning supply with demand boosts business productivity and improves individual well-being through better job satisfaction and wages. In addition, a robust skills assessment system is essential for monitoring and enhance the labour market, as outlined in the European Skills Agenda for sustainable competitiveness, social fairness, and resilience (European Commission, 2022).

The debate on this topic includes the 2022 ad hoc module of the Labour Force Survey conducted by Istat, dedicated to skills applied in one's job. This module consists of additional questions beyond the standard part of the survey, aimed at exploring a specific theme and administered simultaneously across all European countries (Istat, 2024).

The objective of this article is to analyse the concentration or absence of specific skills among the employed population in Italy. Data from the labour force survey allow to highlight differences that emerge across key socio-demographic variables such as gender, educational attainment, citizenship, as well as variables closely related to work such as occupation.

2. Data and methods

The ad hoc module on *Job skills* was included in the Labour Force Survey conducted in 2022. It was administered to persons aged 16 to 74 years old, whether employed or not employed for less than two years. Participants were asked to indicate the time spent on a series of activities, referring to their usual situation in their main job.

Specifically, it involves information on the amount of time spent using digital tools, reading technical documents, performing complex calculations, doing physically demanding activities or tasks requiring dexterity and precision, participating in communicative and training-related interactions, and aspects defining the work approach such as autonomy level, task repetitiveness and standardization of tasks (follow strictly defined procedures).

The measurement of time spent on performing a specific activity has been used as a proxy for a specific skill. The ability to perform certain activities has therefore been observed from the perspective of output, or what is actually practiced in the course of one's work.

The skills addressed in the module are categorized into thematic areas or domains. For instance, reading and calculation skills fall under the broader area of cognitive abilities, which are processes through which an individual registers, retains, retrieves, and uses information. Similarly, physical strength and dexterity fall under the domain of manual skills, highlighting abilities related to physical tasks and precise coordination of movements (Table 1).

For all questions in the module, respondents' perceptions were captured, meaning they chose the response option they considered most appropriate, using a five-point scale[2]. The degree of usage was then grouped into three levels: high, low, and absent[3]. Each skill can be practiced either synchronously or asynchronously with respect to another.

The analysis of skills, focusing on employed persons, examines both socio-demographic characteristics (age, gender, educational attainment, citizenship) and the occupation performed, using the Classification of Occupations (CP2011)[4].

Following an initial descriptive analysis, the study further explores the relationships between the modes of the main variables using Multiple Correspondence Analysis (MCA). Finally, the findings concerning the Italian context are compared with those of other European countries.

[2] Specifically, for all questions asking "How much time did you spend doing ...?", the response options were: a) all or most of the work time; b) half of the work time or slightly more ; c) some of the work time; d) little of work time; e) none of the work time. For questions asking "To what extent can you...?", the response options were: a) to a very large extent; b) to a large extent; c) to some extent; d) to little extent; e) no extent.

[3] The "high" level groups together the responses: a) all or most of the work time and b) half of the work time or slightly more, as well as the responses: a) to a very large extent and b) to a large extent.
The "low" level groups together the responses: c) some of the work time and d) little of work time, as well as the responses: c) to some extent and d) to little extent.
The "absent" level corresponds to the response: e) none of the work time or e) no extent.

[4] The Istat Classification of Occupations (CP), which is linked to the International Classification of Occupations ISCO-08, is the tool for classifying labour market occupations into specific occupational groupings, which is useful for the communication, dissemination and integration of statistical and administrative data on occupations, while ensuring comparability at international level. In this work, focused on the year 2022, the CP2011 Classification of Occupation is used. Since 2023, Istat has adopted the CP2021, the result of a revision of the previous version (CP2011) and further alignment with ISCO-08.

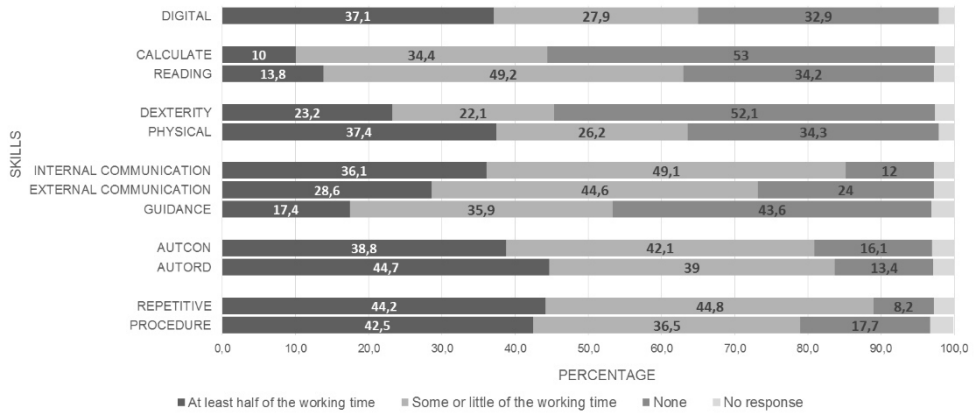
Table 1 - *The professional skills of the employed*

DIGITAL SKILLS	
DIGITAL	time spent on working on digital devices, i.e. use of digital technologies: computer, tablet, phablet or smartphone.
COGNITIVE SKILLS	
READING	time spent on reading work-related manuals and technical documents, i.e. documents that require specific skills or knowledge to be understood (technical complaints, acts, contracts, technical guidelines, etc).
CALCULATE	time spent on doing relatively complex calculations in main or last job, i.e. manipulation and transformation of numeric information.
MANUAL SKILLS	
PHYSICAL	time spent on doing hard physical work, i.e. manual tasks that primarily require energy and strength or intense muscular power.
DEXTERITY	time spent on tasks involving finger dexterity, i.e. ability to make precisely coordinated movements that implies a certain extent of accuracy (coordination of small muscles, synchronization of hands and fingers, with the eyes
SOCIAL SKILLS	
INTERNAL COMMUNICATION	time spent on interacting with people from the same enterprise or organization, i.e. verbal communication, situation where two or more people communicate with each other simultaneously.
EXTERNAL COMMUNICATION	time spent on interacting with people from outside the enterprise or organisation i.e. verbal communication, situation where two or more people communicate with each other simultaneously.
GUIDANCE	time spent on advising, training or teaching other people in formal and informal setting and on any topic, situation where two or more people communicate with each other simultaneously
AUTONOMY SKILLS	
AUTCON	degree of autonomy in influencing the contents of the job, i.e. how far the respondent can influence the way work is carried out in choosing contents.
AUTORD	degree of autonomy in influencing the order of the tasks, i.e. how far the respondent can influence the way work is carried out in term of sequence of work tasks.
MANAGEMENT SKILLS	
REPETITIVE	repetitiveness of tasks, i.e. this respondent implements similar tasks in the same way (work task that is implemented frequently without any variation/change/adaptation).
PROCEDURE	tasks precisely described by strict procedures; it provides a measure of standardization of tasks in different types of jobs (strict procedures are intended as rules specifying the timing and order of actions, the method to be used to perform a task and the use and communication of its results).

3. Skills and socio-demographic characteristics of the employed

Regarding digital skills, it emerges that 37.1% of employed persons use digital devices for at least half of their work time, while 32.9% never use such devices (Figure 1[5]). Those most likely to extensively use digital tools for work are persons aged 30-44 while the 15-29 age group has the highest percentage of those who do not use this ability (36.5%). Besides, the percentage of women using digital skills for at least half of their working time is significantly higher than that of men (42.1%, compared to 33.4%).

Figure 1- *Employed persons by time spent on different skills in their job. Year 2022 (age group 15-74, percentage of total employment).*



Source: Istat, Labour Force Survey

Cognitive skills in reading and calculate are used for more than half of the work time by one in five employed, with no significant differences by gender or age. Among the various age groups, adults aged 30 to 44 are the most likely to use reading and calculate skills, with 14.5% and 11.5%, respectively, spending at least half of their working time on these tasks. In contrast, these skills are less used by younger (16-29 years) and those over 60. About 39% of younger workers and nearly 37% of those aged 60 and over do not use reading at work, while around 56% of both age groups do not use calculate.

Manual skills are assessed based on intensity (physical) and precision (dexterity). Regarding physical, 37.4% of employed persons use this skill for half or more of their work time; while dexterity is applied by 23.2%. These skills are more

[5] For further details, see also https://www.istat.it/wp-content/uploads/2024/06/Stat_Focus_Competenze-professionali.pdf

prevalent among men: 41.6% spend at least half of their work time on physically demanding tasks (compared to 31.6% of women); and 25.3% are engaged in jobs requiring finger dexterity (compared to 20.2% of women). While no significant variations can be found analysing data by age groups. Employed persons who devote at least half of their time to activities requiring relational skills account for 47.9%. Regarding verbal interactions with colleagues within the same organization, about 36% of workers use this skill for most of their work time, nearly half have a low level of use, and only 12% do not use it at all. Communications with external parties are generally less common: among all employed, 28.6% is frequently engage in verbal exchanges with people from outside the enterprise, while 24.0% never do.

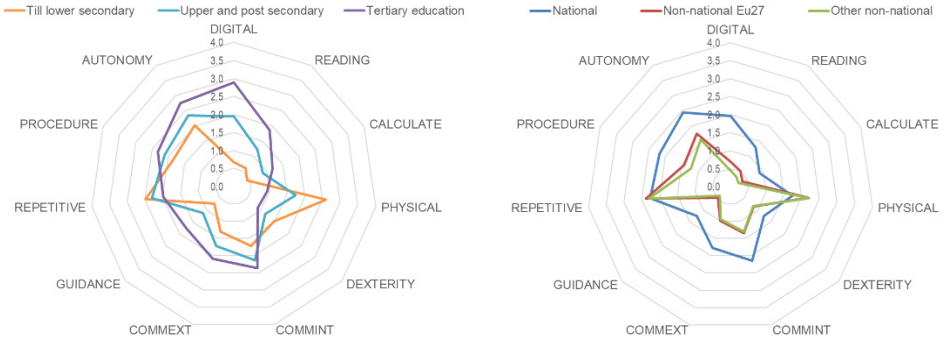
Training, teaching, and providing advice (both formally and informally) are activities performed by over half of employed, but only 17.4% do these activities frequently during their work time; 43.6% never do them.

The two dimensions of autonomy – one concerning the sequence of tasks and the other regarding freedom in defining contents – appear closely associated, although the former is generally more prevalent. 44.7% of employed have significant influence over the order of tasks they perform within their job, while 38.8% have significant influence over contents. Both components increase with age: 28.9% of persons aged 15-29 have a high degree of autonomy in deciding the order of tasks, and 24.7% in defining content, with a notable increase to 43.7% and 37.4%, respectively, for those aged 30-44. This further rises to 48.0% and 41.6% for persons aged 45-59, and reaches 54.2% and 49.1% for those aged 60 and above.

Regarding aspects related to work management, the study examines the repetitiveness of tasks and the frequency of activities that require adherence to procedures. Specifically, 44.2% of employed persons indicate performing repetitive tasks for half or more of their work time, with a higher proportion among women (47.1% versus 42.1% among men). Additionally, 42.5% of employed persons report frequently following standardized procedures.

Relevant differences are found when analysing the data by levels of education and citizenship. In particular, as Figure 2 shows, a strong link can be observed between the level of education attained and the average intensity of using/applying a given skill. The higher the level of education, the more time is spent using digital devices, reading or performing complex calculations, engaging in autonomous tasks, executing protocols with responsibility, conducting training or communication activities, while less time is spent on repetitive tasks or those requiring precise finger movements or intense physical work. Similarly, foreign workers have significantly lower scores compared to Italians across all dimensions - without appreciable variations between Eu27 and non-Eu citizens - except for physically demanding activities and the degree of task repetitiveness, where they surpass them.

Figure 2 - *Employed persons by time spent on different activities in their job by citizenship and by level of education. Year 2022 (mean score).*



Source: Istat, Labour Force Survey

4. A multivariate analysis

Multiple Correspondence Analysis (MCA) was conducted to investigate the associations between the modes of the main variables related to skills and those concerning the key characteristics of the employed. In particular, the analysis focused on the level of use of cognitive skills (synthesis of reading and calculate), physical skills, dexterity, social skills (synthesis of internal and external communication), digital skills and training. Variables related to the frequency of procedural application, task repetitiveness, and degree of autonomy were also included. For all listed skills, modes were categorized into three levels: absent, low, and high. Other considered characteristics included age, sex, education level, citizenship, employment status (employee/self-employed), and occupation at the first digit level (see Appendix for details).

The first two dimensions explain a variance of 89.5%, with the first dimension explaining 76.5% and the second dimension explaining 13%.

The horizontal axis discriminates between advanced intellectual/digital skills (on the left) and manual/unskilled abilities (on the right). Well-represented on this axis are the modes related to digital, cognitive, and manual (as well as social) skills in the high and absent categories, and educational attainment ranging from a university degree to lower secondary education (Figure 3).

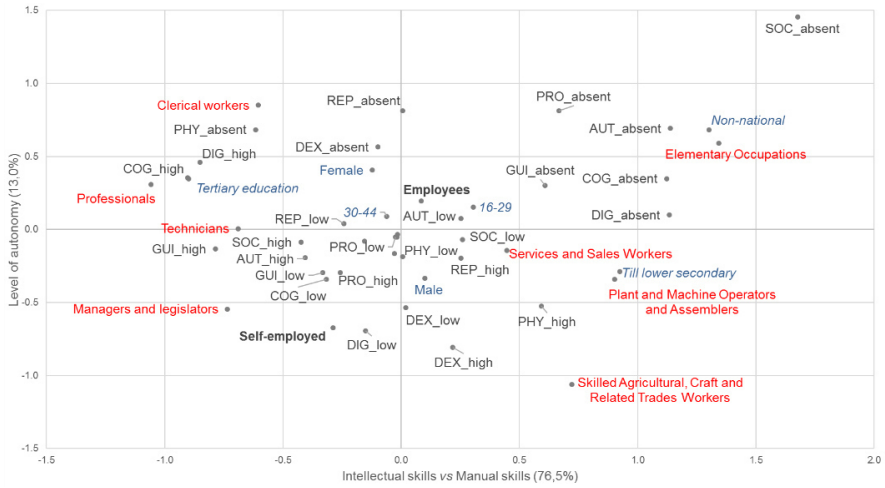
The vertical axis is primarily characterized by modes related to employment status, particularly self-employment, as well as modes of physical (high/absent) and dexterity (high/absent) skills, and occupational categories such as artisans and clerical workers. It identifies the dichotomy between work autonomy (upper part)

and non-autonomy (lower part), specifically represented by clerical workers on one side and artisans on the other, in their more detailed versions.

In the first quadrant, we find individuals with absent digital and cognitive skills, lack of autonomy, non-national, and unskilled professions such as ushers, warehouse workers, and domestic workers, also characterized by absence of social interaction.

The second quadrant identifies profiles with high levels of digital and cognitive skills, while physical and dexterity skills are absent, with low or absent repetitiveness. In this quadrant, we find professionals (such as science, engineering and information and communications technology professionals, teachers, legal, social and cultural professionals and business and administration professionals) and clerical workers. Technicians (such as science and health associate professionals, information and communications technicians, business and administration associate professionals) are positioned on the x-axis, predominantly on the left side. Clerical workers spend more time using digital equipment, but are also those who need to perform complex calculations during their work, along with legislators, managers and entrepreneurs.

Figure 3 - *Multiple Correspondence Analysis.*



Source: Istat, Labour Force Survey

The third quadrant is characterized by managers and legislators and also by self-employment, with profiles defined by high autonomy, strong communication skills, both internal and external, low levels of use of cognitive and digital skills, though not absent.

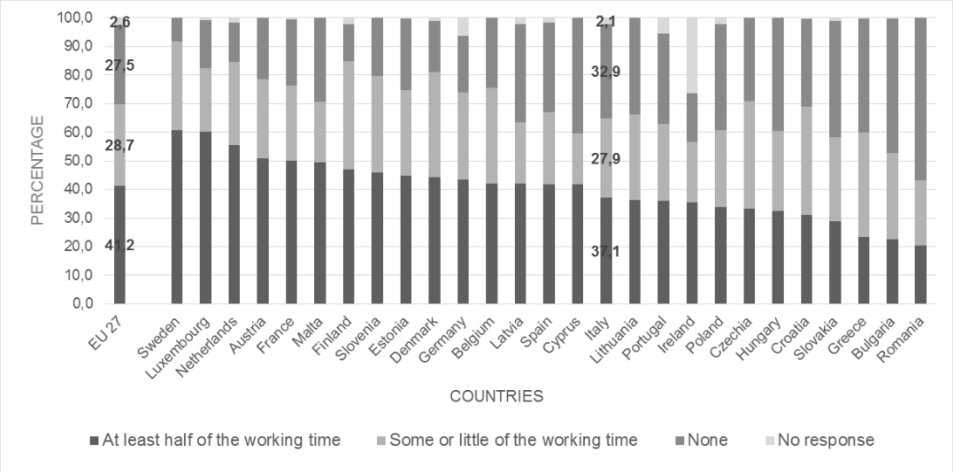
The fourth quadrant defines the professional profiles of skilled agricultural, craft and related trades workers, associated with high levels of dexterity and physical

strength, such as handicraft and printing workers, precision-instrument makers and repairers, musical instrument makers and tuners, jewellery and precious metal workers, food, processing, woodworking, garment, leather and other craft and related trades workers. It also includes services and sales workers (shop salespersons, personal care workers in health services) and plant and machine operators and assemblers with lower education levels (up to lower secondary education) and characterized by a high degree of task repetitiveness.

5. Digital, cognitive and manual skills: Italy in the European scenario

The Job Skills module, harmonised among European countries, allows for data comparison enabling to observe Italy's ranking in relation to the Eu average and other countries. In 2022, 41.2% of employed persons in the Eu27 aged 15-74 reported using digital devices for at least half of their working time and nearly 30% for the entirety or most of their working time (Figure 4).

Figure 4 - *Employed persons by working time spent on using digital devices in Eu27. Year 2022 (age group 15-74, percentage of total employment).*



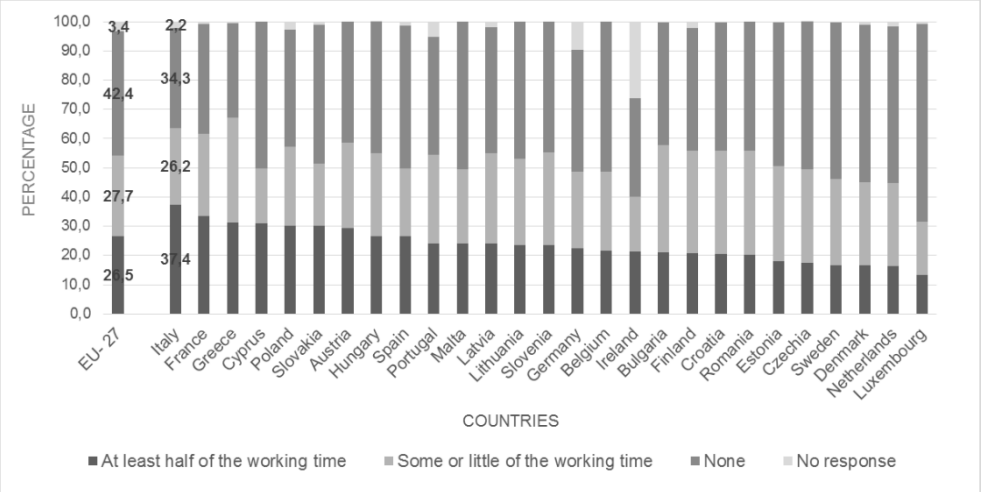
Source: Eurostat (ad hoc extraction 2022 Eu-LFS module)

With 37.1%, Italy had a share similar of the Eu27 of employed persons aged 15-74 who reported using digital devices for at least half of their working time. At country level, Sweden and Luxembourg had the highest share with 60% of employed persons who spent all or at least half of their working time using digital devices. Conversely, Greece, Bulgaria, and Romania had the lowest rates.

However, in Italy, approximately one third of the employed do not use digital devices (32.9%), higher than the European average of 27.5%.

Referring to the cognitive skills, within the Eu, one in ten employed persons aged 15-74 dedicated more than half of their working time to performing relatively complex calculations and approximately three out of twenty employed dedicated more than half of their working time reading work-related documents. The highest percentages for both measures were recorded in Austria, France, Luxembourg, Latvia and Malta. Italy is very close to the European average both for calculation and for reading. In addition, for many European countries, there is a strong positive correlation between the percentage of employed persons who spend more than half of their working time reading documents and the percentage of those engaged in complex calculations.

Figure 5 - *Employed persons by working time spent on doing hard physical work in Eu27. Year 2022 (age group 15-74, percentage of total employment).*



Source: Eurostat (ad hoc extraction 2022 Eu-LFS module)

Regarding the physical skills, nearly one fourth of employed persons in the Eu aged 15-74 reported doing hard physical work for the entirety or at least half of their working time. With 37.4%, Italy had the highest share of employed people aged 15-74 in the Eu who spent all or at least half of their working time doing hard physical tasks, followed by the France and Greece and Cyprus, with percentages also exceeding 30% (Figure 5). In contrast, in the same situation were less than one in seven of their counterparts in Luxembourg (13.3%), followed by the Netherlands (16.4%), and Denmark (16.7%).

Also for dexterity skills, Italy has a higher percentage of employed persons who use them for the majority of their time compared to the Eu, which stands at 16.8%; Italy has 23.2%, placing second after Slovenia.

6. Conclusions

The 2022 ad hoc module provided a comprehensive overview of the skills applied in the Italian labour market and facilitated a comparison between the Italian and European situations.

Italy ranks among the lowest in Europe for the use of digital skills, which is a significant concern in an increasingly digital world. Conversely, Italy stands out for its high use of physical skills, and also dexterity, characteristic of craft professions. This highlights a distinctive aspect of the Italian labour market, which remains deeply rooted in traditional craftsmanship.

Italy experiences high levels of skill mismatch compared to other European countries. Many Italian employed are in jobs that do not align with their educational qualifications, leading to underutilization of their skills. This mismatch is particularly pronounced in Italy's productive system, which is predominantly composed of small, family-run businesses with limited propensity to invest in research and development. Consequently, the relatively few graduates in Italy often have difficulty finding suitable employment that matches their qualifications.

In summary, Italy faces major challenges in aligning its workforce skills with the demands of the modern labour market. Addressing the digital skills deficit and reducing the high levels of skill mismatch are crucial steps for Italy to enhance its productivity and competitiveness. By fostering a culture of continuous learning and encouraging investment in research and development, Italy can better leverage its rich heritage of craftsmanship while also embracing the digital future.

Appendix

Multiple Correspondence Analysis: partial contributions to inertia for the column points.

Variables	Modes	Dim1	Dim2
DIGITAL	DIG_absent	0.1060	0.0014
	DIG_high	0.0699	0.0363
	DIG_low	0.0017	0.0649
COGNITIVE	COG_absent	0.0958	0.0163
	COG_high	0.0402	0.0110
	COG_low	0.0129	0.0265
PHYSICAL	PHY_absent	0.0345	0.0756
	PHY_high	0.0330	0.0460
	PHY_low	0.0000	0.0043
DEXTERITY	DEX_absent	0.0014	0.0783
	DEX_high	0.0027	0.0666
	DEX_low	0.0000	0.0299
PROCEDURE	PRO_absent	0.0206	0.0545
	PRO_high	0.0074	0.0170
	PRO_low	0.0001	0.0005
AUTONOMY	AUT_absent	0.0339	0.0224
	AUT_high	0.0218	0.0087
	AUT_low	0.0061	0.0010
REPETITIVE	REP_absent	0.0000	0.0251
	REP_high	0.0072	0.0078
	REP_low	0.0070	0.0004
SOCIAL	SOC_absent	0.0366	0.0490
	SOC_high	0.0222	0.0017
	SOC_low	0.0079	0.0010
GUIDANCE	GUI_absent	0.0416	0.0184
	GUI_high	0.0292	0.0015
	GUI_low	0.0104	0.0145

Variables	Modes	Dim1	Dim2
EDUCATION	Upper and post secondary	0.0000	0.0002
	Till lower secondary	0.0579	0.0100
	Tertiary education	0.0546	0.0146
EMPL. STATUS	Employees	0.0014	0.0135
	Self-employed	0.0048	0.0462
OCCUPATION	Managers and legislators	0.0043	0.0042
	Professionals	0.0469	0.0070
	Technicians	0.0207	0.0000
	Clerical workers	0.0116	0.0408
	Services and Sales Workers	0.0097	0.0018
	Skilled Agricultural, Craft and Related Trades Workers	0.0184	0.0707
	Plant and Machine Operators and Assemblers	0.0156	0.0040
	Elementary Occupations	0.0467	0.0160
CITIZENSHIP	National	0.0054	0.0027
	Non-national	0.0458	0.0223
SEX	Female	0.0017	0.0337
	Male	0.0014	0.0276
AGE	16-29	0.0024	0.0011
	30-44	0.0003	0.0011
	45-59	0.0000	0.0006
	60 and over	0.0000	0.0015

References

BRUNELLO G., WRUUCK P. 2021. Skill shortages and skill mismatch: a review of the literature, *J Econ Surv.* 2021, Vol. 35, pp. 1145–1167.

EUROPEAN COMMISSION. 2022. *European Skills Agenda for sustainable competitiveness, social fairness and resilience*. European Skills Agenda - Employment, Social Affairs & Inclusion - European Commission (europa.eu)

GALLETTI F., GUALDI F. 2017. *Skills Mismatch in Italia. Analisi e scelte di policy in uno scenario in rapida evoluzione*. Milano: Action Institute.

ISTAT. 2024. *Le competenze professionali nel mercato del lavoro italiano. Anno 2022. Statistica focus*.

ISTAT. 2023. *Rapporto Annuale 2023. La situazione del paese*, Roma.

ISTAT. 2022. *Livelli di Istruzione e ritorni occupazionali. Anno 2021, Statistiche Report*.

OECD. 2018. *OECD Skills Strategy Diagnostic* Report: Italy 2017, OECD Skills Studies, Paris: OECD Publishing.

Barbara BOSCHETTO, Istat, boschett@istat.it
Elisa MARZILLI, Istat, marzilli@istat.it
Silvia MONTECOLLE, Istat, montecol@istat.it
Alessia SABBATINI, Istat, sabbatini@istat.it

LABOUR MARKET AND GENDER ROLE STEREOTYPES: THE EFFECTS OF THE PANDEMIC[1]

Arianna Carra, Paola Maddalena Chiodini, Concettina Impallomeni, Elena Sorba, Flavio Verrecchia

Abstract. Women have always participated in the labour market, though their roles and the social recognition have varied across different periods and cultures. Despite major changes from the industrial revolution, the world wars, women's rights movements, and laws promoting equal opportunities, gender stereotypes and inequalities persist and seem deeply rooted even in societies regarded as "advanced" and democratic. From an occupational and social perspective, the pandemic may have been a turning point. The shift towards flexible work and increased demand in certain sectors may have given women more job opportunities and encouraged them to rethink gender stereotypes. This study aims to explore the social and occupational impacts of the pandemic. To achieve this, the study will use official statistical data and apply normative, positive, and statistical analysis methods.

1. Introduction

The idea of male "dominion" over women traces back to the corruption of the original design: "*Your desire will be for your husband, and he will rule over you*" (Genesis 3:16), to the extent that, even today, traits of the female universe are well represented by the metaphor of "The Chained Elephant". One need only consider gender stereotypes, and that "terrible day for her story" when a young woman, accepting these stereotypes and their real-world consequences, resigns herself to the fate of her grandmother and, later, her mother—whose stories become her own. She "never ever tried to test his powers again", even as an adult. *I closed my eyes, and I tried to imagine the small, newborn elephant, chained to the ground. The small elephant would push, pull and struggle with all his strength, trying to free himself, but he would fail. Despite all his efforts, he would fail again and again, because that stake and chain was too big for his strength. […] Then one day would come that he would just give up, and accept his fate […] He never ever tried to test his powers again. [...] We do the same thing to ourselves that the elephant did, we printed into our minds the message: "I can't – I can't and I will never can."* (Bucay, 2008). This paper aims to study the impact of COVID-19 on gender inequality.

[1] Even if the article is the joint responsibility of the authors, the following individual contributions can be distinguished: paragraph 2 and 5 (up to figure 2) is attributable to Chiodini; paragraph 3 and 4 (third subsection after table 1B onwards) is attributable to Carra; paragraph 4 (first subsection) is attributable to Impallomeni; paragraph 4 (second subsection up to table 1B) is attributable to Sorba; paragraph 5 (after Figure 2 onwards) is attributable to Verrecchia.

While several dimensions related to gender inequality can be observed through official statistics and administrative sources—such as the presence of stereotypes, different contributions to domestic and paid work, and horizontal and vertical segregation in the labour market—only some of these dimensions have recent data available from which initial insights into the impact of COVID-19 on gender disparities can be drawn. After the literature review (para. 2) and an overview of the data and methods used in the paper (para. 3), gender inequality will be studied through official statistics (para. 4) and original findings (para. 5). Concluding remarks will be presented in the conclusions.

2. Literature review

A stereotype is a conventional pattern of speech or attitude—a cliché—a preconstituted, generalized and simplistic opinion, not based on personal evaluation of individual cases rather mechanically repeated, often regarding people, events, or situations. However, recognizing that we are using them is not always straightforward. Stereotypes are part of common sense and learned culture, reflecting socio-historical beliefs and conditions. Not only do they enshrine power differences between social groups, but they also shape expectations regarding the groups' social behaviour. This, in turn, leads those who are subjected to the stereotype to conform to these expectations (Farinella, 2019). This creates a normative function that tends to resist change. Stereotypes generate prejudices that not only define what people are but also dictate what they should be. A classic example of this is gender stereotypes and the associated gender gap, which in the labour market result in both horizontal segregation—where women are concentrated in specific sectors—and vertical segregation, where it is significantly more difficult for women to advance in their careers. The accepted and internalized stereotype becomes reality to such an extent that, in the fourth quarter of 2022, the female employment rate in Italy (Chamber of Deputies, 2023) was the lowest among the member states of the European Union. The international policy framework highlights key strategies, such as the 2030 Agenda, the EU Gender Equality Strategy 2020-2025, and the European Parliament and Council Directive (EU) 2023/970 of May 10, 2023, on equal pay. At the national level, commitments are reflected in the National Recovery and Resilience Plan (PNRR), the Gender Equality Strategy 2021-2026, and the Gender Budget. It becomes clear that women's freedom in the old continent is still not on par with that of men. Despite better academic performance in both school and university, women continue to face challenges in entering and remaining in the labour market (AlmaLaurea and AlmaDiploma University Consortium, 2024). One observed effect of the gender gap is low fertility. Italian women are often forced to either abandon or, at the very least revise and scale back their reproductive plans in order to avoid marginalisation in the labour market. COVID-19 (Ferrario, Profeta, 2020; Ferrario, Profeta, 2021) has exacerbated pre-existing and persistent gender inequalities that were already the subject of international debate prior to the pandemic (EU Parliament Resolution, 2016; UN,

Sustainable Development Agenda, Goal no. 5 "Gender Equality"). Menniti and Demurtas (2012) observe that the time spent on housework is influenced by the level of education. In fact, as women's level of education increases, the amount of time spent on such activities tends to decrease. In contrast, an inverse pattern is observed in men. Similarly, higher income leads to the same result (Kofman et al., 2000; Sarti, 2006, Farvaque, 2013) namely, a reduction in the time spent on domestic activities, which are often delegated to external workers (i.e. domestic workers). This is particularly true in families with high income, where the woman also holds a high level of educational qualification (Lagomarsino, 2004; Farvaque, 2013; Todesco, 2013). It follows that in couples where both partners have high educational qualifications, more equitable gender dynamics emerge, where it is not assumed that the woman should be responsible for housework. However, women continue to spend more time on routine domestic work than men, even when they have high levels of education and income (Carriero, Todesco, 2016) highlighting how deeply entrenched the gender stereotype is that women are expected to care for the home and family. Nyman, Reinikainen and Eriksson (2018) argue that even in couples where there is an apparent commitment to gender equality, there remains a tendency to reproduce and perpetuate gender inequalities.

3. Data and methods

The data on gender stereotypes come from the surveys "on gender stereotypes and the social image of violence" conducted by Italian National Institute of Statistics (Istat) in 2018 and 2023 (the data of which are provisional as, at the time of their publication, the survey had not yet been completed (Istat, 2023) and the final estimates have not yet been released). This is a survey that Istat implemented on the basis of a Cooperation Agreement with the Department of Equal Opportunities (Dipartimento delle pari opportunità e Istat, 2019). On both occasions, the samples originally consisted of approximately 15,000 individuals aged between 18 and 74, who were interviewed using the CATI technique in 2018 and invited to complete a questionnaire using the CAWI method in 2023, with the option of a CATI interview if they were unable to respond online. In 2018, individuals were selected from families interviewed using the CATI technique in the Labour Force survey during the period June-November 2018 (two-stage random sampling without stratification of the I and II stage units). In 2023, respondents were drawn from municipal registry lists to create a random and statistically representative sample of the population residing in Italy. For the purposes of this work, the focus will be exclusively on the stereotype that men are less suited to housework. Data from the Labour Force Survey (Istat, 2024) are also used for ex-post forecasts and to understand the type of outlier caused by COVID-19 on female employment. For the study of segregation, rather than labour force data, it is necessary to use official statistics from administrative databases on businesses (Istat, 2022). This is because, in terms of economic activity, a correct classification is provided by sources where the sectoral

specialization is declared by the entrepreneur and verified by Istat according to international standards. Two distinct methodological approaches were used to test the effects of COVID-19 on stereotype spillovers and related gender inequalities. The first employs the Cochran-Mantel-Haenszel (CMH) test (Cochran, 1954; Mantel and Haenszel, 1959). The CMH test is a generalization of the Mantel-Haenszel (MH) test. It is used to study the association between two binary variables while controlling for a third nominal confounding variable. It allows testing the association between a binary predictor (treatment) and a binary outcome in the presence of stratification. In practice, it tests the null hypothesis of equality of the odds ratios (ORs) of the contingency tables under consideration against the alternative hypothesis that at least one of the ORs considered is statistically different. The CMH is a commonly used technique in biostatistics, particularly in case-control studies. The second approach aims to identify any structural change in the dynamics of female participation in the labour market using an automatic modelling routine based on the TRAMO method (time series regression with ARIMA noise, missing values and outliers) and the specification of different types of outliers (Gómez and Maravall, 1997).

4. Official statistics on gender inequality

Entrenched stereotypes and cultural legacies continue to sustain significant gender differences, despite a slow trend toward convergence. Consequently, the increasing participation of women in the labour market highlights their growing presence in society while reflecting dual roles: the traditional role of wife and homemaker on one hand, and that of a worker on the other. This combination of roles reveals distinct repercussions on women's life policies and the organisation of life time. Labour participation, with an employment rate of 52.5 percent in 2023 (up from 48.9 percent in 2017), demonstrates the growing presence of women in the labour market.

However, significant gender differences persist. Horizontal segregation—referring to the concentration of female workers in specific sectors—remains evident, particularly in non-market service sectors, which account for 56 percent of the total. These sectors include, among others, education, health care, social assistance, and artistic, sporting and entertainment activities (Table 1A). Although women have gradually gained ground in tertiary education and the labour market over the years, the traditional division of gender roles has largely remained unchanged. Despite higher productivity in sectors where women are most employed — such as services — wages in these areas remain disproportionately lower compared to other sectors. (Table 1B). Regarding vertical segregation—i.e. the concentration of female workers in certain job positions—despite the fact that Marilyn Loden coined the term “glass ceiling” in 1978, over 40 years ago, women managers still represent only 15.5 percent of the total while the share of female workers is 40.7 percent (Table 1A).

Table 1A – *Female employees and managers, by economic activity, Italy, 2017 (% of total males and females).*

NACE	*Female employees (%)*	*Female managers (%)*
Total	40.7	15.5
Industry in the strict sense (B-E)	27.5	11.5
Construction (F)	10.9	7.1
Wholesale and retail trade, transportation and storage, accommodation and food service activities (G-I)	43.7	17.4
Other service activities (J-S excluding O)	56.0	19.3

Table 1B – *Wages and salaries per employee and apparent labour productivity, by economic activity, Italy, 2017 (euro per employee).*

NACE	*Salaries and wages (euro)*	*Productivity (euro)*
Total	26,058.18	47,152.19
Industry in the strict sense (B-E)	31,958.77	70,197.35
Construction (F)	25,988.80	37,207.28
Wholesale and retail trade, transportation and storage, accommodation and food service activities (G-I)	23,431.72	37,965.56
Other service activities (J-S excluding O)	22,965.74	42,301.04

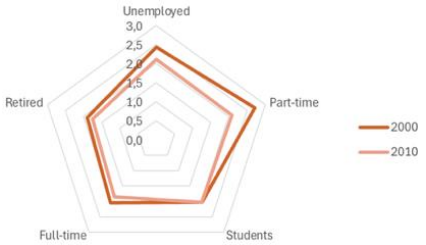
Notes: Based on our elaborations on Istat Asia-employment data. NACE Rev.2: [B] Mining and quarrying, [C] Manufacturing, [D] Electricity, gas, steam and air conditioning supply, [E] Water supply, sewerage, waste management, [F] Construction, [G] Wholesale and retail trade, repair of motor vehicles and motorbikes, [H] Transportation and storage, [I] Accommodation and food service activities, [J] Information and communication services, [L] Real estate activities, [M] Professional, scientific and technical activities, [N] Rental, travel agencies, business support services, [P] Education, [Q] Human health and social work activities, [R] Arts, sporting activities and amusement and recreation activities, [S] Other service activities.

Women remain bound by cultural legacies and stereotypes, which lead them to make more consistent contributions to domestic work, even when employed full-time. In 2010, women devoted on average twice as many hours to housework and family care as men (more than 5 hours and 21 minutes compared to 2 hours and 22 minutes), even -and especially- when the man was unemployed or working part-time (Figure 1A). The unequal gender division within housework is a widespread phenomenon across almost all European countries. Housework and family care responsibilities are predominantly shouldered by women. The 2014 Istat Time-Use Survey also revealed that, on average, 3 hours and 46 minutes per day were spent on unpaid work. This includes "activities related to caring for one's home and the people living in it —whether children, adults or elderly family members—as well as organised voluntary work, informal assistance between families and all the travel associated with these activities" (Istat, 2019). However, the extent of gender-specific involvement revealed significant disparities: in general, men participated in 74.6 percent of cases, contributing an average of 2 hours

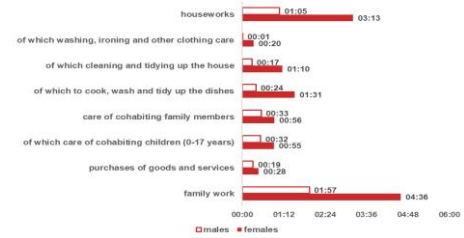
and 16 minutes, while women were involved in 92.3 percent of cases, dedicating an average of 5 hours and 9 minutes. This disparity, as can easily be inferred, affects both the opportunities to engage in other areas of daily life, such as paid work, and leads to a reduction in available leisure time. When focusing solely on the housework of couples where both partners were employed and the woman was aged between 25 and 64, the imbalance remained, particularly in the domain of domestic work (Figure 1B).

Figure 1 – *Average daily time spent on family work.*

A. by self-declared labour status, Italy 2010 (indexes, males=1)



B. persons in couple with both partners employed (female 25-64), by type of activity, Italy 2013-14 (hh:mm)



Notes: Based on our elaborations on Harmonised European time use surveys (A), Istat time use survey data (B).

5. Results of analyses of the impact of COVID-19 on gender inequality

In this study, two applications were used to examine the structural impact of COVID-19 on stereotypes and women's participation in the labour market. According to the data of the 2023 edition of the survey Gender stereotypes and the social image of violence (Istat, 2023), agreement with the statement that men are less suited to do housework has significantly decreased. The general average is now 21.4 percent. Among men, the proportion of those who "strongly or somewhat agree" falls to 24.6 percent. Among women, the change is significant, both in quantitative terms "settling" at 18.3 percent (a decrease of 14 percentage points compared to 2018) and in structural terms (in 2018, more women than men agreed with the stereotype).

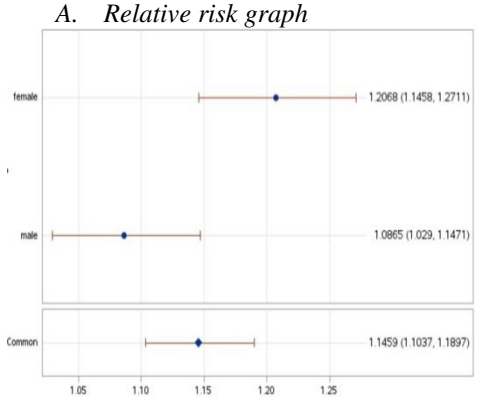
Due to data limitations, the analysis relies solely on percentage compositions. The pandemic period (hereafter COVID) is treated as an exposure factor for changes in the status of the stereotype2. According to this strategy, it is assumed that subjects of both sexes were "unexposed" in 2018 and "exposed" in 2023. Their response to exposure is coded as "change" (i.e. they no longer believe that men are less suited to do housework) or "no change" (i.e. they still believe that men are less suited to do housework). Cochran-Mantel-Haenszel (CMH) test allows for the observation of the overall relative risks in order to assess the effectiveness of COVID exposure. The CMH yields a strongly significant p-value (<0.0001), indicating that the association between COVID exposure

[2] The "COVID" label encompasses a multiplicity of factors that manifested themselves over the period.

and response remains strong even after adjustment for gender. The relative risk graph (Figure 2A) displays the relative risks and confidence intervals for both genders and overall (common) relative risk. The probability of change due to COVID exposure shows an overall improvement of about 15 percent across the entire population, with significant gender differences (about 9 percent for men vs. about 21 percent for women). In addition, considering odds ratios (Figure 2B), the homogeneity test—i.e., the Breslow-Day Test—is used to assess the hypothesis that all strata are equal. The Breslow-Day Test p-value (0.0018) provides strong empirical evidence in favour of the alternative hypothesis of non-homogeneity, i.e., there are differences in the odds ratios strata. In other words, gender differences regarding the stereotype in question were evident.

Figure 2 – *COVID exposure response for the stereotype men are less suited to do housework.*

A. Relative risk graph



B. Statistics

	Statistic	*Value*	*95% confidence Limits*	
Female	Odds ratio	2.13	1.73	2.62
	Relative risk	1.21	1.15	1.27
Male	Odds ratio	1.35	1.11	1.65
	Relative risk	1.09	1.03	1.15
Common	Odds ratio	1.68	1.46	1.94
	Relative risk	1.14	1.10	1.19

Notes: Based on our elaborations on Istat data. Control by gender. Risks calculated for response equal to "Change". Wald 95% confidence limits. Odds ratios and relative risks are calculated by the MH and CMH method.

While COVID may have influenced the perception of stereotypes, there is no empirical evidence of structural changes in female labour market participation. The growth paths of female labour market participation seem to have started earlier and are consistent with the dynamics of more critical turnover in the working population. Using both monthly (Figure 3A) and quarterly (Figure 3B) data, the automatic modelling routine based on the TRAMO method identifies outliers, which are characterized as Level Shift (LS) outliers, i.e. models are specified that reflect structural changes in female employment expressed by multiple LS outliers identified exclusively in the pandemic year. However, further in-depth analyses reveal that these changes are temporary. In contrast to this result, an alternative strategy of using correctly specified forecasting models is to use them in terms of non-expected outcomes (Verrecchia, 1993).

In particular, when considering data up to and including 2020, the automatic model specification identifies an LS-type outlier (Figure 4A), which results in forecasts with a level shift that over time underestimate the cyclical trend component of female employment. On the contrary, forecasts based solely on pre-COVID data (Figure 4B) reject the structural change hypothesis in favour of TC (Temporary Change) outliers. Confirmation of this result can also be obtained by specifying ex-ante LS-type outliers versus TC-type outliers, even when considering data from the pandemic year. In this case, the automatically identified outliers (Figure 3B) were specified, if significant, as additive outliers (AO), i.e., outliers whose duration is limited to a period defined by the following expression

$$\mathrm{AO}_t = \begin{cases} 1 & t = t_o \\ 0 & t \neq t_o \end{cases} \tag{1}$$

and, in addition, in the period immediately following, a level shift outlier was alternatively specified (Figure 4C), i.e., whose duration spans the entire period, including and following the event, with uniform intensity

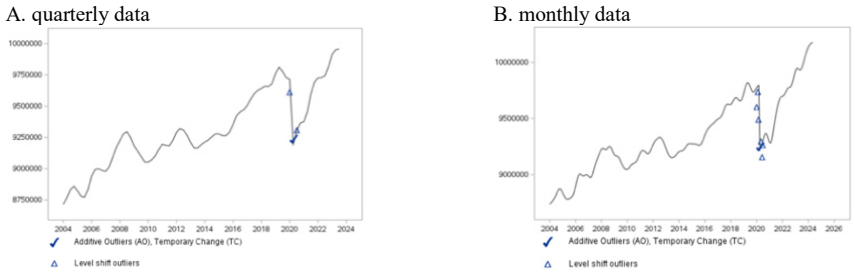
$$\mathrm{LS}_t = \begin{cases} -1 & t < t_o \\ 0 & t \geq t_o \end{cases}, \tag{2}$$

and of the temporary change type (Figure 4D), i.e., whose duration spans the entire period, including and following the event, but with intensity halved in each successive period

$$\mathrm{TC}_t = \begin{cases} 0 & t < t_o \\ \dfrac{1}{2^{(t+i-t_0)}} & t \geq t_o \quad (con\ i = 0,\ 1,\ 2,\ \ldots) \end{cases}. \tag{3}$$

The forecasts confirm the previously obtained results, therefore, the analyses show that COVID had a temporary, rather than structural, effect on female employment. These dynamics can be summarized (Figure 5A and 5B).

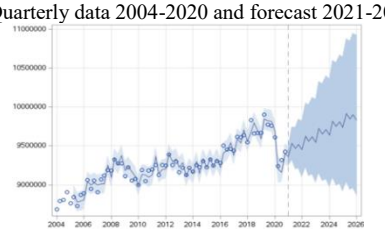
Figure 3 – *Female employment, trend cycle component and outliers, Italy 2004-2023.*



Notes: Based on our elaborations (SAS X11) on Istat Labour Force Survey data.

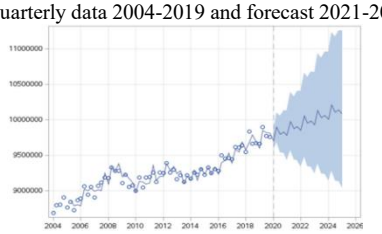
Figure 4 – *Female employment, automatic selection model and forecasts, Italy.*

A. Quarterly data 2004-2020 and forecast 2021-2026



Notes: LS 2020Q2, logarithmic transformation, SARIMA(0,1,1)(0,1,1).

B. Quarterly data 2004-2019 and forecast 2021-2025



Notes: logarithmic transformation. SARIMA(0,1,1)(0,1,1).

C. Quarterly data 2004-2020 and forecast 2021-2026

Notes: Ex-ante outliers specification (AO: 2020Q1 not significant): AO: 2020Q2, 2020Q3, LS: 2020Q4. Logarithmic transformation. SARIMA(0,1,0)(0,1,1).

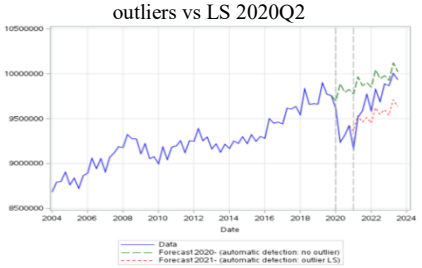
D. Quarterly data 2004-2020 and forecast 2021-2026

Notes: Ex-ante outliers specification (AO: 2020Q1 not significant): AO: 2020Q2, 2020Q3, TC: 2020Q4. logarithmic transformation. SARIMA(0,1,0)(0,1,1).

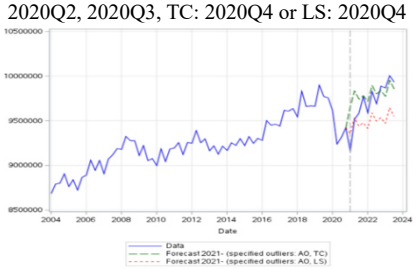
Notes: Based on our elaborations (SAS X13) on Istat Labour Force Survey data.
Data *95% confidence interval* *Forecasts* *beginning of forecasts*

Figure 5 – *Female employment data and forecasts synthesis, Italy, 2004-2023(Q3).*

A. Models with automatic outliers detection: no outliers vs LS 2020Q2



B. Models with ex-ante outliers specification: AO: 2020Q2, 2020Q3, TC: 2020Q4 or LS: 2020Q4



Notes: Based on our elaborations (SAS X13) on Istat Labour Force Survey data.

6. Conclusion

The pandemic period has acted as a catalyst for the widespread adoption of new technologies enabling remote work, the emergence of novel forms of employment, and

a significant shift towards flexible working arrangements (Rapanelli, 2023, p. 187). Additionally, it has prompted a reevaluation of gender stereotypes (see para. 5, Figure 2), largely due to the work environment itself—the domestic sphere—(Del Boca et al., 2020, p. 1013) which has exposed the emperor's new clothes. From this perspective, the decline in employment, which was also observed among the male population during the pandemic (ISTAT, 2021, p. 12), coupled with the lockdown periods that not only affected parents but also children and young people—who were engaged in distance learning and required, especially at younger ages, parental support— may have shed light on and raised awareness of the unequal distribution of domestic and caregiving responsibilities (Rapanelli, 2023, p. 196). The Istat findings and the analysis presented (para. 4) highlight how stereotypes may be far from disappearing. The disruptive impact of the pandemic, along with the subsequent lockdown, has contributed to a more accurate understanding of the unequal (para. 5, figure 2) and generally unjustified distribution of domestic and family responsibilities. However, the long-term effects of this shift should be assessed over time. Furthermore, new employment opportunities for women appear to arise not so much from exogenous events such as COVID (para. 5, figure 5), but from the challenges facing contemporary society. The demographic decline and the associated challenges regarding the structure and turnover of the active population are examples of factors that accelerate the shift away from traditional family roles, driven by the growing demand for greater female participation in the labour market (Istat, 2014, p. 1; Poletti, 2023). From a qualitative perspective, however, the new forms of work introduced by COVID could represent both an opportunity and a double-edged sword for women. On one hand, they might be encouraged to transition from personal services to corporate services (where remote work is more prevalent), leading to a reduction in occupational segregation (para. 4). On the other hand, the potential for better work-life balance could reinforce their attachment to the 'pole' of domestic responsibilities (Rapanelli, 2023, p. 195). Some limitations must be considered when interpreting the results of the present study. Indeed, although the comparison of the levels of sharing of stereotypes provides interesting insights (para. 5, figure 2), please note that the data for 2023 remain provisional. Future official data will allow further development of the study.

Acknowledgements

The authors are grateful to the participants of the 60th SIEDS conference and to the referees and editors for their suggestions and advice, which led to a substantial improvement of the work.

References

BUCAY J. 2008. *El elefante encadenado*, Barcelona, Serres edicione SL.

CAMERA DEI DEPUTATI 2023. L'occupazione femminile. https://documenti.camera.it/leg19/dossier/pdf/PP004LA.pdf

CARRIERO R., TODESCO L. 2016. *Indaffarate e soddisfatte. Donne, uomini e lavoro familiare in Italia*. Roma: Carocci.

COCHRAN W. G. 1954. "Some Methods for Strengthening the Common Chi-Squared Tests" *Biometrics*, Vol. 10, No. 4, pp. 417–51.

CONSORZIO UNIVERSITARIO ALMALAUREA E ALMADIPLOMA 2024. Gender gap tra formazione e mercato del lavoro https://www.almalaurea.it/news/focus-gender-gap-2024

DEL BOCA D., OGGERO N., PROFETA P. E M. ROSSI 2020. "Women's and men's work, housework and childcare, before and during COVID-19" *Rev Econ Household 18: 1001–1017.*

DIPARTIMENTO DELLE PARI OPPORTUNITÀ e ISTAT 2019. *Gli stereotipi sui ruoli di genere e l'immagine sociale della violenza sessuale - anno 2018*. 25 novembre 2019. Roma: Istat.

FARINELLA D. 2019 Ciclo di seminari "Violenza di genere: conoscerla, prevenirla, riconoscerla e contrastarla" Università degli Studi di Messina - Dipartimento di Scienze Politiche e Giuridiche organizzato

FERRARIO T., PROFETA P. 2020. COVID: un paese in bilico tra rischi e opportunità. Donne in prima linea, in Laboratorio Futuro, Istituto Toniolo.

FERRARIO T., PROFETA P. 2021. COVID: la crisi più dura per le donne in un Paese ancora senza parità https://www.laboratoriofuturo.it/ricerche/covid-la-crisi-piu-dura-per-le-donne-in-un-paese-ancora-senza-parita/

FARVAQUE N. 2013. *Developing personal and household services in the EU: A focus on housework activities*, Report for the DG Employment, Social Affairs and Social Inclusion, ORSEU.

FILÌ V. 2021. Le difficili libertà delle donne tra gender wage gap, soffitti di cristallo e bassa fecondità, Lavoro diritti europa – Rivista nuova di diritto del lavoro 2.

GÓMEZ, V., MARAVALL, A. 1997. *Guide for Using the Programs TRAMO and SEATS, Beta Version*. Madrid: Banco de España.

ISTAT 2014. Giovani generazioni nel mercato del lavoro in Italia. https://www.istat.it/it/files/2014/10/focus-giovani.pdf

ISTAT 2019. *I tempi della vita quotidiana. Lavoro, conciliazione, parità di genere e benessere soggettivo*. Roma: Istat [ebook]

ISTAT 2021. *Esame del disegno di legge n. 2144*. https://www.istat.it/it/files//2021/04/Istat-DL-SOSTEGNI-aprile-2021.pdf

ISTAT 2023. *Stereotipi di genere e immagine sociale della violenza: primi risultati*. Dati provvisori maggio-luglio 2023. Statistiche today, Roma: Istat.

KOFMAN E., PHIZACKLEA A., RANGHURAM P., SALES R. 2000. *Gender and International Migration in Europe*. London: Routledge.

LAGOMARSINO F. 2004. *Esodi e Approdi di genere. Famiglie transnazionali e nuove migrazioni dall'Ecuador.* Milano: Franco Angeli.

MANTEL N., W. HAENSZEL 1959. "Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease." *Journal of the National Cancer Institute* Vol. 22, No.4, pp. 719–48.

MENNITI, A., DEMURTAS, P. 2012. Disuguaglianze di genere e attività domestiche. *IRPPS Working Papers,* Vol. 47, pp. 2-45.

NYMAN, C., REINIKAINENB, L., ERIKSSON, K. 2018. The tension between gender equality and doing gender Swedish couples' talk about the division of housework. *Women's Studies International Forum*, Vol. 68, pp. 36-46.

POLETTI F. 2023. In Italia 143 senior ogni 100 junior, ma per i giovani lavoro in crescita, Il Sole 24 Ore, Econopoly, 31.05.2023. https://www.econopoly.ilsole24ore.com/2023/05/31/lavoro-giovani-istat/?refresh_ce=1

RAPANELLI, I. 2023. "Nuovi lavori, nuove alienazioni, nuove strategie di sopravvivenza. Ridefinire il lavoro nell'era post pandemica." *Heteroglossia. Quaderni di Linguaggi e Interdisciplinarità*, 19, 185-202.

RISOLUZIONE DEL PARLAMENTO UE del 13 settembre 2016, P8_TA (2016) 0338.

TODESCO L. 2013. *Quello che gli uomini non fanno*. Il lavoro familiare nelle società contemporanee. Roma: Carocci.

UNITED NATIONS. 2015. Sustainable Development Agenda, Goal n. 5 "Gender Equality".

VERRECCHIA F. 1993. *Alcuni aspetti dello stato di salute della popolazione lombarda negli anni '80*, Università degli studi di Milano, Milano.

Arianna CARRA, Istat, North-West Area Office, carra@istat.it
Paola Maddalena CHIODINI, Università degli Studi di Milano-Bicocca – Dipartimento di Statistica e Metodi Quantitativi, paola.chiodini@unimib.it
Concettina IMPALLOMENI, Istat, North-West Area Office, impallomeni@istat.it
Elena SORBA, Istat, North-West Area Office, sorba@istat.it
Flavio VERRECCHIA, Istat, North-West Area Office, verrecchia@istat.it

TIMING OF PARENTHOOD AND CHILDREN'S EDUCATIONAL OUTCOMES: A NEW APPROACH FOCUSING ON EDUCATION AS A POSITIONAL GOOD

Alessandra Trimarchi

Abstract. As the timing of parenthood is steadily increasing in many Western countries, scholars have been interested in the consequences of later motherhood on children's educational outcomes. Previous studies have shown that, in several contexts, the effect of maternal age on children's education is positive. The mechanisms behind this positive effect, however, are unclear, since both the availability of higher socio-economic resources of older mothers and positive period trends in education may explain the gradient. In this study, I argue that to clarify the association between maternal age and children's education, it is important to consider a relative measure of children's education, focusing on education as a *positional good*. By means of the French survey *Formation et Qualification Professionnelle* (FQP) 2014-2015, and a siblings fixed-effects design, I estimate the effect of maternal age on children's level of education measured in absolute (highest level of education attained) and relative (highest level attained relatively to others in the same age group) terms. Results show that, also in France, maternal age is positively associated with children's – absolute – level of education. Still, the positive gradient disappears when children's education is measured in relative terms. These findings support the argument according to which the positive gradient in children's education by maternal age is substantially driven by educational expansion.

1. Introduction

Parents' socio-economic characteristics are considered important determinants of the level of education that an individual attains, contributing to an increase in educational inequalities (Breen and Müller 2020). Differential demographic behaviour by socio-economic characteristics is at the roots of the reproduction of educational inequalities (Breen et al. 2019; Choi et al. 2020). The timing of parenthood, and maternal age more specifically, received great attention among scholars as one of the dimensions affecting children's education. As timing of motherhood continues to rise, concerns about its effects on children's educational outcomes have been increasing too (Myrskylä et al. 2017; Grätz and Wiborg 2024).

Most research shows a positive effect of maternal age on children's education, which is in contrast with explanations related to biological mechanisms such as reproductive ageing. Reproductive ageing would imply a negative effect of maternal

age on child education because of the risks related to preterm births, lower birth weight, or risks in cognitive disabilities (Cohen 2014). Instead, there exist several reasons maternal age may be positively associated with children's level of education.

First, younger mothers tend to have lower human capital contrary to older mothers, who might have accumulated socio-economic resources and life experiences useful at parenting (Kalmijn and Kraaykamp 2005; Powell et al. 2006; Fishman and Min 2018). Next, using Swedish data and a siblings fixed-effects design, Barclay and Myrskylä (2016a) showed that advanced maternal age could be beneficial for children's education because of benefits derived from positive period trends in educational expansion. A later born may be more likely to obtain a tertiary level of education due to enhancements in access to higher education.

To account for period trends, Barclay and Myrskylä (2016a) included the year of birth of the child as a control variable in the siblings fixed-effects model. This approach has been criticized though. The reason is that maternal age and children's year of birth are linearly dependent in the siblings fixed-effects model, hence it is not possible to identify the effect of one or the other (Keiding and Andersen 2016; Kravdal 2019). Thus, while the siblings fixed-effects model is the ideal technique to account for socio-economic resources of the family where children grow up, it could be less efficient in testing mechanisms relating to period trends. A possible alternative would be to apply a multilevel-multiprocess model of mothers' fertility and children's education (Kravdal 2019), which, however, remains demanding in terms of data availability.

In this study, I argue that to analyse the association between maternal age and children's educational level, considering period effects, it is important to use the siblings fixed-effects model in combination with alternative measures of children's education. None of previous studies focused on measuring education in relative terms, i.e., interpreting education as a *positional good* (Hirsch 1976). This implies that the value of an individual's level of education depends on the level of attainment of others in the same age group. This is an important gap, especially when examining the link between maternal age and children's educational outcomes.

I fill this gap by testing the effect of maternal age on children's educational level measured in absolute and relative terms, applying a paired siblings fixed-effects model to French data from the survey *Formation et Qualification Professionnelle* (FQP) 2014-2015. Results show that, also in France, maternal age is positively associated with children's – absolute – level of education. Still, the positive gradient disappears when children's education is measured in relative terms.

2. On the link between maternal age and children's educational attainment

2.1. Previous findings

Previous studies on the effect of maternal age on children's educational attainment showed a positive effect in the Netherlands (Kalmijn and Kraaykamp 2005), and in Sweden (Barclay and Myrskylä 2016a). Both studies applied a siblings fixed-effects model. This model accounts for family unobserved characteristics, shared by the siblings, that are assumed to be constant over time.

Moreover, these studies were the solely accounting for period trends. The Dutch study included in the model the proportion of tertiary educated of a birth cohort. However, the inclusion of this variable did not lead to any conclusive result because of an overcontrol bias, given that the within-family variation is substantially driven by siblings' timing in enrolment (Grätz and Wiborg 2024). The Swedish study has been also criticized because of the issue of linear dependencies between independent variables (Kravdal 2019). In sum, in both studies, it was not possible to disentangle the effect of maternal age from periods trends.

Several other studies have found a positive effect of maternal age on children's outcomes (Powell et al. 2006; Cantalini et al. 2020; Grätz and Wiborg 2024), while a study also found a nonlinear relationship (Fishman and Min 2018), showing a disadvantage for children of teenagers and very old mothers. Still, the level of analysis remained mainly descriptive since unobserved confounders, and models' specification issues may have contributed to alter the estimates.

2.2. Research question and hypotheses

Do children born to older mothers have a *real* educational advantage net of other mother-related characteristics? In line with previous findings, it is expected that also in France children born to older parents are more likely to attain a high level of education relative to children born to younger parents. I formulate the first hypothesis, according to which there is a positive gradient in children's educational level by maternal age, when education is measured in absolute terms (**H1**).

The positive effect, however, could disappear when we consider education as a *positional good*, i.e., when the value of an individual's education depends on the level of education that other individuals of the same age acquire. Over time, due to educational expansion, more and more graduated individuals have easier access to tertiary education, and the educational distribution across children's birth-cohorts changes. Thus, the positive gradient in children's educational attainment by maternal age could be substantially driven by positive trends in educational expansion. Therefore, according to the second hypothesis, it is expected that when measuring children's education in relative terms, to account for positive period trends, the positive gradient by maternal age disappears or flattens (**H2**).

3. Data and methods

3.1. Sample selection

To test these hypotheses, I use data from the French survey *Formation et Qualification Professionnelle* (FQP) 2014-2015 ((FQP), INSEE, 2015). The survey took place from April 2014 to December 2015 with a sample of 26,861 people born between 1950-1992. The survey deals with topics related to social mobility, school to work transition, the relationship between education, occupation, and income. Beyond information about the mother and the father, respondents were requested to answer questions about a sibling, who, during the interview, was randomly chosen among all siblings (if any).

The sample is constituted by respondents with at least one sibling, respondents without siblings (N = 2,690), or with missing information about sibling's year of birth (N=174) or sibling's educational level (N=2,224) were excluded from the analysis. It is more likely that by 25 years old, individuals have reached their final educational attainment, hence, if respondents (or their sibling) were younger than 25 years old at the time of the survey, they were excluded (N = 2,441). Respondents with missing information about mother's year of birth (N = 893) were also dropped from the analysis. Respondents were dropped from the analysis if there were any inconsistencies between the age of the mother and that of the children (N = 14), or between siblings' ages (N = 12), or in case the mother was younger than 15 years old at birth (N = 208), or older than 50 years old (N = 15). Overall, the sample totalled of 18,190 pairs of siblings, i.e., 36,380 units of analysis, born between 1927-1990.

The main outcome variable is the highest level of education reached at the time of survey by the respondent and the respondent's sibling. The absolute level of education is operationalized in six values, in terms of the International Standard Classification of Education (ISCED) from 2011: (1) individuals without a degree or with primary studies only; (2) individuals who completed the lower-secondary level; (3) individuals who obtained a high school diploma, i.e., upper-secondary level; (4) individuals who obtained a diploma with two more years after finishing the upper-secondary level; (5) lower-tertiary level graduates, i.e., those who attained a bachelor degree; (6) individuals who completed a master or a higher degree. Then, a dichotomous variable indicating whether the individual obtained at least a lower-tertiary degree has been created.

Besides measuring the highest level of education in absolute terms, I have also used a relative measure, which accounts for educational expansion across cohorts (Bukodi and Goldthorpe 2016; Triventi et al. 2016). There is not only one - best - way of constructing a relative measure, since it depends on the data, the research question, and the analytical method applied (Bukodi and Goldthorpe 2016). In line with previous work, I have re-operationalized the absolute variable using quartiles,

namely collapsing the categories of the variable based on changes across cohorts, according to the proportions of cohort members holding a certain degree. Hence, the new variable consists of four values ('1' is the lowest quartile, '4' is the highest quartile, defined by birth cohort). Then, I have constructed a dichotomous variable indicating whether the individual is in the highest quartile or not. Note that birth-cohorts refer to categories of multiple years (below 1949; 1950-1959; 1960-1969; 1970-1979; 1980-1990)[1].

The main covariate of interest is maternal age at birth of the respondent and the respondent's sibling, and it is operationalized in six categories (15-19; 20-24; 25-29; 30-34; 35-39; 40+) to account for nonlinearities. I have also included siblings' sex and birth order as control variables, since both variables have been shown to be relevant in analysing children's educational outcomes (Barclay and Myrskylä 2016a). Table 1 summarizes the distribution of outcome variables and covariates in the sample considered.

Table 1 - *Description of the variables in the sample considered.*

Outcome Variables	**N**	**%**
Absolute Educational Level		
Did not obtain a tertiary degree	29309	80,6
Obtained a tertiary degree	7071	19,4
Relative Educational Level		
Not in the highest education quartile	30033	82,6
Highest education quartile	6347	17,4
Independent Variables		
Maternal age at birth		
15-19	2212	6,1
20-24	11272	31,0
25-29	12295	33,8
30-34	6877	18,9
35-39	2862	7,9
40+	862	2,4
Sex		
Male	17870	49,1
Female	18510	50,9
Birth order		
First born	11954	32,9
Second or higher order	24426	67,1
Total	36380	100

Notes: Own elaboration on FQP 2014-2015 data, N paired siblings = 18190.

[1] Conclusions remain the same when using single birth-years (results available upon request).

3.2. Analytical strategy

I have applied a paired siblings fixed-effects design, which allows to compare siblings within the same family to each other, taking into account all unobserved family related variables that are constant over time. This approach permits to estimate the effect of maternal age at birth on children's educational attainment, net of unobserved family characteristics. Using the binary variable as outcome, I apply a linear probability model (LPM) with fixed-effects, formally, it can be written:

$$Y_{ij} = \alpha + \beta X_{ij} + U_i + e_{ij}$$

Y_{ij} corresponds to the best ordinary least squares (OLS) estimate for $\hat{P}(Y_{ij} = 1)$, the probability of acquiring the highest educational level for sibling j of mother i, X_{ij} relate to covariates whose values change across siblings, U_i represents normally distributed systematic differences between families/mothers, $U_i \sim N(0, \sigma_u^2)$, e_{ij} corresponds to the normally distributed within-siblings error of j^{th} order for the i^{th} mother $e_{ij} \sim N(0, \sigma_e^2)$.

While with binary dependent variables, the logit model could be preferrable, recently, researchers have acknowledged that the application of a LPM could be a better analytical strategy facilitating the interpretation of results, and the comparison of models (Mood 2010; Timoneda 2021). Nevertheless, I have also estimated a logit fixed-effects model, and the conclusions remain substantially the same to those reported here (results available upon request).

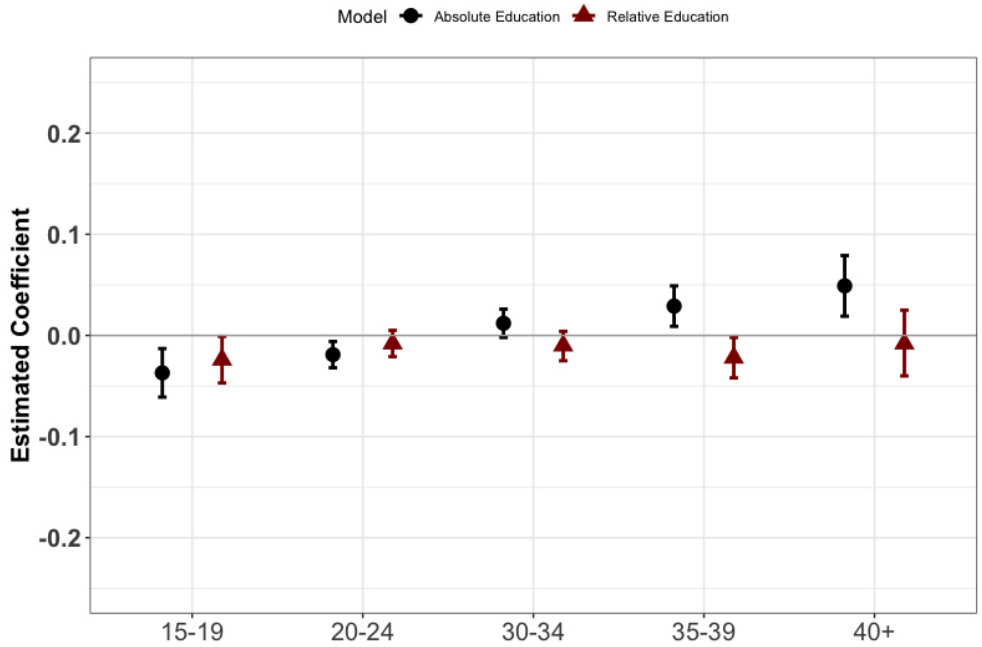
4. Results

This section describes results obtained for the two outcomes. In the first model, the outcome variable is the probability of reaching a tertiary degree, thus it focuses on children's educational attainment in *absolute* terms. The second model, instead, focuses on children's educational attainment in *relative* terms, and the outcome is the probability to be in the highest quartile of the educational distribution relative to the individual's birth-cohort group. By using a fixed-effects design, the effect of maternal age is not biased by other - measured or unmeasured - background factors.

Figure 1 shows the effect of maternal age at the birth of the child for each model. In line with the first hypothesis (**H1**), according to which there is a positive gradient in children's educational level by maternal age, when education is measured in absolute terms; I have found that children born to mothers who are younger than 25 years old at birth are less likely to acquire a tertiary degree relatively to children born to mothers who are 25-29 years old at birth (the reference category). Additionally, children born to mothers who are older than 29 years old at birth are more likely to acquire a tertiary degree relatively to children in the reference category.

This result for the French context is also in line with what has been found earlier in other contexts. The effect of mother's socio-economic background is ruled out because of the siblings fixed-effect design. Consequently, the positive association found can be explained by positive trends in educational expansion. Later born are more likely to acquire a tertiary degree because they can benefit from educational expansion processes.

Figure 1 – *Effect of maternal age on the probability to acquire a tertiary degree (black estimates), and the probability to be in the highest quartile of the educational distribution relatively by birth cohort (red estimates). The reference category are children born to mothers aged 25-29 years old at the birth of the child.*



Notes: Own elaborations on FQP 2014-2015 data, error bars indicate 95% confidence intervals. Models control for sex, birth order, and family fixed effects.

While it is not possible to test this mechanism by including children's year of birth in the model, which would be collinear with mother's age at birth in the sibling fixed-effects design (Barclay and Myrskylä 2016b), it is possible to analyse the effect of maternal age at birth on children's education measured in relative terms. This approach will account for the fact that over time the proportion of tertiary graduates tends to increase.

Results showed that, in line with the second hypothesis (**H2**), the positive gradient in children's education by maternal age substantially disappears. Children born to

teenager mothers, and to mothers who are 35-39 years old at birth are less likely to be in the highest level of the educational distribution, relative to children born to mothers who are 25-29 years old at birth. Hence, the effect of maternal age turns out to be curvilinear (inverse U-shaped), but mostly not statically significant.

Table 2 – *Full models results for the outcome of absolute education (left panel), and relative education (right panel).*

	Absolute Education			**Relative Education**		
	Coef.	95% Confidence Interval		Coef.	95% Confidence Interval	
Female (Ref. Male)	0,01	-0,002	0,017	0,01	0,002	0,021
Second or higher order (Ref. First born)	-0,03	-0,039	-0,017	-0,04	-0,054	-0,031
_cons	0,21	0,199	0,223	0,20	0,193	0,217
sigma_u	0,33			0,31		
sigma_e	0,31			0,31		
rho	0,52			0,49		

Notes: Own elaborations on FQP 2014-2015 data.

Table 2 shows full model results for the remaining covariates. In line with previous findings, females are more likely than males to acquire a higher level of education in both absolute and relative terms, even if the sex of the sibling it is not statically significant in the model of absolute education. Next, in line with the literature, results show that first born children are more likely to acquire a higher level of education rather than second or higher order children.

Moreover, Table 2 shows the "rho" parameter, which is known as the intraclass correlation, and it indicates the proportion of the variance which is due to U_i, the difference across mothers. This is about 52% in the model of absolute education, and 49% in the model of relative education, values which are pretty similar to what has been found earlier (Kalmijn and Kraaykamp 2005).

As robustness checks, I have run stratified models by mother's educational level. The fixed-effects design already accounts for mothers' socio-economic background, in its time-invariant aspects. However, it is plausible that among lower educated mothers there is higher variation in siblings' year of birth than higher educated mothers, since the latter tend to have shorter birth intervals (Cigno and Ermisch 1989; Bartus et al. 2013). Obviously, sample sizes change in the stratified analyses, affecting the magnitude of coefficients and confidence intervals, still, the main conclusions remain the same to those reported here (see Table A1 in Appendix).

5. Conclusions and discussion

As the timing of parenthood is steadily increasing in many Western countries, scholars have been interested in the consequences of later motherhood on children's educational outcomes. Previous studies have shown that, in several contexts, the

effect of maternal age on children's education is positive. The mechanisms behind this positive effect, however, are still unclear, since both socio-economic resources of mothers and positive period trends in education may explain this gradient.

In this study, I shed light on the association between maternal age and children's education by applying a paired siblings fixed-effects design on French data. Differently from previous studies, however, I also measure children's educational outcomes in relative terms, i.e., interpreting education as a *positional good*. Considering also the relative aspect of education allows to account for positive period trends in education and, as a result, it helps clarifying the association between maternal age and children's educational outcomes.

In line with *H1*, I have found that there is a positive educational gradient in children's education by maternal age also in the French context. This gradient cannot be linked to family characteristics of the siblings (e.g., mother's educational level, family size, etc.), which are time-invariant, and are addressed by the fixed-effects design. Thus, the most plausible explanation is that this positive gradient is related to time trends in educational expansion.

In line with *H2*, when children's education is measured in relative terms, the positive gradient by maternal age disappears. This occurs because the process of educational expansion is taken into account by measuring individuals' level of education in relation to the level that others in the same birth cohort have obtained. Overall, these results support the argument that maternal age at birth does not seem to have an effect *per-se* that is independent from positive period trends in education.

This study also presents a few limitations. For instance, due to data constraints, I could not analyse differences across all siblings in a family, given that the survey gathers information only on one sibling randomly chosen during the interview (INSEE 2015). Moreover, the level of education and the year of birth of the respondents' sibling are self-reported by respondents, as a result, measurement error could be higher than studies which rely on siblings' population register data (e.g., Barclay and Myrskylä 2016a). Still, the fact that the choice of the sibling is random prevents from biases derived by selecting a preferred sibling, which would be closer to the respondent in terms of age.

Next, the siblings fixed-effects design allows to control only for *time-invariant* unobserved family characteristics. While this is a great added value relatively to previous studies, the lack of time-varying information about household income and socioeconomic resources of the mother could somewhat affect the estimates. Still, additional stratified analyses by mother's education led to the same conclusions as those presented here.

Moreover, this innovative approach that combines a paired siblings fixed-effects design with a relative measure of education could be limited to fully account for the role of long-term trends in educational expansion, because the typical age-gap

between siblings is relatively small. Even so, results consistently show the importance of considering education as a *positional good*.

Overall, this study has contributed theoretically and methodologically to previous literature on the effect of maternal age on children's educational attainment. Previous studies have mentioned the role of educational expansion as possible mechanism for the positive effect of maternal age, ruling out the effect of measured and unmeasured family characteristics by also using a fixed-effects design. Still, none of previous studies has properly empirically tested the role of changes over time in the educational distribution.

This study showed that *relative education* matters. Period trends such as educational expansion may have unclear outcomes for children born to older mothers. For instance, if structural conditions are difficult, many qualified individuals have to compete on the labour market, thus educational expansion may be considered harmful to a later born. Educational expansion, differently from other positive period trends (such as medical progress, and related improvements in life expectancy), may have more ambiguous effects on children's outcomes.

Here, I have highlighted the importance of *relative education* when testing the effect of maternal age on children's outcomes. Future studies should focus more often on education as a *positional good* to determine the role of demographic behaviours in the intergenerational transmission of education.

Appendix

Robustness checks: Stratified analyses by mothers' level of education

Table A1 – *Effect of maternal age on absolute education (left panel), and relative education (right panel), stratified by mothers' educational level.*

Mothers with at least high-school diploma (N = 5160)	**Absolute**			**Relative**		
Maternal age at birth (Ref. 25-29)	Coef.	95% Confidence Interval		Coef.	95% Confidence Interval	
15-19	-0,08	-0,14	-0,02	-0,04	-0,10	0,02
20-24	-0,02	-0,05	0,01	-0,01	-0,04	0,02
30-34	0,01	-0,02	0,04	-0,03	-0,06	0,00
35-39	0,08	0,02	0,14	-0,02	-0,08	0,03
40+	0,10	-0,01	0,21	-0,06	-0,17	0,05
Mothers with a lower-secondary degree (N = 4894)	Coef.	95% Confidence Interval		Coef.	95% Confidence Interval	
15-19	-0,04	-0,08	0,01	-0,01	-0,06	0,03
20-24	-0,01	-0,03	0,02	0,00	-0,03	0,02
30-34	0,03	0,01	0,06	0,01	-0,02	0,04
35-39	0,04	0,00	0,08	-0,02	-0,06	0,02
40+	0,04	-0,03	0,10	-0,02	-0,10	0,05

Table A1 (cont.) – *Effect of maternal age on absolute education (left panel), and relative education (right panel), stratified by mothers' educational level.*

Mothers with no degree or a primary level (N = 5953)	Coef.	95% Confidence Interval		Coef.	95% Confidence Interval	
15-19	-0,01	-0,05	0,02	-0,01	-0,05	0,02
20-24	-0,02	-0,03	0,00	0,00	-0,02	0,02
30-34	0,01	-0,01	0,03	-0,01	-0,03	0,01
35-39	0,00	-0,03	0,02	-0,02	-0,05	0,00
40+	0,05	0,01	0,08	0,01	-0,02	0,05

Notes: Own elaborations on FQP 2014-2015 data.

References

BARCLAY K., MYRSKYLÄ M., 2016a. Advanced Maternal Age and Offspring Outcomes: Reproductive Aging and Counterbalancing Period Trends, *Population and Development Review*, Vol. 42, No. 1, pp. 69–94.

BARCLAY K., MYRSKYLÄ M., 2016b. Online Comment and Rejoinder, *Population and Development Review*, Vol. 42, No. 4, pp. 711–711.

BARTUS,T., MURINKÓ L., SZALMA I., SZÉL B., 2013. The effect of education on second births in Hungary, *Demographic Research*, Vol. 28, No. January, pp. 1–32.

BREEN R., ERMISCH J., HELSKE S., 2019. Educational reproduction in Europe: A descriptive account, *Demographic Research*, Vol. 41, No. December.

BREEN R., MÜLLER W., 2020. *Education and intergenerational social mobility in Europe and the United States*. Stanford, CA: Stanford University Press.

BUKODI E. GOLDTHORPE J.H., 2016. Educational attainment - relative or absolute - as a mediator of intergenerational class mobility in Britain, *Research in Social Stratification and Mobility*, Vol. 43, pp. 5–15.

CANTALINI S., GUETTO R., PANICHELLA N., 2020. Parental age at childbirth and children's educational outcomes: evidence from upper-secondary schools in Italy, *Genus*, Vol. 76, No. 1, p. 8.

CHOI S., TAIJI R., CHEN M., MONDEN C., 2020. Cohort Trends in the Association Between Sibship Size and Educational Attainment in 26 Low-Fertility Countries, *Demography*, Vol. 57, No. 3, pp. 1035–1062.

CIGNO A., ERMISCH J., 1989. A microeconomic analysis of the timing of births, *European Economic Review*, Vol. 33, pp. 737–760.

COHEN P., 2014. Parental Age and Cognitive Disability among Children in the United States, *Sociological Science*, Vol. 1, pp. 102–110.

FISHMAN S.H., MIN S. 2018. Maternal Age and Offspring's Educational Attainment, *Journal of Marriage and Family*, Vol. 80, No. 4, pp. 853–870.

FORMATION ET QUALIFICATION PROFESSIONNELLE (FQP), 2015, INSEE [producers], ADISP [distributor].

GRÄTZ M., WIBORG Ø.N., 2024. Parental ages and the intergenerational transmission of education: evidence from Germany, Norway, and the United States, *European Societies*, pp. 1–28.

HIRSCH F., 1976. *Social Limits to Growth*. Harvard University Press.

INSEE. 2015. *Enquête Formation et Qualification Professionnelle 2014-2015, Méthodologie*
https://www.insee.fr/fr/metadonnees/source/operation/s1563/documentation-methodologique (accessed 17/07/2024).

KALMIJN M., KRAAYKAMP G., 2005. Late or later? A sibling analysis of the effect of maternal age on children's schooling, *Social Science Research*, Vol. 34, No. 3, pp. 634–650.

KEIDING N., ANDERSEN P. K., 2016. Linear effects of maternal age and period trends cannot be distinguished, *Population and Development Review*. doi: 10. 1111/padr.12025.

KRAVDAL Ø. 2019. Taking birth year into account when analysing effects of maternal age on child health and other outcomes: The value of a multilevel-multiprocess model compared to a sibling model. *Demographic Research*, Vol. 40, pp. 1249–1290.

MOOD C. 2010. Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It, *European Sociological Review*, Vol. 26, No. 1, pp. 67–82.

MYRSKYLÄ M., BARCLAY K., GOISIS A. 2017. Advantages of later motherhood. *Der Gynäkologe*, Vol. 50, No. 10, pp. 767–772.

POWELL B., STEELMAN L.C., CARINI R.M., 2006. Advancing Age, Advantaged Youth: Parental Age and the Transmission of Resources to Children, *Social Forces*, Vol. 84, No. 3, pp. 1359–1390.

TIMONEDA J.C., 2021. Estimating group fixed effects in panel data with a binary dependent variable: How the LPM outperforms logistic regression in rare events data, *Social Science Research*, Vol. 93, p. 102486.

TRIVENTI M., PANICHELLA N., BALLARINO G., BARON, C., BERNARDI F., 2016. Education as a positional good: Implications for social inequalities in educational attainment in Italy, *Research in Social Stratification and Mobility*, Vol. 43, pp. 39–52.

Alessandra TRIMARCHI, University of Messina, alessandra.trimarchi@unime.it

DETERMINANTS OF STUDENTS' PROFICIENCY: SCHOOL FEATURES AND PERSONAL CHARACTERISTICS VS SES IN KOSOVO

Luigi Maria Solivetti

Abstract. This study tests whether the school context and students' personal characteristics improve male and female students' proficiency irrespective of the effects of families' socioeconomic inequalities. The object of the investigation is a country presenting relatively poor educational results: Kosovo.

The study used the PISA 2022 dataset and focused on the math scores of female and male Kosovo students. The statistical analyses employed balanced repeated replication regression models by students' gender.

The results show that SES is a significant but not unmitigated determinant of the students' proficiency. School features affect the students' scores, while students' personal features have an even larger impact on them. Girls' math scores are equivalent to boys' – which is uncommon – but are differently affected by the proficiency determinants.

These findings suggest that school features and personal traits change students' proficiency even when pupils, on average, come from a context negatively affecting their capabilities and present relatively low proficiency. Moreover, averagely lower socioeconomic conditions and the probable persistence of traditional gender roles are not synonymous with the gender gap in proficiency. Still, girls and boys are different in how they respond to proficiency factors, and therefore gender-specific policies are advisable to improve their proficiency.

1. Introduction

The Programme for International Student Assessment (PISA) is an OECD survey designed to assess worldwide education systems. PISA evaluates how well 15-year-old students can use their reading, mathematics, and science knowledge to tackle real-world problems. The 2022 survey (the most recent one) included 80 countries. For each student, PISA recorded more than one thousand variables measuring his/her proficiency, personal opinions and attitudes, and family socioeconomic characteristics. Moreover, PISA collected two other datasets containing hundreds of variables: one about the schools' features and the other about teachers' qualities, opinions, and attitudes.

Since the first PISA survey in 2000, numerous studies have tried to predict students' proficiency scores using micro and macro determinants. Scholars have particularly focused on three domains: the student's socioeconomic status (hereafter

SES), the school features and teachers' qualifications and attitudes, and the student's personal characteristics.

In the present study, we intend to test the impact of determinants belonging to the said three domains on the student's proficiency in a country whose students scored relatively low in the PISA 2022 survey: Kosovo. There has been a high interest in countries presenting poor proficiency results (Okpala, Okpala and Smith, 2010; Woldemichael, Semela and Tulu, 2023). Indeed, identifying the factors associated with lower educational results is crucial to reducing international inequalities in knowledge and development. Besides, comparing SES determinants of proficiency to other factors would be potentially more constructive when the country under investigation is a relatively poor one, such as Kosovo, which has the lowest per capita GDP in Europe (in 2024, 16.7K). Indeed, it would be essential to investigate whether, in less developed countries, non-material proficiency factors counterbalance the effects of an averagely unfavourable SES.

Moreover, it would be stimulating to check the impact of gender on proficiency in an adverse socioeconomic context. All over the world, girls perform, on average, better than boys in reading but worse in math and science, the branches of learning leading to better-paid and prestigious jobs. We expect a substantial gender gap in educational outcomes in less developed societies, owing to the persistence of traditional gender roles. There is, indeed, an inverse cross-country association between GDPs per capita in PPP and the gender gaps in math (Anghel, Rodríguez-Planas and Sanz-de-Galdeano, 2020). However, this association evaporates when we consider only non-OECD countries. PISA 2022 data show that none of the few countries presenting neatly higher math scores for girls belongs to the group of most advanced countries. In addition, regardless of the gender gap in proficiency, it would be worth investigating other potential dissimilarities revolving around gender, such as the different impacts that the same factors could have on girls' and boys' scores.

2. Review of the literature and present study's hypotheses

Current literature has shown great interest in the students' SES. Cross-country, a close correlation (r = 0.73) exists between average GDP per capita and students' average proficiency (World Bank, 2023). However, the correlation for the lowest GDP tertile countries is weaker than that for the highest tertile countries: r = 0.21 vs 0.32. The relationship between students' SES and proficiency is also positive at the micro level. For all the students in PISA 2022, we found a correlation of 0.48 between SES and math. The correlation is distinctly weaker for the students in the group of countries with lower socioeconomic conditions, although in these countries, the students' SES relative standard deviation is – as expected – higher. This suggests that, in underprivileged countries, the relationship between SES and proficiency is not only weaker but also less affected by inequalities in SES.

As a trend, a higher SES would entail more expenditures in goods and opportunities capable of increasing the children's human capital. This, in turn, would favour better educational results (Okpala, Okpala and Smith, 2010). In a higher SES house, there would be more traditional cultural goods, such as newspapers, magazines and books, and more new tools, such as Internet access, computers, and cell phones: tools that tend to increase students' performances (Schleicher, 2023: 34). High-SES students would also benefit from the opportunities created by their parents' expenditures on non-material goods: expenditures on their children's learning and skills, such as private tutoring, better schools etc. Besides, higher SES usually entails a higher level of parental involvement in their children's studies as well as attitudes that encourage children to pursue higher educational achievements. In turn, students from low-income families would be negatively affected not only by their inadequate learning environments at home but also by the pressure to contribute to household income (Post and Pong, 2009; Cooper and Miralay, 2022).

Ultimately, literature has ascribed the impact of the family's SES on the student's proficiency to various specific factors. Moreover, no consensus has emerged about how to measure the various elements underlying the SES acronym: social class, poverty, affluence, power, and culture (Pokropek, Borgonovi and Jakubowski, 2015). Numerous variables (from income to possessions, education, occupation and attitudes) and also various combinations of variables have been used for this purpose. For instance, PISA 2022 used a composite index (ESCS) derived from three indicators: parents' education, parents' occupational status and home possessions.

Secondly, the student's proficiency is associated with his/her experiences at school. Teacher-related factors – such as staff shortages, teacher absenteeism, teachers' qualifications, commitment, and attitudes – would impact the student's learning (Ammermüller, 2004). The same would occur with schools' characteristics, such as rural/urban, private/public, their availability of libraries and other material resources, and their strictness (Shukakidze, 2013). The substandard level of school infrastructures and teachers' working conditions and salaries seems to have affected Kosovari students' proficiency (Lutfiu and Hoxha, 2024). Moreover, the school's social environment would also affect the student's proficiency: the school's social context generates interactions that, in turn, shape the student's perception of self and sense of belonging. For instance, facts such as being bullied negatively influence students' proficiency (Schleicher, 2023).

In lower-income countries, school-related factors are expected to have a deeper impact on students' proficiency. In wealthier countries, students can benefit from educational materials and a good learning environment at home, as well as from private tutoring. In poorer countries, instead, school is the primary source of the student's learning opportunities (Shukakidze, 2013). Ultimately, the chances of educational success for students from an underprivileged background often depend

on a good teacher and a good school. However, one should consider that a school's quality is not independent of the student's SES. As we already remarked, higher-SES families tend to cherry-pick their children's schools. Therefore, an analysis of proficiency based on the school's and teachers' qualities would miss the indirect influence of the student's SES.

Thirdly, students' personal features cannot but influence their school performances. Non-cognitive attributes, namely skills and characteristics concerning the student's attitudinal, behavioural, emotional, and motivational dispositions, seem particularly momentous. Hattie (2009) found that "engagement and motivation", "self-concept", and "anxiety" were the most relevant ones. In turn, other scholars emphasised the relevance of "performance self-efficacy", e.g. the perception of one's performance capability, in predicting educational scores (Lee and Stankov, 2018). Other individual factors can influence proficiency. For instance, student labour can negatively impact educational scores (Post and Pong, 2009).

Given all of this, we advance the following hypotheses:

H1. Even within a country with an averagely low proficiency, there is substantial variance in the students' scores: the precondition for further investigations.

H2. In such a country, it is possible to identify a relationship between the family's SES and the student's proficiency, but other factors also correlate with proficiency.

H3. School-related factors and the student's personal characteristics affect proficiency, acting as forces partially independent of SES.

H4. SES, school-related factors, and the student's personal characteristics can affect the proficiencies of girls and boys differently.

3. Methods

The present study is based on the PISA 2022 data (OECD, 2023) and, in particular, the Kosovo students' sample. We preliminarily merged the Student Questionnaire with the School and the Teacher Questionnaires. Kosovari students' proficiency has been our response variable. We focused on the scores in mathematics, the field of knowledge particularly apt to pass through cross-country differences in culture and language. For each student, PISA provides not a single score but ten plausible values as well as weights. To conduct statistical analyses on PISA proficiency data, it is therefore necessary to use specific procedures, as shown by the notation (1) for our main regression models:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \, [fw(w_i), rw(w_{1,i}, \dots, w_{nth,i}), vce(brr_i), fay(0.5)] \quad (1)$$

where y = plausible values for student i; x = explanatory variable; ε = error; *fw* = final survey weight incorporating any student-level trimming; *rw* = replicated weights (N = 80) for the balanced repeated replication (BRR) procedure taking into account sampling and non-response errors; *vce* = variance estimator for the said

procedure; *fay* = Fay's adjustment that perturbs the replicated weights by a specific per cent. The final results are the mean of the regressions of the ten plausible values. To conduct our analysis, we used the *repest* and *pisatools* applications for Stata.

As for the explanatory variables, we had to make a strict selection, owing to the vast number of measures concerning students' familial and personal characteristics, schools' features, and teachers' qualifications and attitudes. Firstly, we excluded all the variables without enough valid observations. We set the minimum threshold at 3500 valid observations, considering that we had to split the dataset into two parts, one for female students and the other for male students, and that using multiple variables in the regression models would inevitably multiply the final number of row-wise missing observations. Then, we excluded the redundant variables. After that, we checked the summary statistics to ascertain central tendencies and variances. Finally, we used a backward stepwise estimation procedure to filter the variables.

Literature on the determinants of proficiency has usually focused on a specific domain (e.g. the student's SES, or schools' and teachers' characteristics). Alternatively, it has analysed the effect on proficiency jointly made by determinants belonging to various domains. In the present study, we intend to follow a different method. Firstly, we will measure the effect of the family's SES on the student's proficiency. Then, we will analyse the impact of factors belonging to the school domain and to the student's personal characteristics. However, we know that the quality of the school is not independent of the student's SES, and there are links between the students' features (e.g. time for studying) and their familial background. Therefore, to avoid underestimating the direct and indirect effects of SES on proficiency, we will analyse the impact of the school and the student's personal characteristics on the proficiency variance unexplained by familial background. By doing this, we expect it would be possible to measure the impacts of the school and student's characteristics *at parity of family and socioeconomic determinants*.

4. Results

In the PISA 2022 tests, Kosovo's students obtained rather analogous scores in mathematics, reading, and science. Concurrently, those scores were markedly lower than the average scores registered for all the world countries, the European countries and even the Balkan countries (Figure 1).

Figure 1 also shows that the Kosovari girls' scores in reading were neatly higher than boys', which is in tune with the trend in the world's countries. Kosovo's girls also scored better than boys in science, while, on average, in the world, girls' scores are slightly lower than boys'. The gap in math between Kosovo's boys and girls, in turn, was almost non-existent, in contrast with what occurs elsewhere. Figure 2 shows two other facts. The math ten plausible values distributions follow the same pattern, not too dissimilar from a normal distribution. Therefore, Kosovari students'

math scores were widely different: their coefficient of variation (σ/μ) is 0.18, not much lower than all Europe's, which is 0.20. In conclusion, nothing hinders a statistical analysis of these distributions to identify the determinants of the variance in the students' scores. Concurrently, the girls' and boys' distributions of scores are not the same. While kurtosis is similar (3.1 vs 3.3), boys' scores show a greater positive skewness (0.52 vs 0.29). This gap is also found in other countries where, however, on average, it is smaller.

Figure 1 – *Kosovari boys' and girls' average proficiency scores compared to some world regions*'.

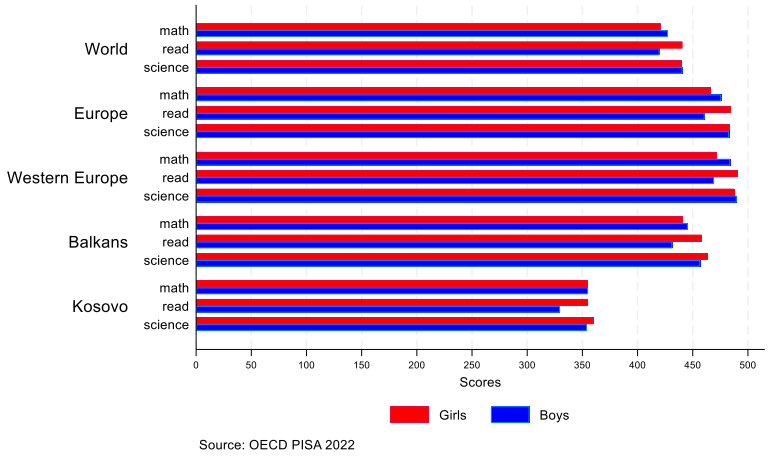
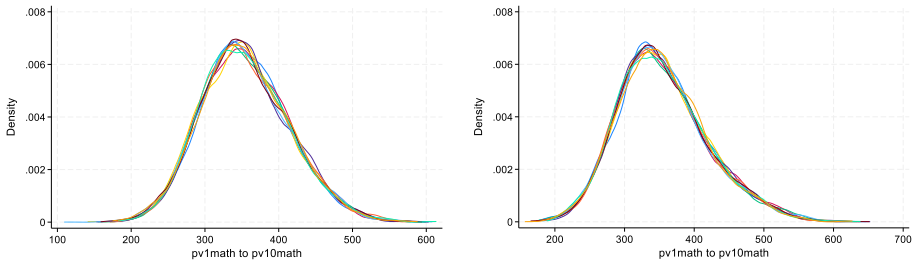


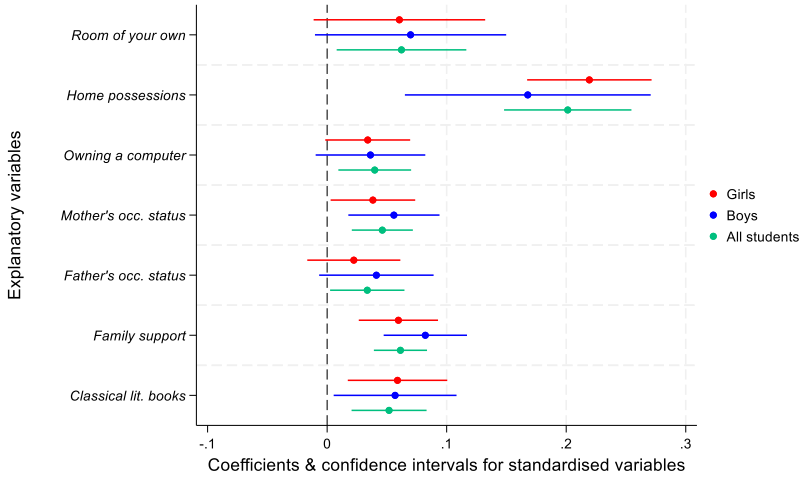
Figure 2 – *Kernel density estimates of Kosovari students' math plausible values.*



Girls Boys

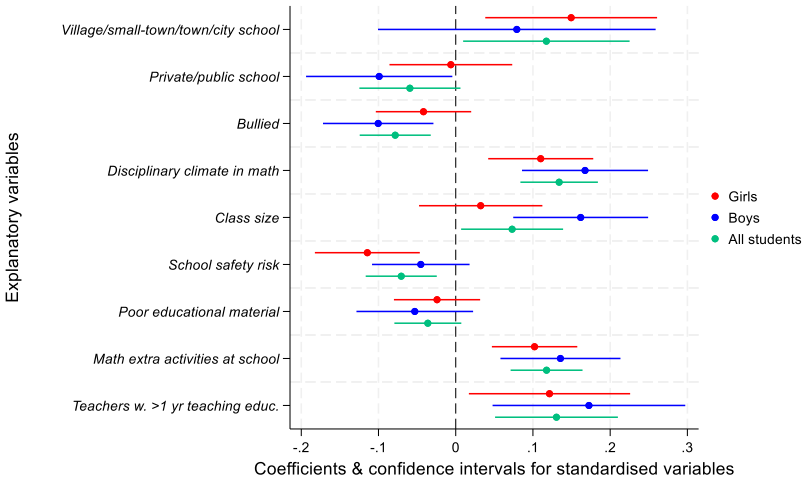
Note: kernel = Epanechnikov; weighted values

Figure 3 – *Balanced repeated replication regression by gender of Kosovari students' math scores on their families' socioeconomic determinants.*



Note: Girls N = 1706, R^2 = 0.14; Boys N = 1389, R^2 = 0.12; All students N = 3095, R^2 = 0.12

Figure 4 – *Balanced repeated replication regression by gender of the residuals of Kosovari students' math scores (from Figure 3 regression) on school determinants.*



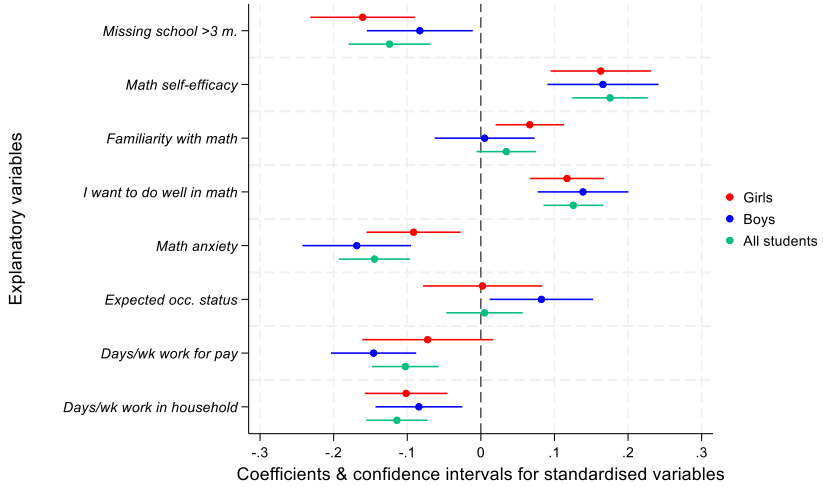
Note: Girls N = 1141, R^2 = 0.08; Boys N = 829, R^2 = 0.13; All students N = 1970, R^2 = 0.09

Figure 3 shows the results of the analysis of the socioeconomic determinants of Kosovari students' proficiency. All the seven explanatory variables directly correlate

with the math scores and are significant when considering girls and boys together. The father's occupational status is less impactful than the mother's. The impact of "home possessions" is particularly large in the case of girls. The models explain 12% to 14% of the proficiency variance.

Figure 4 shows the regressions by gender of the residuals of students' math scores (from the previous regression of Figure 3) on school determinants. The degree of urbanisation, the disciplinary climate in mathematics, the extra activities at school, and the teachers' education directly impact the students' proficiency. The same occurs with class size, which directly correlates with urbanisation. Instead, being bullied and the school safety risk are associated with lower scores. When we split the sample into two groups, girls and boys, it emerges that school urbanisation is significant only for girls, while being bullied significantly reduces only boys' proficiency. School safety risk significantly impacts only girls' proficiency, while being in a public school significantly reduces boys' proficiency. All in all, the school and teachers' characteristics explain 8% to 13% of the variance in proficiency left unaccounted by the SES determinants.

Figure 5 – *Balanced repeated replication regression by gender of the residuals of Kosovari students' math scores (from Figure 3 regression) on their personal determinants.*



Note: Girls N = 1307, R^2 = 0.15; Boys N = 1019, R^2 = 0.19; All students N = 2326, R^2 = 0.16

Figure 5 shows how the students' personal features impact the residuals of the regression of math scores on the family's socioeconomic determinants. Math self-efficacy and the students' determination to do well in math correlate with higher scores. For instance, every increase in the students' determination to do well in math

(4-level variable) averagely adds 14 points to their (residual) math scores. Instead, math anxiety negatively affects proficiency scores, and the same occurs with missing school for an extended span of time, working for pay and working in the household after school. We notice some gender differences: the expected occupational level is significant only for boys, while familiarity with math is significant only for girls. Besides, the anxiety impact on proficiency is smaller in the case of girls. In turn, working for pay only significantly impacts boys' proficiency. Apropos of this, we also found that, in Kosovo, male students work for pay significantly more than female students. Kosovari boys also work significantly more than boys in Western Europe. All in all, in Kosovo, personal characteristics explain 15% to 19% of the variance left unexplained by the socioeconomic determinants.

5. Discussion

Despite the comparatively low level of proficiency, the sample of Kosovo's students considered by PISA 2022 presents a well-differentiated distribution of their math scores. The distribution of the girls' scores shows a right tail less extended than the boys'. This feature is shared, on average, by all the countries but is more marked in Kosovo. This means that, among Kosovari girls, the incidence of top scorers is smaller.

As to the role of SES, our results confirm the findings of previous scientific literature as to its significant effect on school proficiency. Our regression model showed that SES measures explained 12% of the Kosovari students' variance in math scores. This figure is approximately halfway between the figures calculated by PISA for all the world countries, where SES accounts for ~5% to ~25% of the national variance in math scores. However, the variance explained by our model is two times the variance (~6%) explained by the PISA basic index of economic, social and cultural status (ESCS) when applied to Kosovo students.

This gap seems to derive from the determinants selected in the present study. While the importance of "home possessions" is evident – and this index is part of the wider ESCS index – "a room of your own" would measure the role of a student's *sacred space*, where one feels safe and comfortable. The specific relevance, among home possessions, of "a computer at home" derives from the opportunities for learning offered by such a device. Both the mother's and father's occupational status would measure the impact of class as something only partially overlapping with material well-being. However, it is noteworthy that the mother's occupational status emerged as more impactful than the father's: this confirms the relevance of the mother-child relationship for the younger generations' upbringing. In turn, classical literature books at home would gauge the family's cultural background as distinct from material well-being. Finally, "family support" would measure the social capital

on which the student can rely at home: something that can be independent of the family's economic status.

As for gender differences, the most relevant one is the larger impact of "home possessions" on the girls' proficiency. Because "home possessions" *per se* can hardly cause these gender differences, we believe that more copious "possessions" mirror non-observed familial socio-cultural factors favouring female education.

Despite these results, the variance explained by our SES model (12%) clearly leaves most of the Kosovari students' variation in proficiency unexplained.

The regression models based on the school's features explained a share of this residual variance, from 8% to 13%. The results highlight the relevance of both the teaching and the school context for the student's proficiency. As for teaching, the impact of variables such as longer teaching education, "disciplinary climate", and extra activities at school emerged. As for the school context, one can infer its importance from the impact of the school's urbanisation level, the fact of being bullied, and the school's safety risk. These models also detected noteworthy gender differences that helped better understand the school's role. Proficiency elasticity to safety risks is higher for girls, suggesting they are more exposed to such risks. Girls' proficiency is also more affected by the school's urbanisation level: this could be ascribed to the fact that the rural context would probably discourage female education. Lastly, being bullied is significant only for the boys' proficiency: this outcome draws attention to gender differences in interpersonal relations.

The regression models revolving around the student's personal characteristics explained a further share (from 15% to 19%) of the variance left unanswered by the SES variables. Math scores proved to be particularly elastic to variables concerning the student's self-confidence and commitment to study, such as math self-efficacy and determination to do well in math. Instead, working for pay and working in the household negatively impacted proficiency. Unsurprisingly, the same occurs with anxiety, inversely correlated with self-confidence as measured by math self-efficacy ($r = -0.34$). Anxiety about one's math proficiency had a larger impact on boys' scores. This conflicts with the fact that, on average, Kosovo's girls assert a higher level of anxiety than boys (0.34 vs 0.13): a difference common to the students of most countries. This contradiction suggests that either girls overrate their anxiety or this last is intrinsically different from boys' as to its effect on learning. In turn, girls' expected occupational status is higher than boys' but is less correlated with their determination to do well in math and has a non-significant impact on their proficiency, as if girls were less confident of their chances of obtaining a better status through better scores.

Ultimately, this study concerned students from the country with the lowest per capita income in Europe. This disadvantaged socioeconomic background seems to have affected the Kosovari students' capabilities, leading to results in proficiency

that, on average, are the lowest in Europe. However, we saw that, on an individual level, although the SES determinants impacted the students' proficiency, they left most of the proficiency variance unexplained.

By focusing on this variance unexplained by the SES determinants, this study was able to analyse the impact of school-related factors and students' personal characteristics on proficiency, bypassing the problem of the indirect effects of SES on proficiency. The regressions showed that schools' features reduced the amount of this unexplained variance, while the students' personal characteristics had an even larger impact on it. These findings attest that – despite a familial socioeconomic context that, on average, was unfavourable – school features and personal characteristics can change students' proficiency. The school emerged not only as the setting in which the family's SES is translated into proficiency – as suggested by current literature – but also as an institution capable of reworking this translation. This finding has important implications for government policies. The role played by the student's personal characteristics, in turn, seems crucial. In converting the students' SES into a personal realisation – as school proficiency is – these characteristics emerged from the present study as even more impactful than that SES. These findings represent an additional contribution to the current literature also because the impacts of school features and students' personal characteristics were calculated on the basis of proficiency values net of family's socioeconomic status.

As for the gender issue, the present results show that Kosovari girls' math scores were substantially equal to boys'. This outcome distinguishes Kosovo from most countries in the world, including those in Europe. Therefore, relatively lower socioeconomic conditions and the probable persistence of traditional gender roles are not synonymous with the gender gap in proficiency. Concurrently, Kosovari boys' proficiency is negatively affected by their greater involvement in work for pay. This fact distinguishes them from West European boys, who work significantly less and score in math neatly more than girls.

All in all, performances in educational achievements are much less closely associated with gender than with SES, personal characteristics and also school features. However, this does not mean that the gender issue is inconsequential. By analysing separately girls and boys, it has been possible to show that their proficiency scores differ in how they respond to factors commonly affecting educational outcomes. While boys' proficiency is particularly affected by being the victim of bullying, girls' is reactive to safety at school and urbanisation. Moreover, the girls' group is comparatively short of top scorers: a fact expected later to affect girls' chances of access to hard-to-obtain positions. These results suggest that any policy meant to improve educational outcomes should include gender-differentiated provisions in tune with the different impacts of specific conditions.

References

AMMERMUELLER A. 2004. PISA: *What makes the difference? Explaining the Gap in PISA Test Scores Between Finland and Germany*. Mannheim: Zentrum für Europäische Wirtschaftsforschung,

ANGHEL B., RODRÍGUEZ-PLANAS N., SANZ-DE-GALDEANO A. 2020. *Gender Equality and the Math Gender Gap*. SSRN Paper. Rochester, N.Y.: SSRN.

COOPER J.J., MIRALAY F. 2022. The Examination of Poverty Effect on Student Performance in Low-Income Countries, *Near East University Online Journal of Education*, Vol. 5, No. 2, pp. 37–46.

HATTIE J. 2009. *Visible Learning: A Synthesis of over 800 Meta-Analyses relating to Achievement*. London: Routledge.

LEE J., STANKOV L. 2018. Non-Cognitive Predictors of Academic Achievement, *Learning and Individual Differences*, Vol. 65, pp. 50–64.

LUTFIU B., LUTFIU HOXHA L. 2024. Socioeconomic Status of Teachers and Its Impact on Teaching Quality, *European Journal of Education and Pedagogy*, Vol. 5, No. 2, pp. 52–58.

OECD. 2023. *PISA 2022 Database*. Paris: OECD. https://www.oecd.org/en/data/datasets/pisa-2022-database.html

OKPALA C.O., OKPALA A.O., SMITH F.E. 2001. Parental Involvement, Instructional Expenditures, Family Socioeconomic Attributes, and Student Achievement, *The Journal of Educational Research*, Vol. 95, No. 2, pp. 110–115.

POKROPEK A., BORGONOVI F., JAKUBOWSKI M. 2015. Socio-Economic Disparities in Academic Achievement, *Learning and Individual Differences*, Vol. 42, pp. 10–18.

POST D., PONG S-L. 2009. Student Labour and Academic Proficiency in International Perspective, *Internat. Labour Review*, Vol. 148, No. 1-2, pp. 93-122.

SCHLEICHER A. 2023. *PISA 2022 Insights and Interpretations*. Paris: OECD.

SHUKAKIDZE B. 2013. Comparative Study: Impact of Family, School, and Students Factors on Students Achievements in Reading in Developed (Estonia) and Developing (Azerbaijan) Countries, *International Education Studies*, Vol. 6, No. 7, pp. 131–143.

WOLDEMICHAEL B., SEMELA T., TULU A. 2023. Mathematics Self-Efficacy, Self-Concept and Anxiety in Relation to Mathematics Achievement: The Case of Kafa Zone School Adolescents, Ethiopia, *F1000Research*, Vol. 12, 1494.

WORLD BANK, 2023. *Average Learning Outcomes* [dataset].

Luigi M. SOLIVETTI, Sapienza University of Rome, luigi.solivetti@uniroma1.it

AN ISTAT SURVEY ON CHILDREN AND YOUNG PEOPLE: ANALYZING FEEDBACK FROM RESPONDENTS[1]

Cinzia Conti, Gabriella Fazzi, Barbara D'Amen, Marco Rizzo

Abstract. Since 2014, Istat has been carrying out surveys in which young people are protagonists. The new edition of the Survey "Children and young people: behavior, attitudes, and future projects" (year 2023) was aimed at collecting information on the living conditions of Italian adolescents and adolescents from a migratory background aged between eleven and nineteen years. In an era of crisis and profound social change, such as the current post-pandemic period, there is a need to further enrich the statistical information collected and analyzed for the implementation of policies that enhance the human capital of the very young and improve their well-being. One initial goal of this contribution is to discuss the innovations introduced in the statistical process to reach young people. Specific strategies adopted by Istat include a web questionnaire optimized for any kind of device (including smartphones), the use of QR code to directly access the form and the availability of the questionnaire in Italian and nine other languages. The main objective of the contribution is to analyze, through textual analysis, the suggestions gathered in an open-ended question placed at the end of the survey. We will use statistical models like Analysis of Specificities, Sentiment Analysis and Cluster Analysis to examine term occurrences, identify the main concepts, and retrieve semantic relationships between them.

1. Introduction

Social science researchers are increasingly surveying young adolescents, but there is limited methodological knowledge on this age group. Much of the existing guidelines come from studies and theories focused on adults (Omrani et al., 2019; de Leeuw, 2011). In 2023 Istat carried out the survey on children and young people: forms of behavior, attitudes and future plans making them the primary respondents (without proxies). The literature, in fact, encourages the use of questionnaires for children starting from the age of eleven, with suitable precautions put in place (Borgers et al., 2000). Many innovations have been introduced in order to minimize the statistical burden on children and to maximize data protection and confidentiality.

[1] This article is the work of the authors. Paragraph 1 was written by Cinzia Conti, paragraph 2 was written by Cinzia Conti and Marco Rizzo, paragraph 3 was written by Marco Rizzo, paragraphs 4, 5, 6, were written by Barbara D'Amen and paragraph 7 was written by Gabriella Fazzi. Conclusions were written jointly by the authors.

The survey is based on a "light" questionnaire[2] that can be comfortably filled out even on a smartphone. The questionnaire includes sections on: demographic information; school life; citizenship and identity; social relationships; opinions about the future; and opinions about men and women. Given the innovative nature of the process, some questions about the respondents' opinions on the survey and the structure of the questionnaire were included at the end. In particular, there is a final open-ended question in which we asked, "*Do you have any suggestions for us for improving the questionnaire?*". Istat was indeed interested in gathering advice directly from their young respondents to enhance the data collection process. In this contribution, we wish to focus specifically on that open-ended question. We are interested in studying both the factors that can influence the propensity to answer (through regression models) and the content of the answers (textual analysis). The literature suggests that open-ended questions in surveys allow respondents to freely express their opinions, adding depth to the results and often providing more reliable measurements than closed questions. However, they require better cognitive and communication skills, which can affect response reliability. Open-ended questions also have a higher likelihood of "don't know" answers and non-responses, especially among young adolescents (Omrani et al., 2019; de Leeuw, 2011).

The first general hypothesis is that technological innovations may make surveys more appealing to young people. This paper aims to present the opinions of young people on the proposed questionnaire, the methods of data collection, and the topics introduced in the form.

The second – and more specific – hypothesis is that the socio-economic characteristics of the respondents and their families, as well as their relational networks, can influence both the propensity to respond to open-ended questions and the content of their answers.

2. Data and methods: a "smart" inclusive survey

Istat started conducting surveys on young people in 2015 with a survey on the integration of second generation migrants. In 2021, it conducted a survey focused on the issues faced by students aged eleven to nineteen during the COVID-19 pandemic. The new edition of the Survey on children and young people (year 2023) aims to collect information on behaviours, attitudes and living conditions of Italian children and children from a migratory background (both born in Italy and immigrants at a very young age) aged between eleven and nineteen (Istat, 2024). The aim of the survey is to highlight both weaknesses and the strengths of the new generations, to allow for a better valorisation of their energies and abilities in terms of policies and

[2] https://www.istat.it/fascicoloSidi/1542/Questionario%20italiano%20(Facsimile).pdf

actions. The research was set up, right from the early planning stages, with adolescents at the centre as active and participating subjects. Children represent a fundamental social subject for building the future and in an era of crisis and profound social change, there is a need to further enrich the statistical information collected and analysed for the implementation of policies that enhance the human capital of the young and to improve their well-being (Conti et al., 2024; UNECE, 2023).

The survey was conducted between 2 October and 20 December 2023 and involved a representative sample of almost 20,500 Italian young people and almost 18,000 young foreign residents in Italy[3].

The survey was carried out using the CAWI technique (Computer Assisted Web Interviewing)[4]. Among the innovations introduced by Istat in this data collection dedicated to young people it is essential to note that the questionnaire could be taken by scanning a QR code. It could also be submitted via PC, smartphone or tablet by accessing the link reported and typing the access code printed in the informative letter. Most of the sample preferred to fill out the questionnaire using a smartphone or a tablet rather than a PC. There is a difference between foreign nationals and Italians: 74.9% of Italians completed the questionnaire via smartphone or tablet compared to 83.1% among foreign nationals. The possibility of accessing via QR code favoured the use of these devices over PCs. Among those who answered through smartphone or tablet, 86.7% used the QR code. Instead, in the 2021 edition of the Survey on children and young people, where the questionnaire could only be accessed through a link, just 49 percent of users filled it out via smartphone or tablet. The introduction of the access via QR code could boost the number of respondents, especially for foreign nationals or for those who do not own a PC.

The survey is based on the consideration that the younger generations are "digital natives" (Sadiku et al., 2017). In the design of the survey internet, social media and smartphones play a central role. The questionnaire is moderately light: it takes about 23 minutes to fill out.

Another important aim of the survey was to be "inclusive" and also to collect information about the most vulnerable groups. To meet the language difficulties of young immigrants, for the first time in the history of Istat, the web questionnaire was available not only in Italian but in nine other languages: Albanian, Arabic, English, French, German, Mandarin, Romanian, Slovenian, Spanish and Ukrainian.

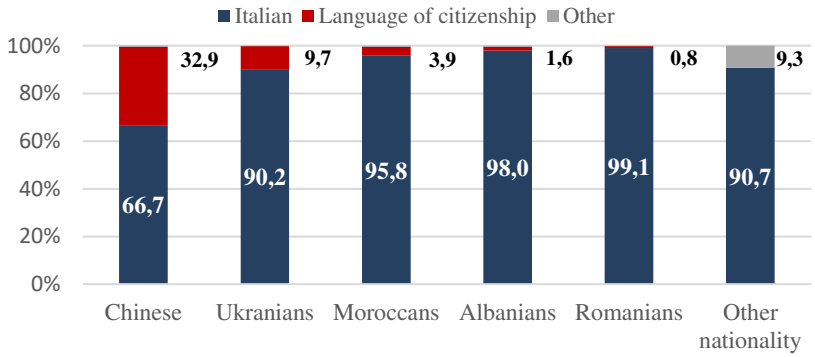
Apart from German that is spoken by national minorities in Italy, Mandarin was the most popular foreign language to be chosen, English was the second and the third

[3] The number of children selected based on optimal allocation was set, with some margin for over-sampling, at approximately 108,000. Stratification was defined by crossing the categories of region, age groups, and citizenship, resulting in a total of 588 levels (Istat, 2024).

[4] The decision to introduce the CAWI technique in population surveys makes it possible to contain the costs of public statistics and to exploit the potential offered by technology to capture segments of the population that are increasingly elusive compared to traditional techniques. (Istat, 2017)

was Ukrainian. About 1,700 foreign respondents filled out the questionnaire in a language other than Italian. Figure 1 shows the questionnaire language selected by respondents with respect to major citizenships. 32.9% of young people from a Chinese background decided to fill out the questionnaire in Mandarin Chinese (more than 800 questionnaires). This fact evidences low integration among young people from a Chinese background, even though more than 87% of the Chinese youth in the sample were born in Italy. Another interesting result is that just 9.7% of Ukrainian respondents selected Ukrainian. This fact is explained by the conflict in Ukraine. In our sample 301 Ukrainian children arrived in Italy for the first time in 2022, compared to an average of 73 new entrants per year for the previous nine years.

Figure 1 – *Distribution (%) of the respondents (11-19 year old) by selected language for the questionnaire and citizenship, year 2023.*



Istat, Survey on children and young people 2023

These results show us that giving respondents the opportunity to fill out a questionnaire in their preferred language facilitates the response of poorly-integrated communities. It also facilitates responses in emergency cases such as a forced migrant community in a country with a different language.

As previously mentioned, the propensity to answer was studied through logistic regression, where the dependent variable represents whether or not the respondents answered the open-ended question. The individual characteristics of the respondents were examined using structural variables: gender, territorial breakdown, citizenship, and age group. Socio-economic characteristics were assessed through the variables: shared bedroom; and mother's educational qualification[5]. Information regarding relational networks was measured using the variables: bullying indicator; school

[5] The mother's educational qualification is as an indicator of the family's socioeconomic status. We preferred this indicator over personal perceptions of their family's economic condition because the latter was not significant and to avoid issues with correlations between independent variables.

performance; frequency of seeing friends; and average hours a day on social media. All variables were derived from questions put to the respondents. The textual analysis, addressed in paragraph four, examines: the content of the answers; the techniques; lexical analysis; sentiment analysis; analysis of positive specificities; and cluster analysis.

3. Factors influencing the propensity to comment on the survey and the questionnaire

To support the textual analysis that will explore the suggestions shared through the above noted open-ended question, we investigated the factors that could most influence the respondents' willingness to answer the open-ended questions related to their opinions on the survey.

The analysis of the results, carried out via a binary regression model, focuses on the hypotheses mentioned in the first paragraph, considering the response variable of the logistic model as 'responds or does not respond' to the question "*Do you have any suggestions for us for improving the questionnaire?*"[6]. The odds ratios do not show values that are extremely distant from one. But all variables show a good level of significance, with p-values never above 0.01. The structural variables on individual characteristics included in the model show a higher propensity to answer among the younger age group: the older the respondents, the lower the propensity to answer the open-ended question. Interviewers with Italian citizenship are slightly more likely to share suggestions compared to foreign students. Gender and territorial breakdown do not have significant effects.

Among the socio-economic characteristics "sharing a bedroom" seems to positively affect the attitude to answer the question. Those who share a room with a relative are 17% more likely to answer than those who have a room for themselves. In the case of children with no brothers or sisters, it is more difficult for them to share a room with other people, but the variable is meant to measure the habit of relatedness and living conditions. As to the mother's educational qualification, we note that the higher the mother's level of education, the lower the propensity to answer the open-ended question.

To explore relational networks, we first examined the issue of bullying. The bullying indicator variable was constructed from seven questions in which the respondent was asked whether, in the past twelve months, he or she had been offended, threatened, ostracized, or defamed, either in person or online. If the respondent had experienced at least one of these aspects monthly, the "bullying

[6] Binary logistic model with logit distribution function. The independent variables are: citizenship, territorial breakdown, age group, gender, shared bedroom, mother's educational qualification, bullying indicator, school performance, frequency of seeing friends and average hours a day on social media.

indicator" is one. This aspect is the most important factor in the decision to answer the open question; youths who have experienced these forms of discrimination were 25% more likely to answer the open question than those who have not. The respondent's level of interaction on social networks and in person were also included in the study; these two aspects seem to go in opposite directions. The less often respondents meet their friends, the less likely they are to respond to the open-ended question. In contrast, those who spend only one or two hours on social media are more likely to respond to open-ended questions compared to those who spend more than three hours. Although the intensity of the estimates is very low, high use of social networks seems to be synonymous with slight isolation, rather than with strong friendships.

Table 1 – *Odds ratios, confidence intervals and p-values of logistic regression on the propensity to answer the open-ended question.*

Variable (ref.)	Odds ratio	Confidence interval	Pr(>|z|)
Mother university (middle school or lower)	0.91	0.86-0.96	***
Mother high school (middle school or lower)	0.95	0.91-0.99	*
Shared bedroom (not shared)	1.17	1.12-1.22	***
High bullying indicator (low)	1.25	1.19-1-32	***
Low school performance (high)	1.09	1.02-1.16	**
Sees friends less than once a week (everyday)	0.88	0.83-0.94	***
Sees friends several times a week (everyday)	0.89	0.84-0.94	***
4+ hours on social (< 1 hour)	0.85	0.78-0.92	***
3-4 hours on social (< 1 hour)	0.85	0.79-0.92	***
1-2 hours on social (< 1 hour)	0.91	0.84-0.98	*

*Note: levels of significance * p<0.05, ** p<0.01, *** p<0.001*

4. Textual analysis of open-ended question

The "Do *you have any suggestions for us for improving the questionnaire?*" open ended question included in the survey "Children and Young People" was analysed through a lexicometric approach via textual analysis.

The answers to the open-ended question were collected in a corpus that underwent a double stage analysis. In the first stage, the corpus was pre-processed and imported into the software TaLTaC2 to carry out a lexical analysis aimed at identifying positive specificities (Bolasco, De Mauro, 2013; Lafon, 1980; Lebart, Salem, Berry, 1997) and sentiment analysis. Following this first stage, the analysis of positive specificities allows for the detection of the main themes related to the different ages of the respondents, while the sentiment analysis revealed the presence of both positive and negative opinions expressed through the use of adjectives. In

the second stage, after lemmatization of the corpus, a cluster analysis was carried out.

According to this analytical strategy, in the first stage of the analysis, the corpus was prepared by removing special characters and replacing uppercase with lowercase letters. Hence, given that the software used for the analysis (TaLTaC2) allows the text to be processed in Italian, all the responses written in other languages were removed. Following this approach, the corpus analysed consists of 15,096 fragments (38.5% of respondents). This corpus was imported into TaLTaC2 and pre-processed by applying a procedure of parsing, normalization and lexicalization, in order to reduce the redundancy and to provide homogeneity among forms. After this preliminary procedure, the corpus consists of 86,980 word tokens and 6,568 word types (see table 2). In order to verify whether the textual data could be statistically processed, the type-token ratio lexical indicator was calculated (see table 2).

Table 2 – *Main characteristics of the textual corpus composed of the answers to the open-ended question.*

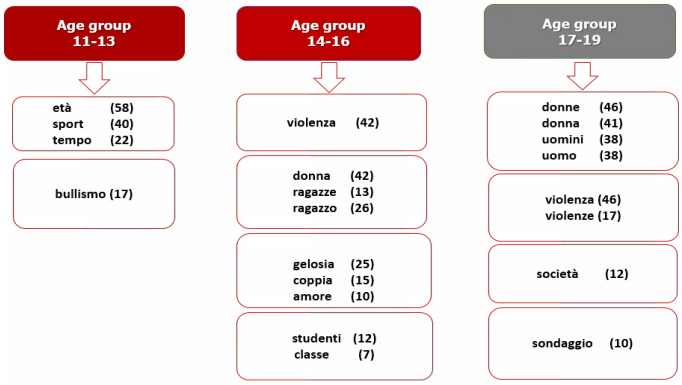
Lexicometric measurements	
Word types (V)	6,568
Word token (N)	86,980
Hapax (V1)	3,512
Type-token ratio index (V/N*100)	7.55%

According to the size of the corpus, the type-token ratio index, since it is 7.55% and < 20% (Bolasco, 1999, p. 203) shows that the corpus can be subjected to quantitative analysis.

5. The positive specificities

In order to investigate lexical differences related to the different ages of the respondents, we analysed the positive specificities for the age variable. In particular, specificity can be used to spot items that are both over- and under-represented in a corpus by applying a statistical test (Lafon, 1984, pp. 65-66). In our study, we are solely interested in positive specificities which highlight items that are over-represented in the corpus. Moreover, given that respondents were classified into three different age groups, we investigated positive specificities for the three modalities of the age variable. This analytical approach allows us to identify three different groups of words (positive specificities), as shown in figure 2.

Figure 2 – *Specificities for each age group from the textual analysis of the open-ended question. Year 2023.*



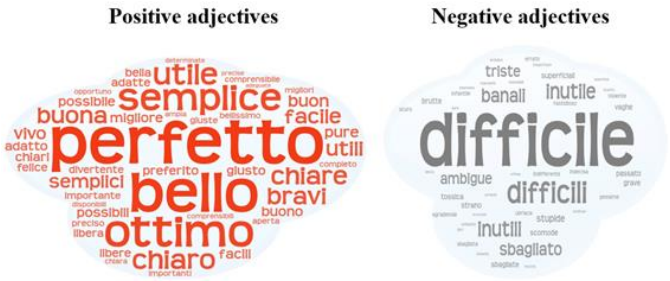
The first group characterizes the 11-13 age group of respondents that suggested improving the survey by introducing new themes such as bullying, cyberbullying, sports, free time, and road safety. Moreover, these respondents suggested using less direct questions applying age-appropriate language. The second group relates the 14-16 age group. The introduction of new themes to improve the survey is also suggested by these respondents including: video games, school life and geopolitical context, eating habits and relationship dynamics. They also expressed the need for a formal revision of the language in order to make it clearer. Finally, the third group characterizes respondents aged 17-19. In their opinions, the survey could have been improved by introducing new themes such as women's rights, men's mental health and themes related to current affairs. These respondents offered other suggestions for improving the survey, mainly related to the survey website and the survey technique. Regarding the website, they suggest modernizing it, in order to make it more suitable to the young age of the respondents. Moreover, they suggested completing the survey entirely online, without paper.

6. Sentiment analysis

Sentiment analysis was conducted using the software TaLTaC2 that automatically recognizes adjectives and classifies them as positive or negative (Bolasco, Della Ratta, 2004). Following this approach, the corpus underwent a grammatical tagging procedure in TaLTaC2 in order to identify the adjectives. This procedure allowed us to tag 429 adjectives, for a total of 2,479 word tokens. Regarding the adjectives, the sentiment analysis identified 61% positive adjectives and 39% negative. Thus, sentiment analysis highlighted a prevalence of positive

opinions of respondents. Figure 3 shows words clouds of both positive and negative adjectives.

Figure 3 – *Word cloud of positive and negative adjectives from the textual analysis of the open-ended question. Year 2023.*



Considering the positive adjectives, respondents defined the questionnaire as "perfetto" ("perfect", 437), "bello" ("beautiful", 109), "chiaro" ("clear", 106), "semplice" ("simple", 94) and "utile" ("useful", 51). On the other hand, for some other respondents the questionnaire was "difficile" ("difficult", 51), "inutile" ("useless", 22), "banale" ("banal", 10) and "sbagliato" ("wrong", 9). These adjectives provide a snapshot of the potential weaknesses of the survey, which will be explored with the following cluster analysis.

7. Cluster analysis

A cluster analysis was conducted (DHC- Descending Hierarchical Classification, Reinert method (1990), IRaMuTeQ software[7]) that allowed for the identification of eight groups of respondents, where the first three groups represent 98.4% of the sample. We will focus our analysis on three big groups, characterized by different socio-demographic aspects and the topics reported in their responses, identified through the lexical content of each cluster. Each cluster is characterized by the use of specific words. The first group (31.4%) has been named "the questionnaire methodologists" for their suggestions aimed at setting up specific actions on the questionnaire, such as making it shorter or improving the graphics. This group includes 33% of the respondents. They are mostly aged between eleven and thirteen, they tend to be male and foreign. Moreover, they use social media one or two hours a day. Economically, they are students from wealthier backgrounds. The second group (34%), named "the qualitative methodologists" asked for more opportunities

[7] We express our gratitude to Francesca Della Ratta for her invaluable assistance in conducting the IRaMuTeQ data analysis.

to express themselves freely: they would like open-ended questions, suggest replacing true/false responses with agree/disagree to better express their opinions, and they wanted more ways to respond to the closed-ended questions. In this analysis, we'll focus our attention on the third group (33.03%), named "the thematic" for their suggestions for introducing new topics. They are mostly female, aged between 14-19 years, and use social media less than an hour a day. They generally reported a low economic status. The words that characterize this group are primarily related to the everyday life of the young, such as family and school.

Table 3 – *Words that mostly characterize the "thematic" cluster in the textual analysis of the open-ended question. Absolute values, percentage values and Chi-square. Year 2023.*

Word	Eff. Tot.[8]	Perc.[9]	X^2[10]	Word	Eff. Tot.[3]	Perc.[4]	X^2[5]
Violenza	104	95.2	185.1	Sicuro	15	73.3	11.0
Bullismo	22	95.4	38.9	Lingua	31	93.5	51.6
Abuso	10	100.0	20.3	Straniero	33	90.9	50.2
Subire	7	100.0	14.2	Cittadinanza	28	92.9	5.5
Vittima	6	100.0	12.2	Razzismo	17	94.1	8.7
Discriminazione	6	100.0	12.2	Sessuale	28	92.9	5.5
Cyberbullismo	5	100.0	10.1	Lgbt	11	100.0	2.3
Mentale	37	100.0	75.5	Sessualità	8	100.0	6.2
Salute	30	96.7	55.2	Sport	52	71.1	4.4
Psicologico	20	100.0	40.7	Ambiente	27	96.3	9.0
Paura	13	92.3	20.7	Droga	14	100.0	8.4
Disagio	9	88.9	12.7	Guerra	12	100.0	4.3
Felice	13	76.9	11.3	Inquinamento	5	100.0	10.1

The analysis shows that the word "futuro" (future) stands out (Perc=95.6; X^2=244.4). Young people are asking more questions about their future projects, but they are also expressing their fears and insecurities.

"Ask questions that are also inherent to fear of the future, understood as climate fear, fear of an uncertain future because of the climate and the collapse of the planet. But also fear of the instability of Italian labor, fear of not being able to have a future."

They are worried about: violence, abuse, bullying, and cyberbullying. Additionally, they also express concerns about themes that usually are not asked to

[8] Number of text segments containing the cited word in the corpus at least once.
[9] Percentage of word occurrence on the text segments of this cluster for its occurrence in the corpus.
[10] Association between word and cluster.

this age group, such as mental health, fear, discomfort, racism, and sexuality. Taking a broader view, there is also a focus on collective issues, with words connected to the environment and war.

8. Conclusions

The study of the propensity to answer the open-ended question on suggestions for the survey shows that, among individual characteristics, age has the greatest effect: the younger the respondent, the greater the propensity to answer. Regarding socioeconomic characteristics, we note that a less comfortable situation, such as sharing a room, has a positive and significant effect on the propensity to respond. Finally, regarding relational networks, it is important to emphasize that adolescents experiencing bullying are more inclined to offer suggestions. The analysis indicated that the survey "Children and Young People 2023" conducted by Istat was largely appreciated by respondents, as highlighted by the sentiment analysis, which shows a higher frequency of positive adjectives in the responses of the youngest, expressing their positive attitude toward the survey. The use of adjectives, in the open-ended question, such as "perfect", "beautiful", "clear" and "useful" confirms respondents' positive opinions. This positive attitude was also confirmed by the low frequency of negative adjectives in the analyzed texts. These findings suggest that the survey should be repeated in the future, maintaining the innovations proposed in the 2023 edition, like the introduction of QR code. However, the analysis of the words included in the responses to the open-ended question identified some improvements, such as the inclusion of new topics, closer to the everyday life of a younger person and to geopolitical and social context. The introduction of these new topics could be added by taking into account the age of the respondents, in order to better represent their experiences and their sensibility.

References

BOLASCO S. 1999. *L'analisi multidimensionale dei dati*. Roma: Carocci Editore.

BOLASCO S., DELLA RATTA-RINALDI F. 2004. Experiments on semantic categorisation of texts: analysis of positive and negative dimension. In *Le poids des mots, Actes des 7es journées Internationales d'Analyse Statistique des Données Textuelles*. Louvain-la-Neuve: UCL, Presses Universitaires de Louvain, pp. 202-210.

BOLASCO S., DE MAURO T. 2013. *L'analisi automatica dei testi: fare ricerca con il text mining*. Roma: Carocci Editore.

BORGERS N., DE LEEUW E., HOX J. J. 2000. Children as Respondents in Survey Research: Cognitive Development and Response Quality. *Bulletin of Sociological Methodology/Bulletin de Mathodologie Sociologique*, Vol. 66, No. 1, pp. 60–75.

CONTI C., FANFONI L., FAZZI G. 2024. Giving children and young people a voice: the experience of the sample survey on children and young people in Italy. *Working Paper n.1*, Expert meeting on statistics on children Geneva, Switzerland, 4-6 March 2024.

DE LEEUW E. D. 2011. Improving data quality when surveying children and adolescents: Cognitive and social development and its role in questionnaire construction and pretesting. In Report prepared for the Annual Meeting of the Academy of Finland: Research Programs Public Health, Finland.

ISTAT 2024. *Nuove generazioni sempre più digitali e multiculturali*, Statistica Report, 20 maggio 2024.

ISTAT 2017. *L'utilizzo della tecnica cawi nelle indagini su individui e famiglie,* eBook letture statistiche.

LAFON P. 1980. Sur la variabilité de la fréquence des formes dans un corpus, *Mots*, Vol. 1, No 1, pp. 127-165.

LAFON P. 1984. *Dépouillement et statistiques en lexicométrie*. Paris: Slatkine Champion.

LEBART L., SALEM A., BERRY L. 1997. *Exploring textual data (Vol. 4)*. Dordrecht: Springer Science & Business Media.

LIU B. 2012. *Sentiment analysis and opinion mining (synthesis lectures on human language technologies)*. San Rafael: Morgan & Claypool Publishers.

OMRANI A., WAKEFIELD-SCURR J., SMITH J., BROWN N.. 2019, Survey Development for Adolescents Aged 11–16 Years: A Developmental Science Based Guide, *Adolescent Res Rev* 4, pp. 329–340

REINERT M. 1990. Alceste, une méthodologie d'analyse des données textuelles et une application: Aurélia de G. de Nerval, *Bulletin de méthodologie sociologique*, Vol. 28, pp. 24- 54.

SADIKU M., SHADARE A., SARHAN M., 2017. Digital Natives. International Journal of Advanced Research in Computer Science and Software Engineering.

UNECE. 2022. Guidance on statistics on children: spotlight on children exposed to violence, in alternative care, and with disabilities.

Cinzia CONTI, Istat, ciconti@istat.it
Gabriella FAZZI, Istat, fazzi@istat.it
Barbara D'AMEN, Istat, barbara.damen@istat.it
Marco RIZZO, Istat, marco.rizzo@istat.it

DISENTANGLING RESIDENTIAL GEOGRAPHIES OF SELECTED FOREIGN POPULATION GROUPS IN A CONTEXT OF LOW IMMIGRATION[1]

Federico Benassi, Alessio Buonomo, Rosaria Simone, Salvatore Strozza

Abstract. Foreign presence is an intrinsically spatial phenomenon, characterized by strong geographical variability across different territorial contexts. This variability becomes progressively more intense as we move down the territorial scale of analysis. This contribution proposes an empirical evaluation of the residential geographies of EU foreign citizens and non-EU foreign citizens residing in the Metropolitan City of Naples (Campania, Southern Italy), a context characterized by a low level of immigration. The empirical analysis uses data from the 2021 permanent demographic census and is carried out using census tracts as the primary statistical units and an *ad hoc* geographical framework ('quartieri' for the municipality of Naples and municipalities for the rest of the Metropolitan City of Naples). The residential geographies of foreign groups are summarized using the dissimilarity index of Duncan and Duncan in the adjusted version proposed by Mazza and Punzo. This index informs us about the level of evenness between the spatial distribution of foreigners compared to that of Italians (who act as the reference group). The adjusted version of the index allows us to control for some biases that normally arise in cases with a low number of foreigners. The geographical distribution of the bias-corrected index for the two foreign population groups is represented by thematic maps. Then, its level of spatial autocorrelation (global and local) is analyzed. The results allow us to appreciate the spatial variability of the foreign population residing in the Metropolitan City of Naples and to provide initial insights into their spatial patterns using a multiscale geographical approach, which serves as a first step toward a better understanding of their residential behaviours.

1. Introduction

The foreign population in Italy has surged in recent decades, jumping from just over a million residents in 2001 to now, 2024 January the 1st, over 5.3 million, comprising 8.8% of total population. Their settlement patterns vary widely, reflecting different adaptation strategies ranging from concentrated to dispersed models (Strozza, 2006). These settlement choices are influenced by various factors, including community ties and labor specialization (Ferrara *et al.*, 2010; Conti *et al.* 2023). The study of settlement patterns of different foreign communities residing in Italy is particularly relevant as it

[1] The authors contributed equally to the conception and realization of the contribution.

provides essential elements for understanding the characteristics of foreign presence in different local and regional contexts, especially in metropolitan areas where spatial inequalities, including residential segregation, are more frequent (Benassi *et al.*, 2023a). The spatial polarization of the foreign population is an issue that is still particularly evident in Italy. The last data provided by Istat certify that in the North reside slightly less than 59% (3,1 million) of the total foreign population while the South count only the 16.9% (897 thousand) of them. The spatial distribution is even more unequal if bear in mind that the 14 Italian Metropolitan Cities[2] host the 37.4% of the total foreign population resident in Italy with the seven Metropolitan Cities of the Centre-North that, alone, host slightly less than 1/3 of the total foreign population resident in Italy. This framework is not new, as shown in previous studies on the subject (Strozza *et al*., 2016).

The paper represents a first attempt to disentangle the residential geographies of selected foreign groups (EU and non-EU foreign citizens) resident in the Metropolitan Cities of Naples (MCN hereafter). Studies on the geographical distribution of foreign population in Italy and its level of residential segregation are numerous (Benassi *et al.*, 2022; Bitonti *et al.* 2023a, 2023b; Conti *et al.*, 2023; Pratschke and Benassi, 2024; Rimoldi *et al.*, 2024) including contributions specifically referred to Southern urban contexts (Benassi *et al.*, 2023b; Busetta *et al.*, 2015; Mazza *et al.*, 2018; Mazza and Punzo, 2016). Nevertheless, to the best of our knowledge, this is the first contribution that use data coming from the permanent Census (2021) and that focussing on a single Metropolitan City of South Italy combine a multiscale approach of analysis using an *ad hoc* geographical partition that include both sub municipal (s.c. 'quartieri') and municipal territorial units and implementing both global and local indexes (Brown and Chung, 2006).

The paper is structured as follows: the next section presents the materials and methods, followed by the results in section 3, and finally, discussions and conclusions are provided in section 4.

2. Geographical contexts of analysis, data and methods

2.1. The geographical context of analysis

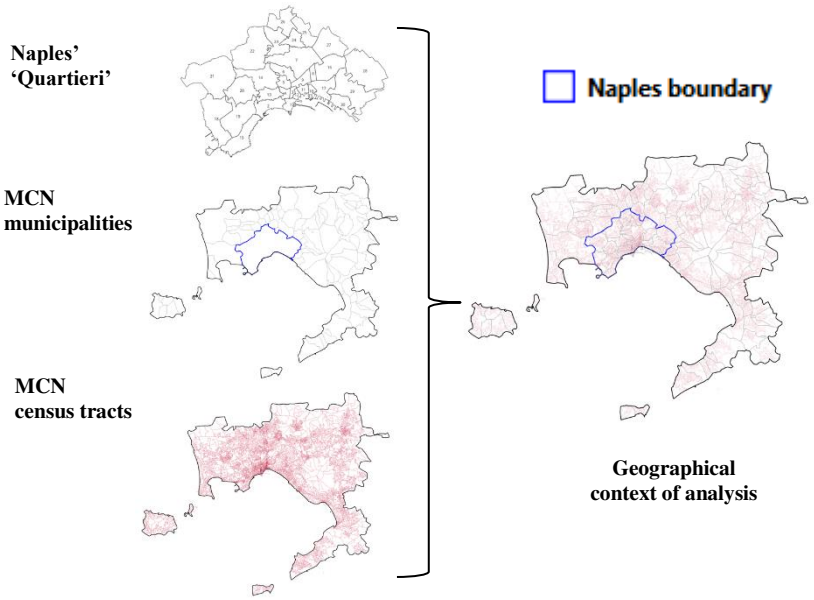
The idea of focusing on the MCN arises from the necessity of limiting the empirical analysis to the urban contexts of the Campania region. Moreover, the relevance of Italian Metropolitan Cities as pole of attraction for international migrations flows is known and underlined in a recent contribution of Buonomo and

[2] These are represented by the province of Milan, Turin, Venice, Genoa, Bologna for the North; Florence and Rome for the Centre; Naples, Bari, Reggio di Calabria, Palermo, Catania, Messina and Cagliari for the South. Please note that Cagliari is the only case in which the Metropolitan City is different from the province.

colleagues (Buonomo *et al.,* 2024). Using MCN assures an adequate minimum number of foreign resident population, and it allows us to better analyse processes (like residential segregation) that should be considered as purely 'urban' (Feitosa *et al.*, 2007).

Basically (Figure 1) we built an *ad hoc* geographical context: the 30 districts ('quartieri') of Naples plus the rest of the municipalities that for the MCN (91) for a total of 121 territorial units (e.g. statistical cases). For each of the 121 territorial units the elementary territorial units are represented by the census tracts (as at 2011).

Figure 1 – *The geographical context of analysis (an* ad hoc *construction).*



Considering the municipality of Naples according to its division into districts (i.e. 'quartieri') allows us to appreciate the internal heterogeneities within the capital city of the MCN, which, as evident, are not few and can significantly contribute to a better interpretation, also in this explorative phase, of the residential geographies of foreign population groups and their possible drivers. About the choices of 'quartieri' as sub municipality units of analysis it is important to provide some other details.
For Naples, 'quartieri' are not proper administrative units, yet these sub-municipalities represent, to a certain extent, the social history of the municipality and they contribute differently to its economic and social dynamics. Moreover, they guarantee a balanced distribution in terms of resident population. In 2021 the resident population in the municipality of Naples was 921,142 with a minimum population resident in the *Porto* district (4,469 residents) and a maximum population resident in

Fuorigrotta district (63,049 residents). For an ovierview on the geographies of the 'quartieri' see the Appendix.

2.2. Data and methods

Data used are from the permanent census (2021) and they refer to resident population (stock) by country of citizenship. In order to have a sufficient number of individuals in each territorial unit (30 districts and 91 municipalities) but also considering the inner heterogeneity inside the foreign population we used two sub population groups: EU foreign citizens and non-EU foreign citizens. Of course, a certain degree of inner heterogeneity remains, but this way to classify the foreign population is quite common in studies on settlement models and residential segregation (Benassi *et al.*, 2020a, Malmberg *et al.*, 2018). In Table 1 are shown the population groups used in the analysis as of 2021.

Table 1 – *Resident population by groups of citizenship and territorial units. MCN, end of 2021 (permanent census).*

Territorial units	EU-foreign	Non-EU foreign	Italians	Total
Naples municipality	4,753	48,687	867,702	921,142
Rest of the MC	11,980	55,887	1,999,367	2,067,234
Total	**16,733**	**104,574**	**2,867,069**	**2,988,376**

The residential geographies of foreign populations are here analysed in terms of evenness (using the Italian population as reference group). Evenness is the first dimension of the Massey and Denton's (1988) conceptual model of residential segregation. Evenness concerns with the differential distribution of foreigners (or, more generally, social groups) in the area units. More specifically, evenness measures of segregation compare the spatial distributions of different groups among units. When majority and minority populations are evenly distributed, segregation is smallest. There are many different evenness measures of segregation, for an overview we remind to the existing literature (Fossett, 2017; Reardon and O'Sullivan, 2004; Tivadar, 2019). The most widely used is the dissimilarity index (*ID*), firstly proposed by Ducan and Duncan (1955a, 1955b) which is calculated according to the formula:

$$ID = \frac{1}{2}\sum_{i=1}^{n}\left|\frac{z_i}{Z} - \frac{y_i}{Y}\right| \quad (1)$$

where *i* is used to identify each of the *n* territorial units (in our case enumeration areas of each single 'quartieri' for the municipality of Naples and of each single

municipality for the rest of the MCN), while z_i and y_i are the total of the foreign group (EU or non-EU) and Italians, respectively, in the *i-th* territorial unit, and *Z* and *Y* are the overall total number of residents of the two groups, respectively.

ID measures the percentage of a group's population that would have to change residence to have the same percentage of that group as the city overall. The index ranges from 0 (absence of segregation) to 1 (complete segregation).

ID is widely known and widespread measures in study on the residential geographies and residential segregation of migrants (Friedman, 2008; Iceland *et al.*, 2013, 2014; Logan and Parman 2017; Malmberg *et al.*, 2018). Despite its widespread adoption, the index also presents some limitations and weakness that have been addressed by many scholars (Morrill, 1991; White, 1983, 1986; Wong, 1993; Reardon and O'Sullivan, 2004; Yao *et al*., 2019).

As clearly explained in Busetta *et al.* (2015), a problem with the Duncan and Duncan's dissimilarity index is that it appears to be an upward biased estimator of systematic dissimilarity. Within plausible assumptions, Allen *et al.* (2009) demonstrate, using Monte Carlo simulations, that random allocation generates substantial unevenness, and hence an upward bias, especially when dealing with: small units sizes, a small minority proportion and a low level of segregation. These are three aspects that potentially characterized the MCN and that cannot be ignored in measuring the residential geographies of the selected foreign groups.

To partially mitigate these problems proper to the *ID*, we resort to the *ID* bias-corrected estimator introduced by Mazza and Punzo (2015), which outperforms many other resampling based bias corrections in terms of both bias and mean square errors[3].

A second aim of the study was to measure the level of global and local spatial autocorrelation of the bias-corrected *ID* for the two foreign populations here observed. Spatial autocorrelation can be detected using the global Moran's *I* (Moran, 1948) to determine whether the bias-corrected ID is similar in spatially adjacent territorial units. In its global form, Moran's *I* computes a unique value for the MCN and is calculated as follows:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \sum_{i=1}^{n} (x_i - \bar{x})^2} \quad (2)$$

where x_i is bias-corrected *ID* in the territorial units *i*, *n* is the number of territorial units (121 in our case), w_{ij} denotes the elements in the spatial weight matrix, and $\bar{x}$ denotes the average bias-corrected *ID*. Global Moran's *I* ranges within the interval [-1,1]. Values greater than 0 correspond to positive spatial autocorrelation, and

[3] The analysis has been carried out also using the classic version of the ID. Results are available from the Authors upon reasonable request.

values smaller than 0 indicate negative spatial autocorrelation. We tested different spatial weight matrices and chose first-order queen (i.e. two territorial units are neighbors if they share a boundary and/or a geographical vertex). The clusters of bias-corrected *ID* were identified using the local version of the univariate Moran's *I* as proposed in Anselin (1995). The univariate local Moran's *I* (i.e., I_i) for the *i-th* territorial unit is calculated as follows:

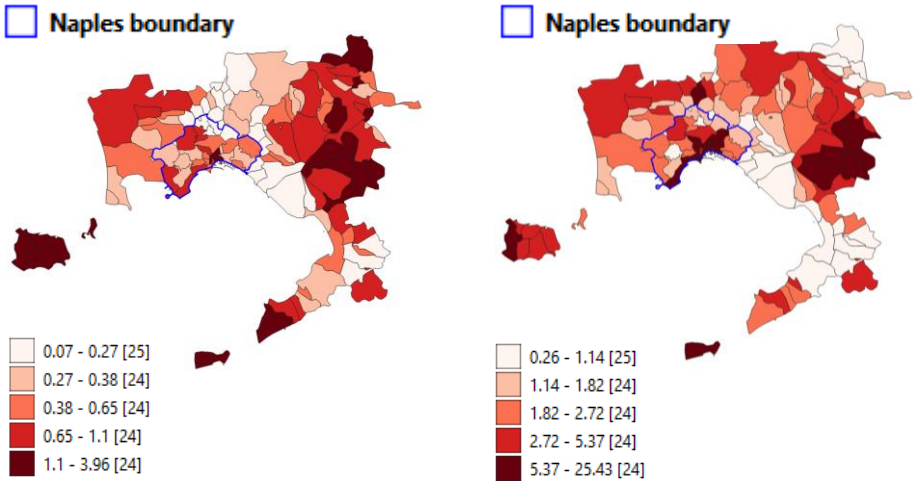
$$I_i = \frac{(x_i - \bar{x})}{\sum_{h=1}^{n}(x_h - \bar{x})^2/(n-1)} \sum_{\substack{j=1 \\ j \neq i}}^{n} w_{ij}(x_j - \bar{x}) \quad (3)$$

where x_i is the bias-corrected ID for the *i-th* territorial unit and x_j's are the bias-corrected ID values in adjacent areas.

3. Results

The geographical distribution of the two populations groups is quite variable underlying different spatial patterns. For a sake of brevity, we report here only the distribution of the ratio between foreign population (EU and Non-EU foreign citizens) and Italians (Figure 1).

Figure 1 – *Ratio between foreign population and Italians. EU foreigners (left panel), Non-EU foreigners (right panel). Percentage values.*



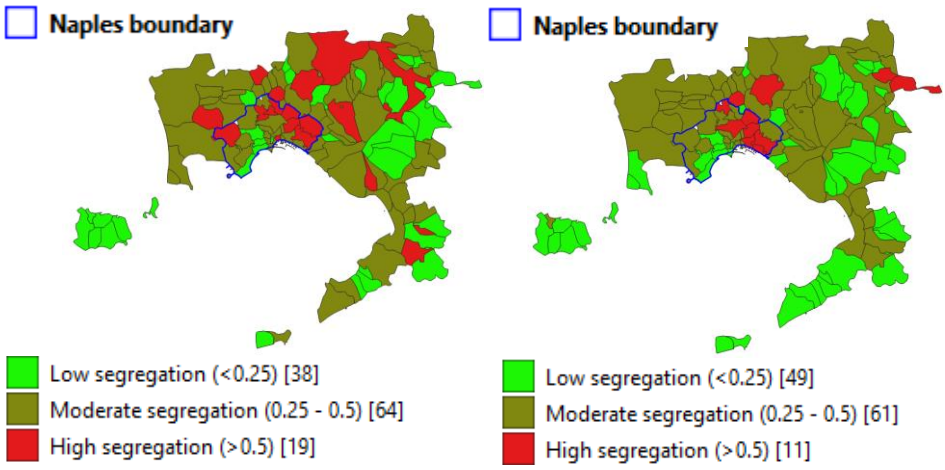
For the non-EU foreign population, the highest values of the indicator are recorded in some districts of the municipality of Naples: *Zona Industriale* (25.4%),

Mercato (24.2%), *San Lorenzo* (22.6%) and *Stella* (21.8%), but also in two neighbour municipalities located across the eastern quadrant of the MCN area, Palma Campania (19.6%) and San Giuseppe Vesuviano (19.4%). The EU-foreign population records the higher values of the indicator outside the Municipality of Naples in some specific municipalities that in some cases are very famous touristic destinations and place with a high level of amenity - Procida (3.9%) and Serrata Fontana (2.3%) – or in more marginal area, located on across the eastern quadrant of the MCN – Poggiomarino (3.3%) and Liveri (2.9%). The district of Naples with the highest level of the indicator is, for this population, *Pendino* that is actually quite far compared to the other contexts (1.6%).

The level of dissimilarity is not particularly high for both populations (0.34 and 0.32 are the average values of bias-corrected ID for EU foreign and non-EU foreign citizens respectively) with a certain degree of spatial variability. The territorial units with high level of dissimilarity (>0.5) are more numerous in the case of EU foreign population (19) compared to non-EU Foreign population (11). Another element of distinction is that in the first case the 47.3% of the territorial units with an ID >0.5 is included in the municipality of Naples. A percentage that in the case of non-EU foreign population rise to 63.6% (Figure 2).

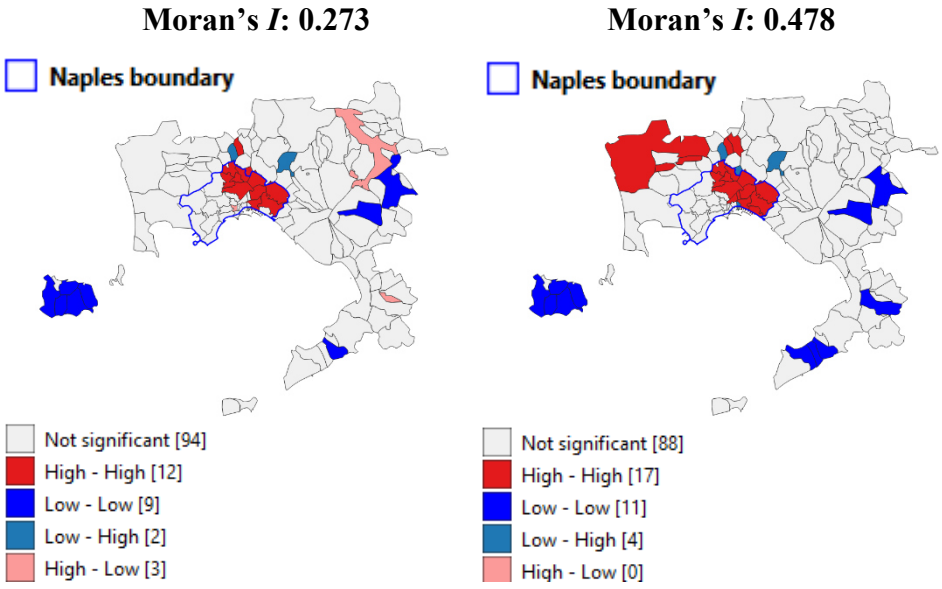
The highest level of residential segregation is recorded, in the case of EU foreign population, in Pimonte (0.82) a quite small municipality of about 6 thousand residents while in the case of non-EU foreigners the highest level of bias-corrected *ID* is the one recorded in Scampia (0.70) in the norther part of the municipality of Naples.

Figure 2 – *Bias-corrected ID, thematic maps. EU foreign population (left panel), non-EU foreign population (right panel).*



The level of global clustering is higher for the bias-corrected *ID* of non-EU foreign population (0.478 versus 0.273) (Figure 3). At the local level it is quite interesting to note that the HH (high-high) clusters inside the municipality of Naples are the same for both foreign populations but in the case of non-EU foreign population there is a further area located on the west upper quadrant of the MCN formed by the municipalities of Giuliano in Campania and Villaricca. What emerges therefore seems to indicate rather contained levels of dissimilarity but with areas where the levels are quite high, tending to cluster particularly within the municipality of Naples where, moreover, the presence of non-EU foreigners is larger than in the rest of the MCN compared to Eu foreigners (46.5% versus 28.4%).

Figure 3 – *Univariate global and local Moran's I. EU foreign population (left panel), non-EU foreign population (right panel).*



p-value ≤ 0.05, Queen contiguity matrix of 1st order

4. Discussion and conclusions

The topic of settlement geographies of the immigrant foreign population and its level of segregation is of primary interest in order to assess the level of integration and social cohesion of the host contexts (Bolt et al., 2010).

In the past, the levels of residential segregation of immigrant foreigners were quite low in Southern Europe compared to those in Central and Northern Europe (Malheiros, 2002). However, some more recent studies, using 2011 census data, have shown how things have changed, both as a result of the 2008 economic crisis and the increasing weight of the foreign population residing in contexts traditionally areas of emigration (Benassi et al., 2020a, 2020b).

This contribution, exploratory and descriptive in nature, fits into this line of study by proposing an analysis of the level of dissimilarity in the distribution of foreigners (both EU and non-EU foreign citizens) residing in the MCN compared to Italians.

The results obtained using a custom-built geography, a robust index, and based on the 2021 permanent census data, have shown that the average level of dissimilarity is quite low (<0.5) for both foreign populations, with a slight advantage for the non-EU population which, however, shows higher levels of clustering of the indicator both globally and locally. Another interesting difference between the settlement geographies of the two populations, compared to those of Italians, is that in the case of the non-EU foreign population, the areas with higher levels of segregation are mostly located within the municipal perimeter, where their concentration is also higher.

The results represent a first step towards further investigations that should concern the estimation of the determinants of dissimilarity levels through appropriate regression models capable of controlling the level of ID autocorrelation and possibly providing local estimates.

Appendix

The "Quartieri" of Naples municipality

Acknowledgements

Contribution conceived and realized as part of the PRIN 2022-PNRR research project "Foreign population and territory: integration processes, demographic imbalances, challenges and opportunities for the social and economic sustainability of the different local contexts (For.Pop.Ter)" [P2022WNLM7]. Funded by European Union – Next Generation EU, component M4C2, Investment 1.1.

References

ALLEN R., BURGESS S., WINDMEIJER F.A. 2009. More reliable inference for segregation indices. *Working Papers 09/2016*, Centre for Market and Public Organisation, Bristol Institute of Publica Affairs, University of Bristol.

ANSELIN L. 1995. Local indicators of spatial association, *Geographical Analysis*, Vol. 27, No. 2, pp. 93-115.

BENASSI F., BITONTI F, MAZZA A., STROZZA S. 2023b. Sri Lankans' residential segregation and spatial inequalities in Southern Italy: an empirical analysis using fine-scale data on regular lattice geographies, *Quality & Quantity*, Vol. 57, No. 2, pp. 1629-1648.

BENASSI F., BONIFAZI C., HEINS F., LIPIZZI F., STROZZA S. 2020a. Comparing residential segregation of migrant populations in selected European urban and metropolitan areas, *Spatial Demography*, Vol. 8, pp. 269-290.

BENASSI F., CRISCI M., MATTHEWS S.A., RIMOLDI S.M.L. 2022. Migrants' population, residential segregation, and metropolitan spaces - insights from the Italian experience over the last 20 years, *Migration Letters*, Vol. 19, No. 3, pp. 287-301.

BENASSI F., IGLESIAS-PASCUAL R., SALVATI L. 2020b. Residential segregation and social diversification: exploring spatial settlement patterns of foreign population in Southern European cities, *Habitat International*, 1010, 102200.

BENASSI F., NACCARATO A., IGLESIAS-PASCUAL R., SALVATI L., STROZZA S. 2023a. Measuring residential segregation in multi-ethnic and unequal European cities. *International Migration*, Vol. 61, No. 2, pp. 341-361.

BITONTI F., BENASSI F., MAZZA A., STROZZA S. 2023a. From South Asia to Southern Europe: A comparative analysis of Sri Lankans' residential segregation in the main Italian cities using high-resolution data on regular lattice geographies, *Genus*, Vol. 79, No. 3, pp.1-27.

BITONTI F., BENASSI F., MAZZA A., STROZZA S. 2023b. Framing the residential patterns of Asian communities in three Italian cities: evidence from Milan, Rome and Naples, *Social Sciences*, Vol. 12, No. 9, pp.1-29.

BOLT G., ÖZUEKREN A. S., PHILLIPS D. 2010. Linking integration and residential segregation, *Journal of Ethnic and Migration studies*, Vol. 36, No. 2, pp. 169-186.

BROWN L.A., CHUNG S.Y. 2006. Spatial segregation, segregation indices and the geographical perspective, *Population Space and Place,* Vol. 12, No. 2, pp. 125-143.

BUONOMO A., BENASSI F., GALLO G., SALVATI L., STROZZA S. 2024. In-between centers and suburbs? Increasing differentials in recent demographic dynamics of Italian metropolitan cities, *Genus*, Vol. 80, No.1, pp. 1-19.

BUSETTA A., MAZZA A., STRANGES M. 2015. Residential segregation of foreigners: an analysis of the Italian city of Palermo, *Genus*, Vol. 71, No. 2-3, pp. 177-198.

CONTI C., MUCCIARDI M., SIMONE M. 2023. Exploring the settlement models of the main foreign communities residing in Italy (2003-2021), *Social Sciences*, Vol. 12, No. 9, 524.

DUNCAN O.D., DUNCAN B. 1955a. A methodological analysis of segregation indexes, *American Sociological Review*, Vol. 20, No. 2, pp. 210-217.

DUNCAN O.D., DUNCAN B. 1955b. Residential distribution and occupational stratification, *American Journal of Sociology*, Vol. 60, pp. 493-503.

FEITOSA F.F., CAMARA G., MONTEIRO A.M.V., KOSCHITZKI T., SILVA M.P. 2007. Global and local spatial indices of urban segregation, *International Journal of Geographical Information Science*, Vol. 21, No. 3, pp. 299-323.

FERRARA R., FORCELLATI L., STROZZA S. 2010. Modelli insediativi delle comunità immigrate in Italia, *Bollettino della Società Geografica Italiana*, Vol. 3, N. 12, pp. 619-639.

FOSSETT M. 2017. *New methods for measuring and analysing segregation*. Cham: Springer Nature.

FRIEDMAN S. 2008. Do declines in residential segregation mean stable neighborhood racial integration in metropolitan America? A research note, *Social Indicator Research*, Vol. 37, No. 2, pp. 920-933.

ICELAND J., SHARP G., TIMBERLAKE J.M. 2013. Sun belt rising: regional population change and the decline in black residential segregation, 1970-2009, *Demography*, Vol. 50, No. 1, pp. 97-123.

ICELAND J., WEINBERG D., HUGHES L. 2014. The residential segregation of detailed Hispanic and Asian groups in the United States: 1980-2010, *Demographic Research*, Vol. 31, No.2, pp. 593-624.

LOGAN T.D., PARMAN J.M. 2017. The national rise in residential segregation, *The Journal of Economic History*, Vol. 77, No. 1, pp. 127-170.

MALHEIROS J.M. 2002. Ethni-cities: residential patterns in the Northern European and Mediterranean metropolises-implications for policy design, *International Journal of Population Geography*, Vol. 8, pp. 107-134.

MALMBERG B., ANDERSSON E.K., NIELSEN M.M., HAANDRIKMAN K. 2018. Residential segregation of European and non-European migrants in Sweden: 1990-2012, *European Journal of Population*, Vol. 42, No.2, pp. 169-193.

MASSEY D., DENTON N.A. 1988. The dimensions of residential segregation, *Social Forces*, Vol. 67, No. 2, pp. 281-315.

MAZZA A., GABRIELLI G., STROZZA S. 2018. Residential segregation of foreign immigrants in Naples, *Spatial Demography*, Vol. 6, pp. 71-87.

MAZZA A., PUNZO A. 2015. On the upward bias of the dissimilarity index and its corrections, *Sociological Methods & Research*, Vol. 44, No. 1, pp. 80-107.

MAZZA A., PUNZO A. 2016. Spatial attraction in migrants' settlement patterns in the city of Catania, *Demographic Research*, Vol. 35, pp. 117-138.
MORAN P.A.P. 1948. The interpretation of statistical maps, *Journal of the Royal Statistical Society*, Vol. 10, pp. 243-251.
MORRILL R.L. 1991. On the measure of geographical segregation, *Geography research forum*, Vol. 11, No. 1, pp. 25-36.
PRATSCHKE J., BENASSI F. 2024. Population change and residential segregation in Italian small areas, 2011-2021: An analysis with new spatial units, *Spatial Demography*, Vol. 12, No. 2, pp.1-30.
REARDON S.F., O'SULLIVAN D. 2004. Measures of spatial segregation, *Sociological Methodology*, Vol. 34, No.1, pp. 121-162.
RIMOLDI S.M., CRISCI M., BENASSI F., RAYMER J. 2024. Intra-urban residential mobility and segregation of foreigners in Rome, *Population Space and Place*, e277.
STROZZA S. 2006. Geographic distribution of foreign communities in Italy: preliminary analyses. In *Proceedings of the XLIII Scientific Conference of the Italian Society of Statistics SIS*, Padua: Cleup, pp. 389-392.
STROZZA S., BENASSI F., FERRARA R., GALLO G. 2016. Recent demographic trends in the major Italian urban agglomerations: the role of foreigners, *Spatial Demography*, Vol. 4, pp. 39-70.
TIVADAR M. 2019. OasisR: an R package to bring some order to the world of segregation measurement, *Journal of Statistical Software*, Vol. 89, pp. 1-39.
WHITE M.J. 1983. The measurement of spatial segregation, *American Journal of Sociology*, Vol. 8, No. 5, pp. 1008-1018
WHITE M.J. 1986. Segregation and diversity: measures in population distribution, *Population index*, Vol. 52, No. 2, pp. 1981-221.
WONG D.W.S. 1993. Spatial indices of segregation, *Urban Studies*, Vol. 30, No. 3, pp. 559-572.
YAO J., WONG D.W., BAILEY N., MINTON J. 2019. Spatial segregation measures: A methodological review, *Tijdschrift voor economische en sociale geografie*, Vol. 110, No. 3, pp. 235-250.

Federico BENASSI, Università di Napoli Federico II, federico.benassi@unina.it
Alessio BUONOMO, Università di Napoli Federico II, alessio.buonomo@unina.it
Rosaria SIMONE, Università di Napoli Federico II, rosaria.simone@unina.it
Salvatore STROZZA, Università di Napoli Federico II, strozza@unina.it

ORGANIZATIONAL CAPABILITIES IN HEALTH CARE INSTITUTIONS: A MULTISOURCE TERRITORIAL STUDY[1]

Francesca Abate, Anna Pia Maria Mirto, Francesco Gaudio, Adriano Pareto, Francesco Paolo Rizzo, Dario Maiolino

Abstract. The Italian debate on territorial healthcare is currently focused on the implementation of various measures under the National Recovery and Resilience Plan (PNRR), particularly Mission 6 on Health, aimed at enhancing human resources and services. This paper examines, from an ex-ante perspective, the potential capability of regional health institutions using an input-output framework applied to personnel classified by socio-demographic characteristics. The analysis draws on two key official data sources: the Annual Accounts from the Ministry of Economy and Finance, and ISTAT Multipurpose Household Survey. A comparison is made between the most recent available data (2022) and that from 2012. The authors introduce a novel composite index of organizational functionality (IFO-Health), which aggregates 10 distinct, minimally correlated, indicators related to the quality of human capital. The findings reveal a persistent disparity in responsiveness of territorial healthcare, particularly in southern regions of Italy, where weaknesses in human capital are relevant. These areas often correspond with regions under Health Deficit Recovery Plans and report lower levels of patient satisfaction with medical and nursing care.

1. Premise: objectives and general framework

The potential capability of territorial health care institutions (National Health System - NHS) is a multifaceted phenomenon. In this study, it is analyzed according to the input/outcome model, which has been applied to personnel classified by socio-demographic characteristics. The study of the quality of human capital in health institutions helps to measure the potential capability of these administrations to respond to demographic and technological changes and transformations and it is a valuable aid for policy makers. This is the context for the analysis of a new organizational strategy towards a balanced development of the quality of human capital, with the aim of enhancing and strengthening the skills and knowledge of public health care employees (Inapp, 2021), due to the effects of an ageing population on turnover planning and the definition of educational supply. The input component is explored through a composite index that measures the potential

[1] The paper has been jointly written by all the authors but § 1 and 8 can be attributed to all the authors, while § 2 to A.P.M. Mirto, § 3 and 7 to F. Gaudio, § 4 to F. P. Rizzo, § 5 to A. Pareto, § 6 to D. Maiolino. Elaborations by Comic software and ArcGis maps have been realised by F. P. Rizzo.

functionality of health care institutions. The composite index has already been tested on municipality according to the same input/outcome model (Istat, 2024). However, it should be noted that the *one-size-fits-all model* may not always be applicable, and that it should be considered the specificities of the compartment, the territories, and the historical context.

2. Policy needs and health evaluation

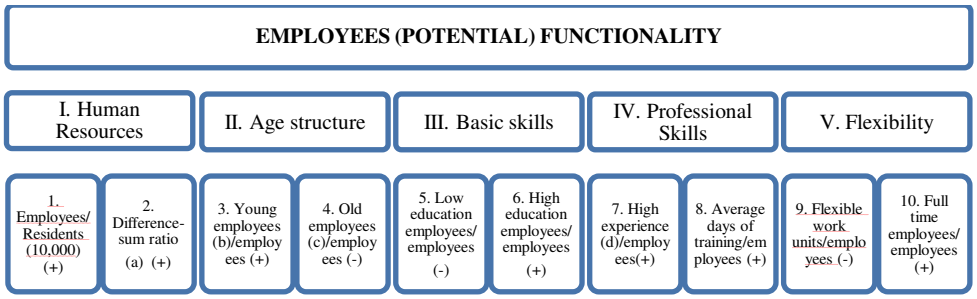
Over the past fifty years evaluation systems - in the public and in the private sector - have undergone a process of incremental evolution which has broadened the scope of analysis from a mainly economic-financial focus to a multidimensional assessment, to an inter-organisational vision (Nuti S. et al., 2021). This promotes the use of methodologies and the collection of information regarding performance to provide all stakeholders with appropriate evidence concerning needs, used resources and achieved results (Walshe and Rundall, 2001). The selected indicators generally cover a plurality of evaluation dimensions, such as population health status, efficiency and sustainability, communication and processes, regional health strategies, internal (employee) evaluation, external (user) evaluation, and clinical care evaluation. This is the context for this contribution, which focuses on one of the organisational components such as human capital. In 2022 the NHS was focused on attempting to capitalise on the opportunities presented by the PNRR investments (S. Anna, 2023). The renewal of regional healthcare systems has long been a subject of national political and scientific debate. The capability of health institutions forms the basis of both planning and evaluation processes. Its analysis contributes to the ongoing argue on ensuring accessible, sustainable and resilient healthcare personnel (Agenas, 2023). In the context of the fundamental principles that underpin the establishment of the NHS (namely, universality, equality and equity), the national intervention in the field of health acts in favour of specific targets of the population characterised by a vulnerable socio-economic situation and contributes to egalitarian access to services (PCM, 2023). Within the Framework of Italy's Partnership Agreement on Cohesion Policy 2021-2027 the National Health Equity Programme (NHEP) was approved. In accordance with the NHEP, the capacity of regional health systems in at least seven regions of the country (Basilicata, Calabria, Campania, Molise, Puglia, Sardegna and Sicilia) should be enhanced. Compared to the rest of the country, these regions have lower levels of compliance with nationally defined standards (LEA, Essential Levels of Care) and greater financial and organisational difficulties in the management of the health service. The following analysis considers NHS bodies, which, unlike other sectors of the public service, are not subject to shift hiring limits but to an expenditure constraint, where and further, more

stringent limitations are laid down in the specific programmes of regions subject to financial recovery plans (MEF, 2024). The operational programmes for the reorganisation and strengthening of the Regional Health Service (later called Deficit Recovery Plans-DRP) came into being with the 2005 Budget Law (311/2004) and were attached to agreements entered by the Ministers of Health and Economy and Finance with the Regions. There are seven regions currently subject to DRP (Abruzzo, Calabria, Campania, Lazio, Molise, Puglia e Sicilia), Liguria and Sardegna concluded their DRP at the end of the first three-year period 2007-2009; Piemonte at the end of the three-year period 2013-2015.

3. Ex-ante evaluation of human resources: framework, individual indicators and regional profiles

From an ex-ante evaluative perspective - hence, at a probabilistic and potential level - a composite index, designated IFO-Health, has been devised to assess the functionality of National Health Service (NHS) institutions. IFO-Health permits the examination of the dimensional and socio-professional adequacy of employees, the primary input variable of these services, and the relationship between these variables and the perceived outcome of the client (satisfaction).

Figure 1 - *Structure of the IFO-Health index.*



(a) (recruited-terminated)/(recruited+terminated); (b) <40 years; (c) >60 years; (d) >20 years

The framework, which has recently been applied to the case of other public institutions (Istat, 2024; Mirto A.P.M., Gaudio F., Abate F., 2023), enhances ten typical proxies of the quality of services to people, with available data provided by the Ministry of Economy and Finance (MEF) Annual Account[2] (Figure 1). The framework is inspired by the input-outcome model and considers the strategic value of personnel in the working process of services. The quantitative endowment, skills

[2] Census survey on public administrations carried out since 1992 by the Ragioneria Generale dello Stato.

and ability to perform sequences of actions functional to the resolution of a socio-technical problem constitute the key element for good organisational functioning (Gori E., Vittadini G., 1999; Gaudio F., 2009). The work considers the most recent available year (2022) by comparing it to 2012.

Table 1 – *Staff of health institutions by region. 2022 (absolute and % values).*

Region	Employees	%	Number per 1,000 residents	Part-time Employees (%)
Piemonte (b)	55,136	8.1	13.0	9.2
Valle d'Aosta	2,138	0.3	17.4	8.2
Lombardia	104,495	15.3	10.5	10.6
Liguria	23,825	3.5	15.8	6.2
Trentino-Alto Adige	17,317	2.5	16.1	31.8
Friuli-Venezia Giulia	20,609	3.0	17.3	6.8
Veneto	63,292	9.3	13.1	11.4
Emilia-Romagna	67,949	10.0	15.3	6.1
Toscana	54,976	8.1	15.0	3.7
Umbria	11,787	1.7	13.8	2.4
Marche	19,700	2.9	13.3	4.3
Lazio (a)	48,502	7.1	8.5	2.1
Abruzzo (a)	14,400	2.1	11.3	1.9
Molise (a)	2,794	0.4	9.6	1.5
Campania (a)	47,404	7.0	8.5	0.4
Puglia (a)	38,892	5.7	10.0	1.4
Basilicata	6,452	0.9	12.0	1.9
Calabria (a)	17,841	2.6	9.7	0.6
Sicilia (a)	43,262	6.3	9.0	1.0
Sardegna (b)	21,084	3.1	13.4	2.0
Italy	***681,855***	***100.0***	***11.3***	***6.2***

(a) Regions subject to Deficit Recovery Plans; (b) Regions exited from Deficit Recovery Plans.
Source: elaborations on MEF – Annual Account.

In 2022, 681,855 people were employed in healthcare, representing 20.8% of the total in the public sector. This denotes a slight increase over 2012 (+1.2%), which can be attributed to the injections aimed at containing the impact of the pandemic. The health sector is characterised by a high concentration of personnel in territorial units, with 185 institutions accounting for 90.3% of the workforce[3]. It also exhibits a considerable degree of differentiation in professional profiles[4].The analysis does not address these specificities; instead, it focuses on diachronic trends and, above all, on territorial differences. This is also in consideration of certain relevant specificities connected to the Deficit Recovery Plans and to the effects on personnel due to financial constraints. In quantitative terms, the contraction in staff numbers recorded until 2019 is mainly ascribable to the regions subject to such discipline (MEF-RGS,

[3] The remaining share: Scientific Care Institutes (20 units, 4.7% of employees); University Polyclinics (10; 2.2%); Environmental Protection Agencies (19; 1.2%); Others (55; 1.6%).
[4] Nursing staff 41.6%; technicians 18.6%; doctors 15.8%; rehabilitation staff 11.9%; administrative staff 9.6%; others 5.7%. See Ministero della Salute (2022).

2024). These regions also present the lowest staff per 1,000 residents and low part-time quotas (Table 1).

4. Main results at the level of individual indicators

An examination of the basic indicators set by the Framework in relation to human resources in the health sector reveals a few important findings and trends (Table 2).

Table 2 – *Individual indicators (with polarity) and trends – 2022 (*).*

Regione	1. Staff per inhabitants (x10,000) (+)	2. Difference-sum ratio (+)	3. % Young employees (< 40 years) (+)	4. % Old employees (> 60 years) (-)	5. % Low education (-)	6. % High education (+)	7. % High experience (+)	8. Average days of training (+)	9. % Flexible work units (-)	10. % Full time employees (+)
Piemonte (b)	129.7	0.09	21.7	14.3	16.1	44.0	46.6	1.3	7.7	90.8
Valle d'Aosta	173.6	-0.21	21.5	9.8	11.9	46.6	39.3	0.4	13.3	91.8
Lombardia	104.7	0.04	26.1	11.5	15.4	48.4	44.2	0.8	5.2	89.4
Bolzano/Bozen	168.3	0.06	21.7	7.8	41.6	35.5	48.9	3.1	17.5	58.1
Trento	153.3	-0.03	29.9	8.2	28.8	45.4	40.1	0.1	6.2	79.1
Veneto	130.5	0.03	27.5	10.7	25.8	48.0	42.1	0.4	2.3	88.6
Friuli-Venezia Giu	172.6	0.11	27.5	10.4	21.1	46.1	29.2	0.6	9.5	93.2
Liguria (b)	158.0	0.22	20.5	15.9	18.0	41.9	42.1	1.7	7.6	93.8
Emilia-Romagna	153.1	0.10	29.5	10.8	10.5	53.6	32.8	2.2	5.9	93.9
Toscana	150.1	-0.01	22.3	15.2	13.5	51.1	38.1	2.9	4.4	96.3
Umbria	137.6	-0.06	22.4	17.2	4.1	65.2	37.5	1.2	8.4	97.6
Marche	132.7	0.00	21.6	14.7	9.3	53.6	35.8	0.3	9.4	95.7
Lazio (a)	84.8	0.19	19.9	18.6	9.5	73.2	39.7	1.6	10.2	97.9
Abruzzo (a)	113.2	0.16	17.7	20.8	9.7	58.7	41.8	0.6	21.5	98.1
Molise (a)	96.1	0.17	13.4	27.6	0.4	67.3	37.1	0.8	11.5	98.5
Campania (a)	84.5	0.29	22.2	24.3	28.2	38.6	36.0	4.6	9.2	99.6
Puglia (a)	99.5	0.00	18.8	20.0	14.1	49.4	27.3	0.8	13.8	98.6
Basilicata	120.0	-0.10	14.7	22.5	6.5	68.4	28.3	1.2	10.4	98.1
Calabria (a)	96.6	-0.02	29.3	25.4	30.0	43.4	32.3	0.2	8.1	99.4
Sicilia (a)	89.9	0.11	12.8	26.1	13.6	52.3	39.0	1.0	19.4	99.0
Sardegna (b)	133.6	-0.24	20.2	21.5	34.5	58.8	34.7	0.3	13.2	98.0
Italia	115.6 ↑	0.03 ↑	23.1 ↑	15.9 ↑	17.2 ↓	50.7 ↑	38.6 ↓	1.4 ↓	8.5 ↑	93.8 ↑

(*) The font color, in the regions, indicates the trend of the 2022 ranking compared to that of 2012: red if decreasing, green if increasing, black if stationary; (a) Subject to Deficit Recovery Plans; (b) Exited from Deficit Recovery Plans.
Source: elaborations on MEF - Annual Account.

Over the period 2012-2022, the ratio of employees to residents per 10,000 inhabitants (ind. 1) increases from 111.7 to 115.6. The regions in the Centre-North have a higher relative endowment than those in the South. In recent years, characterized by the unblocking of recruitment in the public sector (MEF-RGS, 2024), positive effects have been reported in the health sector (as well as in education), both in terms of endowment and, consequently, in terms of age composition. Regarding the first aspect, in 2022 the difference-sum ratio (ind. 2) takes a slightly positive value (0.03), reversing the sign of 2012 (-0.15). This dynamic implies a slight change in the age structure of the workforce compared to 2012. The share of employees under 40 years of age (ind. 3) increased (from 17.4% to 23.1%), with a higher incidence in the north-eastern regions, except for Alto Adige, and a lower one in the south, except for Calabria (from 9% in 2012 to 29.3%). The share of employees aged over 60 also increased over the decade (ind. 4), from

6.2% to 15.9%, with high shares in the Mezzogiorno and rather low shares in the northern regions (especially the North-East). During the decade under review, the share of employees with a low level of education (ind. 5) fell from 22% in 2012 to 17.2% in 2022. The best picture is found in the central regions. At the same time, the proportion of employees with a high level of education (ind. 6) is generally increasing: in 2022, more than one in two employees in the health sector will have a tertiary level qualification (36.4% in 2012). The highest increase will take place in the central regions, and in Lazio this share will be almost 3/4. An unfavorable trend can be observed in the dimension of professional skills. The unblocking of turnover and the positive trend reported by the surplus index have led to a tendency for the share of employees with more than 20 years' experience (ind. 7) to fall from 42.4% in 2012 to 38.6% in 2022, particularly in many southern regions. In 2022, the average number of days of training (ind. 8) also decreases, presumably also as a result of the restrictions imposed by the pandemic, from around 1.8 days in 2012 to almost 1.5 days in 2022. The dynamics of employment flexibility are mixed. The ratio of flexible work units to the total number of permanent employees shows a significant increase (from 4.8% to 8.5%) over the period considered. The regions with the highest values are in the south, those with the lowest in the north, except for Bolzano and Valle d'Aosta. On the other hand, as regards the proportion of full-time employees (ind. 10) - which has slightly increased at national level (from 91.9% to 93.8%) - the southern regions have consistently higher values than the central and northern regions, in contrast to the trend in other public sectors (Istat, 2024).

5. The composite index: methodology and robustness

The composite index was constructed using the Adjusted Mazziotta-Pareto Index (AMPI), which is a partially compensatory aggregation function (Mazziotta and Pareto, 2016). Such a choice is advisable whenever adequate performance on any of the individual indicators is considered critical to overall performance.

The most original aspect of this index is the method of normalization, the so-called “constrained min-max method” (Mazziotta and Pareto, 2021). It normalizes the range of individual indicators in a similar way to the classic min-max method but uses a common reference (base value) that makes it possible to define a ‘balancing model’ (i.e. the set of values that are considered balanced). In this way, it is possible to compare the values of the units both spatially and temporally with respect to a common reference that does not change over time. Let us consider the matrix $\mathbf{X}=\{x_{ijt}\}$ with 21 rows (Italian regions), 10 columns (individual indicators), and 2 layers (years) where x_{ijt} is the value of individual indicator j, for region i, at year t. A normalized matrix $\mathbf{R}=\{r_{ijt}\}$ is computed as follows:

$$r_{ijt} = 100 \pm \frac{x_{ijt} - x_{j0}}{\max_{it}(x_{ijt}) - \min_{it}(x_{ijt})} 60$$

where $\min_{it}(x_{ijt})$ and $\max_{it}(x_{ijt})$ are, respectively, the overall minimum and maximum of indicator *j* across the two years (goalposts), x_{j0} is the Italy average in 2012 (base value) for indicator *j*, and the sign ± depends on the polarity of indicator *j*.

Denoting with $\mathrm{M}_{r_{it}}$, $\mathrm{S}_{r_{it}}$, $\mathrm{cv}_{r_{it}}$, respectively, the mean, standard deviation, and coefficient of variation of the normalized values for region *i*, at year *t*, the composite index is given by:

$$\mathrm{AMPI}_{it} = \mathrm{M}_{r_{it}} - \mathrm{S}_{r_{it}} \mathrm{cv}_{r_{it}}$$

where:

$$\mathrm{M}_{r_{it}} = \frac{\sum_{j=1}^{10} r_{ijt}}{10} \qquad \mathrm{S}_{r_{it}} = \sqrt{\frac{\sum_{j=1}^{10} (r_{ijt} - \mathrm{M}_{r_{it}})^2}{10}} \qquad \mathrm{cv}_{r_{it}} = \frac{\mathrm{S}_{r_{it}}}{\mathrm{M}_{r_{it}}}.$$

The version of AMPI with a negative penalty was used, so that an unbalance among indicators has a negative effect on the value of the index.

The robustness of the index was assessed through an influence analysis aimed at studying whether and to what extent the ranking of the Italian regions changes when a single indicator is removed from the initial set.

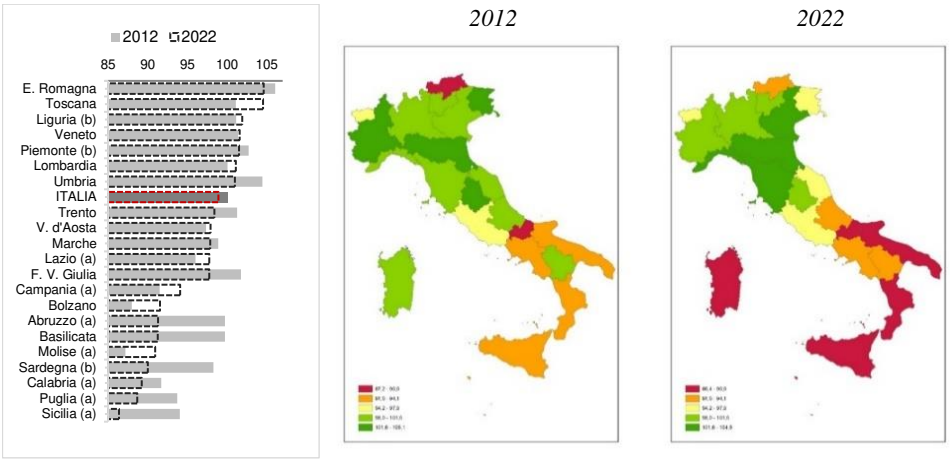
Starting from the 10 individual indicators, 10 replications were performed by eliminating a different indicator each time and calculating the values of the composite index based on the remaining 9 indicators. For each replication, the ranking of the Italian regions was created and the absolute difference of rank between the position in the original ranking and the position in the ranking corresponding to the 9 indicators was calculated for each region. Finally, the Mean Absolute Difference of rank (MAD) was obtained for each indicator removed. In 2012, the MAD varies between 0.38 (high education) and 2.10 (young employees) positions, while in 2022 it varies between 0.48 (high education) and 2.29 (high experience). Moreover, the standard deviation of MAD is 0.537 in 2012 and 0.516 in 2022. Therefore, the variability of the rankings appears to remain constant over time.

6. Composite index: main evidence and territorial gaps

The composite index IFO-Health offers an analytical overview of the set of individual indicators covered by the operational model. In order to facilitate geographical and temporal comparisons, the 2012 Italy value (100) has been indexed.

At the diachronic level, a general downward shift in the (potential) functionality of the NHS workforce over the decade under consideration is confirmed. This finding is consistent with what has been observed with a similar approach in other areas of the public sector (Istat, 2024). In particular, the decline in the national average value (from 100 in 2012 to 98.9 in 2022) can be attributed to the significant reduction observed in 12 regions, with a particularly pronounced decline evident in much of southern Italy (Figure 3). Concurrently, a clear divergence can be observed between the IFO values in the Centre-North – particularly in the main northern regions, which are above the national average and generally progressing – and those in the South, all of which are far from the average and generally in regression despite the existence of deficit recovery plans. Within this framework, there are a few exceptions, which will be mentioned later. Nevertheless, these observations do not alter the substance of a dualistic model at a territorial level, as showed by the related cartographic evidence (Figure 3), which demonstrates a notable strengthening over the decade in question. As already pointed out in the analysis of the individual indicators, most influential in the widening of the gap between North and South are the increase in the ratio of employees to residents in the North compared to the South, the indicator relating to the share of employees over 60 years of age (significant growth in the South) and the ratio of flexible work units to total permanent employees, which grew more in the South. Upon closer examination, several noteworthy elements emerge. In 2022, Emilia-Romagna is confirmed as the most virtuous region, although its IFO score has declined slightly. Toscana has demonstrated clear progression, advancing six positions since 2012 and becoming the second-best performer. At the opposite end of the distribution is Sicilia, which has experienced notable regression. In regions with a more pronounced potential functionality gap in terms of employees compared to the national average, all cases affected by Deficit Recovery Plans stand out. The Province of Bolzano, however, has improved its ranking, moving from 20th to 14th. Among the regions exhibiting a positive trend are Lazio, Campania and Molise, which have experienced the highest growth rate (3.8). Conversely, the remaining regions of Abruzzo, Calabria, Puglia and Sicilia have demonstrated discrepancies, which appear to be worsening.

Figure 3 – *IFO-Health by region.*



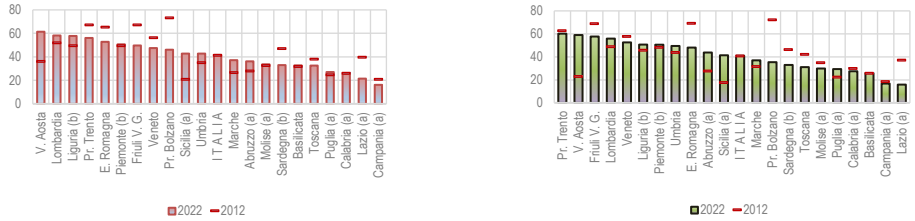
Source: Elaborations on MEF - Annual Account.

In Mezzogiorno, Basilicata, Abruzzo and Sardegna which are already very critical cases, all recorded a contraction of more than 8 points. Other regions such as Sicilia (- 7.7), Puglia (- 5.1) and Calabria (- 2.5) helped confirmed that the negative trend with the gap compared to the regions of the North has widened further. Significant reductions were also observed in Umbria (- 3.5), Friuli-Venezia Giulia (- 4.0) and the Province of Trento (- 2.9).

7. Composite index and perceived outcome (satisfaction)

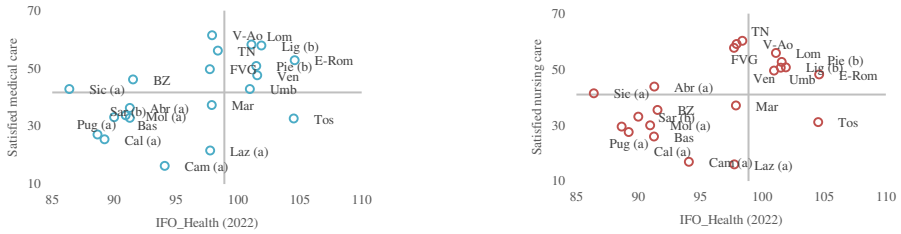
Patient satisfaction is a relevant outcome dimension: it is the perceived quality resulting from a subjective assessment between what was expected and what was received (or perceived as such). It is functional for continuous improvement, subject to certain limitations related to its subjective nature (Ovretveit J., 1992; Pagano A., Vittadini G., 2004). Despite a critical context marked by the pandemic emergency, in Italy there has been a very slight increase over ten years in the number of people who are very satisfied with the assistance they receive, which exceeds 4 out of 10 cases both in the medical field (from 41.3% to 41.7%) and in the nursing field (from 40.4% to 41%, Figure 4).

Figure 4 - *People very satisfied with medical care (left) or nursing care (right), by region (*). 2012 and 2022 (% values).*



(*) Persons with at least one hospitalisation in the three months preceding the interview; (a) Regions subject to Deficit Recovery Plans; (b) Regions exited from Deficit Recovery Plans.
Source: Elaborations on MEF - Annual Account.

Figure 5 - *Organisational functionality (IFO-Health) and persons highly satisfied with medical (left) and nursing (right) care (*) by region. 2022 (absolute and % values).*



(a) Regions subject to Deficit Recovery Plans; (b) Regions exited from Deficit Recovery Plans
Source: elaborations on MEF data and Istat, Multipurpose Survey - Aspects of Daily Life.

Satisfaction decreases from North to South, with some significant exceptions: Tuscany - below the national average and in decrease - and the Province of Bolzano, in visible retreat. Excluding Sicilia (where satisfaction improves) and Abruzzo in part, the problematic regions are all long-standing subject to recovery plans: Puglia (since 2010), Calabria (since 2009, now a commissioner), Lazio (since 2007) and Campania (since 2007). The Province of Trento, Valle d'Aosta, Lombardia and Liguria are virtuous cases. The Province of Bolzano, and the rest of the north-east, report a regression on which the pandemic event may have had an influence, which hit those territories hard during the onset and most acute phase (ISTAT-ISS, 2022). Similarly to what was observed in the case of the municipalities (Istat, 2024), there is a visible link between the IFO-Health index and patient satisfaction, while is quite evident in certain contiguous groupings of the South and the Islands, on the one hand, and of the

Centre and above all of the North, on the other, with a certain consistency between the two performance areas considered (Figure 5). The main southern regions, except for Sicilia, tend to be in the third quadrant (low satisfaction and low functionality), both for the medical sector (Figure 5, left) and for the care sector (Figure 5, right). On the other hand, the main regions of central and, above all, northern Italy, which are characterised by high functionality, also report virtuous levels of satisfaction and generally place themselves in the first quadrant. Apart from a few exceptions (above all Lazio, Sicilia and Toscana), there is a rather clear differentiation on a territorial basis, a kind of dual model that can be extended to the position of each region regarding the deficit reduction plans: the regions that are still subject to these measures converge in the group characterised by low satisfaction and low functionality.

8. Final remarks and future perspectives

During the last decade, the decline in the IFO-Health national average value can be attributed to the significant reduction observed in 12 regions, with a particularly pronounced drop evident in much of southern Italy - which are far from the average and generally in regression despite the existence of Deficit Recovery Plans.

Future perspectives of this contribution are related to the deepening of the context in which the analysis is carried out, considering the end of the pandemic, PNRR measures and the consequences for employment in health care institutions (MEF, 2024; CE, 2022). One of the main improvements in the model considers the time series analysis of employment by type of contract, gender and qualification to study the development and improvement of skills and capacity for workforce planning at national and regional level. Finally, the study of all model components (process, output and outcome) is a crucial issue also to evaluate the segment of health institutions by type of public body (Istat, 2024).

References

AGENAS. 2024. La struttura e la metodologia del sistema di Agenas per il monitoraggio della performance delle aziende ospedaliere, *Monitor,* 2024.

ANESSI PESSINA E, CANTU' E. 2007. L'aziendalizzazione della sanità in Italia, Rapporto OASI.

EUROPEAN COMMISSION. 2022. Decisione della UE su approvazione del programma "PN Equità nella Salute 2021-2027".

GAUDIO F. 2009. *Monitoraggio e valutazione dei servizi sociali*. Dispense, Università della Calabria, Facoltà di Scienze Politiche.

GORI E., VITTADINI G. 1999. *Qualità e valutazione nei servizi di pubblica utilità*, ETAS Libri.
INAPP. 2021. Rapporto dell'Istituto nazionale per l'analisi delle politiche pubbliche. Lavoro, formazione e società in Italia nel passaggio all'era post Covid-19. Luglio.
ISTAT. 2024. Comuni: vincoli strutturali e opportunità del PNRR, Statistiche Focus.
ISTAT-ISS. 2022. Impatto dell'epidemia Covid-19 sulla mortalità totale della popolazione residente. Anni 2020-21 e gennaio 2022.
MAZZIOTTA M., PARETO A. 2020*, Gli indici sintetici*, GIAPPICHELLI ed. Torino.
MAZZIOTTA M., PARETO A. 2022. Normalization methods for spatio-temporal analysis of environmental performance: Revisiting the Min–Max method, *Environmetrics*, DOI: 10.1002/env.2730.
MEF-RGS. 2024. Commento ai principali dati del Conto Annuale del periodo 2013-2022.
MINISTERO DELLA SALUTE. 2022. *Personale delle ASL e degli Istituti di ricovero pubblici ed equiparati.*
MIRTO A.P.M., GAUDIO F., ABATE F. 2024. Ageing and functionality in public administration employment: a case study for Italian municipalities, *RIEDS,* Vol. 77, No. 2.
NUTI S., VINCI A., VOLA F. 2021. Stato e regioni: le performance dei sistemi sanitari regionali, *Italian Journal of Social Policy*, Vol. 2.
OVRETVEIT J. 1992. *Health service quality*, Blackwell Scientific Publication.
PAGANO A. E G. VITTADINI (a cura di). 2004. Q*ualità e valutazione delle strutture sanitarie*, ETAS.
PCM-DIPARTIMENTO PER LE POLITICHE DI COESIONE. 2023. Accordo di partenariato Italia, 2021-2027.
SANT'ANNA – Scuola Universitaria Superiore Pisa. 2023. Il sistema di valutazione della performance dei sistemi sanitari regionali, Report 2022.
WALSCHE K., RUNDALL T.G. 2001. Evidence-Based Management: From Theory to Practice in Health Care, *The Milbank Quarterly*, Vol. 79, No. 3, pp. 429-457.
ZAMAGNI S., VENTURI P., RAGO S. 2015. Valutare l'impatto sociale. La questione della misurazione nelle imprese sociali, *Impresa sociale,* No. 6, dicembre.

Francesca ABATE, Istat, zagara5812@gmail.com
Anna Pia Maria MIRTO, Istat, mirto@istat.it
Francesco GAUDIO, Istat, gaudio@istat.it
Adriano PARETO, Istat, pareto@istat.it
Francesco Paolo RIZZO, Istat, frrizzo@istat.it
Dario MAIOLINO, Istat, maiolino@istat.it

ENHANCING ENVIRONMENTAL AND HEALTH STATISTICS THROUGH ARTIFICIAL INTELLIGENCE: A COMPARATIVE STUDY OF IMPUTATION TECHNIQUES[1]

Simona Cafieri, Francesco Pugliese, Mauro Sodani

Abstract. In an increasingly globalized world, addressing health, environmental sustainability and social inequalities is crucial and requires an integrated approach involving national statistical offices. The latter is increasingly called upon to develop statistical frameworks to facilitate informed policy-making. However, incomplete or missing data in questionnaires or registers may compromise the accuracy and reliability of results.

The main objective of this study is to assess the effectiveness of different imputation methods using machine learning (ML) and artificial intelligence (AI) techniques in dealing with missing data in social surveys. To this end, a comparative analysis of different imputation techniques has been carried out, based on real datasets from the Istat Multi-purpose Household Survey, where missing data are common. Preliminary results suggest that ML/AI-based imputation methods outperform traditional statistical techniques in terms of performance and robustness.

The aim is to improve imputation techniques in official statistics to improve data quality on critical issues.

1 Introduction

The rise of artificial intelligence (AI) is having a significant impact on official statistics. AI methods provide solutions to data incompleteness and support informed decision-making (Sun et al., 2023). National Statistical Institutes have developed frameworks to support policy decisions, particularly in relation to environmental sustainability, health and social inequalities (Rigo, 2022).

In Italy, these issues are of particular importance as they form the basis of the BES (balanced and sustainable well-being) indicators that underpin the government's economic and financial planning document (Istat, 2024). It is worth noting that a significant proportion of these indicators are based on survey data, which are prone to inaccuracy and unreliability if incomplete or missing responses are not addressed. The application of AI is a promising solution to the problem of missing data in surveys. Such methods can be used to predict and impute missing values, thereby improving the overall quality of statistical datasets. Traditional techniques such as

[1] This work is the result of a close collaboration among the authors.

mean or median imputation often introduce bias, whereas AI-based methods can provide more accurate and impartial estimates. Several machine learning and deep learning models can be employed to impute missing data in official statistics. A substantial body of research has already been conducted to apply these methods, with positive results (De Fausti et al., 2023). This paper explores how AI can improve the quality of official statistics, focusing on the potential of machine learning (ML) and deep learning (DL) to improve the accuracy and reliability of health and environmental data.

2 Related Works

According to the literature, many deep learning and machine learning algorithms have been considered for 'data imputation'. Support vector machines (SVMs), a supervised learning method that identifies optimal hyperplanes for class separation in high-dimensional spaces, have been used to deal with missing values (Honghai et al., 2005). In addition, decision trees, which divide data into branches for decision-making, are evaluated for interpretability and used for data imputation (Nikfalazar, 2020).

XGBoost (Extreme Gradient Boosting) improves model performance by combining predictions from multiple estimators and is known for its efficiency (Mitchell, 2017; Rusdah & Murfi, 2020). The k-Nearest Neighbours (KNN) algorithm classifies data points based on their nearest neighbours and is popular for missing value imputation (Guo et al., 2003; Pujianto et al., 2019).

Linear regression models relationships between variables and is often used for predictive analysis (Montgomery et al., 2021). Random forest (RF) constructs multiple decision trees to improve model accuracy (Breiman, 2001) and has been used to impute missing values (Tang & Ishwaran, 2017). Long short-term memory (LSTM) networks capture long-term dependencies in sequential data and have been used to impute missing values (Hochreiter & Schmidhuber, 1997; Yuan et al., 2018). Gated recurrent units (GRUs) provide a simplified architecture for a similar task (Dey & Salem, 2017; Wang et al., 2022).

3 Methods

One of the main sources of social and household health data in Italy is the Aspects of Daily Life (AVQ) survey, carried out annually by Istat (Istat, 2022). AVQ represents an integral component of a unified system of social surveys. Indeed, collecting data is indispensable for understanding the daily lives of individuals and households. The survey provides information on the habits of citizens and the problems they face in everyday life through interviews with a sample of 20,000 households, representing approximately 50,000 individuals. Since 2018, the survey

has been carried out using a sequential CAWI/CAPI mixed-mode technique. The survey investigates a range of social aspects, including education, employment, family and social life, leisure time, political and social participation, health, lifestyles, access to services and other factors relevant to the study of quality of life. These topics are investigated from a social perspective, with particular consideration given to behaviors, motivations and opinions as key elements in the definition of social information. The survey is included in the National Statistics Plan, which collates the statistical investigations that are required for the Country. However, it is not uncommon for questionnaires to be incomplete, which can affect the precision and dependability of the resulting data.

To address this issue, we designed and implemented an imputation experiment by using the AVQ dataset from the 2021 survey comprising 735 variables. The presence of missing values in the dataset, frequently represented as blanks or NaN, is incompatible with scikit-learn estimators, which require all values to be numeric and significant. A fundamental approach is the complete case analysis, whereby rows (dropNA) or columns with missing values (list-wise deletion) are excluded. Nevertheless, this may result in a significant reduction of the available information. An effective strategy is to impute missing data by inferring it from the available data. Conversely, techniques such as the use of central tendency measures (mean, median, etc.) can be employed. This approach appears to be relatively straightforward and robust. However, there is a risk of underestimating or overestimating the true values, which could introduce bias into the resulting estimates. This phenomenon occurs when an algorithm produces results that are systematically biased due to incorrect assumptions, which are typically present in the data set or in the machine learning process.

In this study, we used missing data imputation techniques known as 'regression imputation'. Essentially, this method estimates missing values using a regressor (e.g. support vector regressor or random forest regressor), with the missing variable as the target and the other variables as inputs. Regression imputation is divided into 'deterministic' and 'stochastic'. The main difference between these two approaches is how the missing values are estimated and how uncertainty is taken into account. In deterministic regression imputation, missing values are estimated using a deterministic relationship between the variables. The trained regression model is used to predict the missing values for incomplete observations. The predicted value is used directly as an estimate for the missing value, hence the term deterministic. Deterministic imputation does not take into account the uncertainty associated with the estimate, so the predicted values are always the same for a given combination of input values. In contrast, stochastic regression imputation incorporates the uncertainty of the estimate into the imputation process. A stochastic noise term is added to the prediction. This noise term can be generated using the distribution of

residual errors from the regression model. For example, if the regression model has a residual variance of sigma squared, noise can be added extracting it from a normal distribution with a mean of 0 and a variance of sigma squared, resulting in a more realistic estimate. However, in our work we only used deterministic regression imputation, with the intention of exploring stochastic imputation adapted to machine learning and deep learning models in the future. As a final observation, we cannot directly impute regression values before preprocessing. The input predictor variables also contain missing data themselves, which would cause issues for the machine learning and deep learning models, as the libraries we used (i.e., scikit-learn or Keras) do not accept null values. Therefore, we imputed all input variables with missing values using a method called "Simple Random Imputation," which involves replacing the missing value with a random value. It is proven that this approach does not significantly affect the final estimate given the large number of variables present in our dataset. It is crucial to emphasize that the dataset for the subsequent machine learning models must adhere to specific rules and requirements to ensure optimal final imputation. Missing data should be random or at least analyzable; for example, it should be Missing Completely at Random (MCAR) or Missing at Random (MAR). In the event that the data is missing not at random (MNAR), the process of imputation becomes more complex. Regarding the percentage of missing data, datasets exhibiting a markedly high percentage of missing values (>50%) may potentially compromise the accuracy of the imputation models. It is optimal for the percentage of missing data to fall within the range of 12% to 50%. In the case of datasets with a missing data percentage lower than 12%, it would be preferable to utilize simpler imputation methods, such as setting the missing values to the mean of the variable or other constant values. Moreover, a significant proportion of the variables must remain complete, as these data points are indispensable for reconstructing the relationships between variables and predicting the missing ones. In particular, when employing deep learning models, it is crucial to ensure that the dataset is sufficiently large to enable effective training of the model, as these models typically require a significant number of samples to generalise effectively. Conversely, for a traditional machine learning model, the inclusion of a greater number of useful variables facilitates a more accurate prediction of the missing values. Additionally, the dataset must be structured in a way that meets the statistical production requirements. For example, existing values must be consistent with domain rules. In the case of a variable representing age, for instance, negative values or values exceeding a realistic threshold should be avoided. Imputation models must be capable of accommodating both qualitative variables, which should be accurately encoded using one-hot or ordinal encoding, and quantitative variables, which should be standardized or normalized in accordance with the models employed. The imputation process may be distorted by extreme outliers, particularly in the case of

traditional models such as KNN or regression. Outliers should be identified and managed, either through transformations or exclusion. Variables must have a stable and meaningful distribution, as this enables the models to accurately capture the patterns. Variables with highly imbalanced distributions may reduce the effectiveness of the imputation. Finally, for models such as linear regression or logistic regression, it is essential to reduce multicollinearity (strong correlation between independent variables), as this could negatively affect the imputation. This can be addressed using techniques such as the Variance Inflation Factor (VIF). In the literature, this method is more efficient than other methods of replacing predictor variables with a zero, the mean, etc. (Kalton & Kish, 1984).

In this work, we trained all the traditional machine learning models described in the previous section: SVM, DT, RF, XGBoost, KNN, and the most recent deep learning models: MLP, LSTM, GRU, CONV1D. The objective of the training was to create models for the imputation of the following health-related variables: The variables of interest were body mass index (BMI) for individuals aged 18 and over. The same models have been trained for the imputation of additional environmental variables, namely SODPOAP (resident satisfaction with household waste collection services). The ML and DL models were evaluated using a range of metrics that are appropriate for regression problems. The Root Mean Squared Error (RMSE), the Mean Absolute Percentage Error (MAPE), the Mean Absolute Error (MAE) and the R2 Score were employed for the assessment of the models. We observed that even though the RMSE yields absolute values, it was an adequate metric for comparing the performance of the different models and for their ranking. Accordingly, the trained models were ordered in descending order of RMSE, with the most effective model identified as the one with the lowest error. Furthermore, models that demonstrated minimal overfitting, as evidenced by a minimal discrepancy between the RMSE values for the training and test sets, were deemed the most optimal. Prior to the commencement of the training phase, a train-test split was conducted, with the test set comprising 20% of the entire dataset. No pre-processing was applied to the traditional machine learning models, given that the data is entirely numeric and lacks qualitative variables. For the deep learning models based on neural networks, all numerical inputs were normalized between 0 and 1. The training phase used Python's Hyperopt framework to optimize hyperparameters for scikit-learn (ML) and Keras (DL) models. Given the strong impact of initial hyperparameters, optimization was integrated directly into each model's training, adding computation time but ensuring optimal model selection, as Hyperopt returns the best model similarly to RandomizedSearch and GridSearch. Hyperopt is more efficient than these methods, using the Tree-structured Parzen Estimator (TPE) and Bayesian inference. TPE classifies parameter sets as "good" or "bad," focusing sampling on "good" regions. Iteratively, Hyperopt refines probability estimates, narrowing the search and

optimizing models faster. All results here were obtained using Hyperopt. The K-nearest neighbour (KNN) method exhibited the shortest training time, at 0.03 seconds, while the Gated Recurrent Units (GRUs) method demonstrated the longest training time, at 143 seconds. The remaining methods exhibited an average runtime of approximately 20 seconds, with the exception of Decision Trees and Linear Regression, which completed their execution in 0.4 seconds. Furthermore, XGBoost is optimized in terms of speed in comparison to traditional Gradient Boosting, which is known for its slow processing time and was therefore not included in this analysis. Regarding the deep learning models (MLP, Conv1D, LSTM, and GRU), the hyperparameters employed included a batch size of 32 (to ensure a sufficient degree of parallelism without unduly limiting the potential for parallelism or causing performance issues if the batch size were too high), a learning rate of 0.1, and 50 training epochs for all models. The deep learning models required an average training time of 100 seconds, except the MLP, which required only 21 seconds.

4 Results

Upon completion of the training phase, it was observed that the combination of the most effective models varied depending on the variable being imputed. This finding may be related to the No-Free-Lunch Theorem (ADAM, et al. 2019), which sets a theoretical limit in machine learning. The No-Free-Lunch Theorem postulates that no single optimal machine learning model exists for every task. Consequently, the strength of our method lies in its capacity to identify the optimal model for each variable (for each variable, the task is completely distinct) to be imputed, which is meticulously selected from a vast array of models. About the health-related variable "BMI", the results for all the models are presented in Table 1.

Table 1 – *Table of metrics of the models trained to predict the variable "BMI".*

MODEL	Training RMSE	Test RMSE
LSTM	0.7489	0.7546
GRU	0.7423	0.7551
CONV1D	0.7637	0.7854
MLP	0.7866	0.8122
SVM	0.8524	0.8823
KNN	0.6954	0.7949
LR	0.6400	0.7936
XG Boost	0.3898	0.7318
RF	0.2666	0.7240
DT	0.0000	1.0143

It can be seen that the Deep Learning model is the most effective. Long Short-Term Memory (LSTM) model. Table 2 presents a comparison of the descriptive

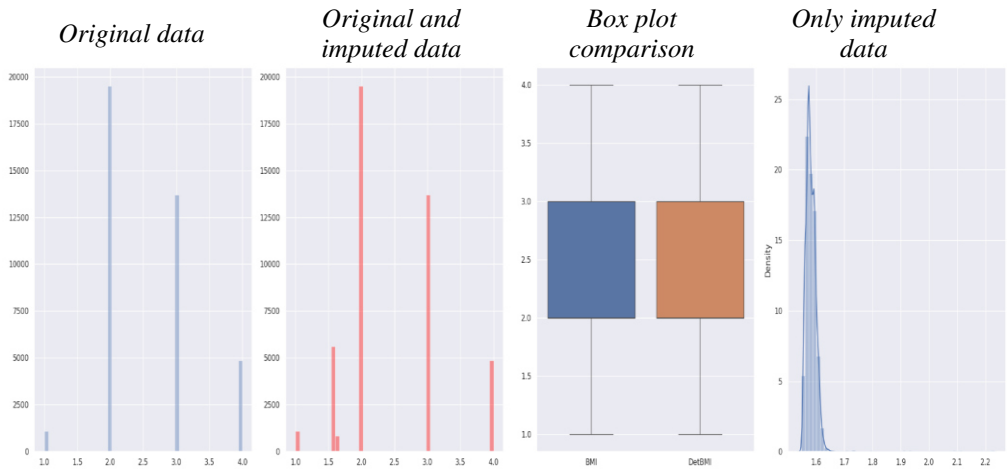
statistics (mean, standard deviation, and quartiles) calculated for the original variable prior to imputation and the imputed variable. This preliminary assessment indicates that the imputed distribution is not markedly disparate from the original distribution.

Table 2 – *Table of comparisons between descriptive statistics of BMI and Imputed BMI (DetBMI).*

MODEL	Mean	STD	MIN	25%	50%	75%	MAX
BMI	2.5695	0.7413	1	2	2	3	4
DetBMI	2.4296	0.7683	1	2	2	3	4

Figure 1 provides a comparison between the original distributions (omitting the nulls in the BMI column) and the distribution with imputed data (the BMI column without nulls plus imputed values).

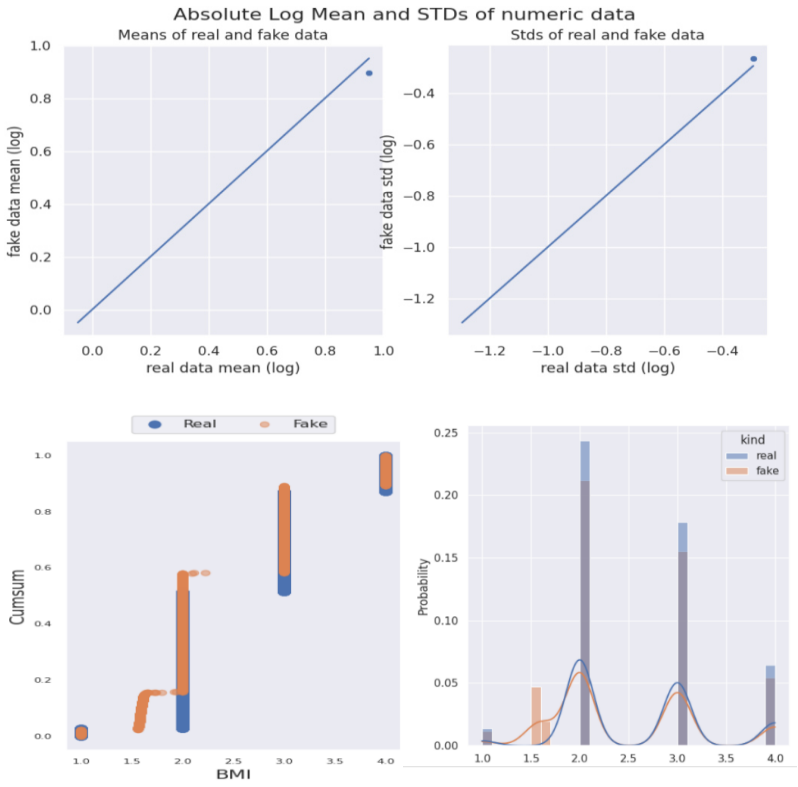
Figure 1 – *Comparisons among distributions charts for the variable BMI.*



Furthermore, an analysis of the box plots of both distributions and the distribution of only the imputed values that replace the nulls is provided. As can be observed, the two distributions are similar, although a slight margin of error is to be expected at this stage. Figure 2 presents a further comparison between the mean and standard deviation of the imputed data and the original data. Furthermore, a comparison between the original and imputed univariate and cumulative

distributions is presented. The distributions are highly similar, which reinforces the assertion that AI models can markedly enhance the handling of missing data.

Figure 2 – *Comparisons among means, standard deviations, cumulate and univariate distributions charts.*



With regard to the environmental variable SODPOAP, Table 3 provides a comparison of the metrics for all models, indicating that the Multi-Layer Perceptron (MLP) is the optimal model in this context. The most recent models exhibit a proclivity for overfitting, as evidenced by the markedly lower error rate on the training set in comparison to the test set. It is also noteworthy that, as anticipated, deep learning models demonstrate superior performance compared to machine learning models on these high-dimensional imputation datasets. Indeed, deep learning models consistently rank among the top performers, irrespective of whether they are recurrent or not. This suggests that the longitudinal (temporal) aspect of the data does not influence the models' performance in this particular dataset. In

comparison to traditional machine learning models, simpler models such as SVM and LR appear to demonstrate superior performance, whereas more sophisticated ensemble models like XGBoost and RF tend to exhibit a higher propensity for overfitting

Table 3 – *Table of metrics of the models trained to predict the variable "SODPOAP".*

MODEL	Training RMSE	Test RMSE
MLP	0.6843	0.6970
GRU	0.6048	0.6190
CONV1D	0.6169	0.6313
LSTM	0.6020	0.6165
SVM	0.6031	0.6189
KNN	0.5347	0.6115
LR	0.5093	0.6195
XGBoost	0.2915	0.5315
RF	0.1964	0.5337
DT	0.0000	0.7211

In Table 4, we present comparisons between descriptive statistics.

Table 4 – *Table of comparisons between descriptive statistics of SODPOAP and Imputed SODPOAP (DetSODPOAP).*

MODEL	Mean	STD	MIN	25%	50%	75%	MAX
SODPOAP	1.8563	0.6260	1	1	2	2	4
DetSODPOAP	1.7666	0.5532	1	1.5	2	2	4

Figure 3 presents a comparison of the original and imputed distributions of SODPOAP, while Figure 4 provides a comparison of the means, standard deviations, univariate, and cumulative distributions. As with the previous results, excellent results are obtained. However, it can be observed that the outcome and behavior of the models change depending on the difficulty level of the variable. The principal advantage of our methodology is the construction of a bespoke model for each variable to be imputed. The results demonstrate that deep learning models, such as long short-term memory (LSTM) and gated recurrent unit (GRU) networks, exhibit high performance in terms of root mean square error (RMSE) on every task, indicating their suitability for handling sequential data. Random Forest and XGBoost also demonstrated satisfactory performance, however, they tended to over-fit, rendering them unsuitable for the task of the imputation of missing values. The support vector machine models demonstrated reliable performance in imputation,

although they exhibited slightly higher root mean square error (RMSE) compared to the top-performing models.

Figure 3 – *Comparisons among distributions charts for the variable SODPOAP.*

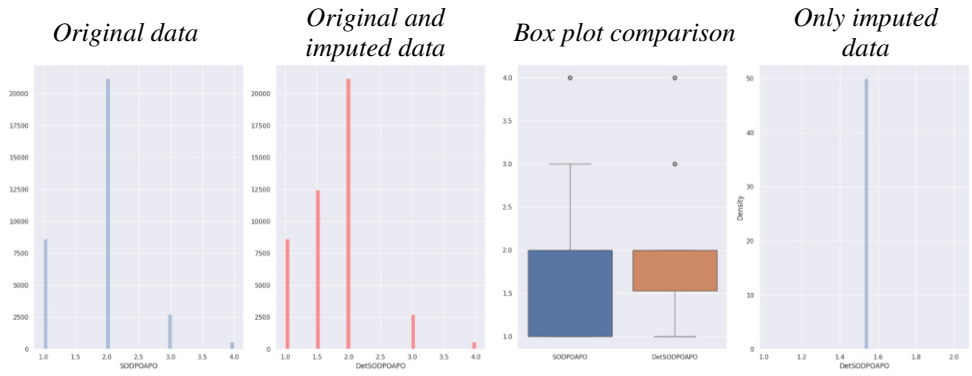
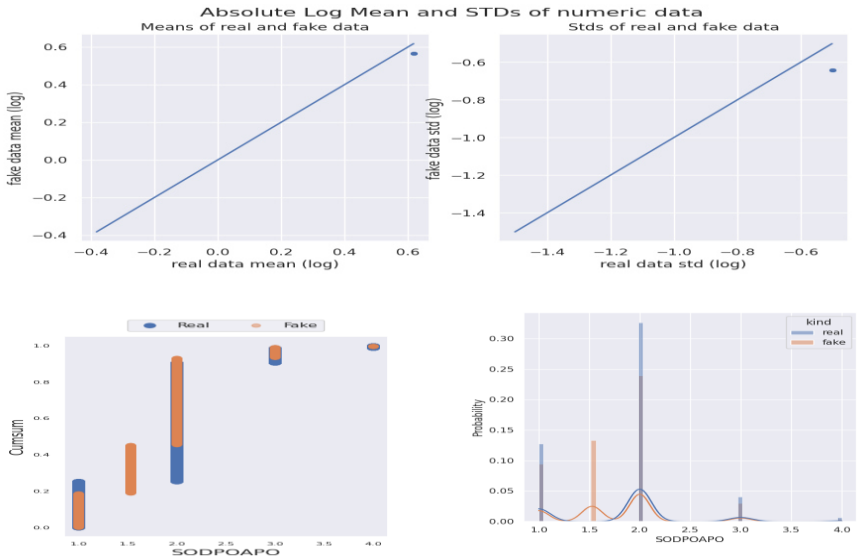


Figure 4 – *Comparisons among means, standard deviations, cumulate and univariate distributions charts.*



5 Conclusion

The examination of health and environmental statistics offers a promising avenue for improving the quality and reliability of data. It should be noted that this research was conducted on a single dataset; however, it could have been performed on multiple datasets. However, we encourage you to consider a few points. The dataset is not a simple one; rather, it is essential for Istat's production needs. The objective was to concentrate on this particular dataset, comprising over 700 variables, to ascertain the limits of AI capabilities in the context of this specific imputation problem. Furthermore, the No-Free-Lunch theorem indicates that no optimal model or AI exists that can perform equally well on all tasks and datasets. All such applications are domain-specific or task-specific, and thus dataset-oriented. Consequently, even if the AI had been tested on all imputation datasets, it would not be possible to guarantee that it would always work. It can be stated with certainty that AI-based imputation methods can effectively address the issue of missing data, thereby enhancing the overall integrity of statistical surveys. However, it is essential to note that the most suitable model must be carefully designed and implemented for each imputation problem, as there is currently no general AI (AGI) that can be applied to all types of problems and datasets. A comprehensive analysis allows for the selection of the optimal model, the imputation of missing data using this model, and the subsequent evaluation of the quality of the imputed data. Future research should prioritize the development of a stochastic regression imputation method that more effectively preserves the variance between the original and imputed distributions, as well as the investigation of advanced models, such as Transformers and Generative Adversarial Networks (GANs), to further enhance the imputation process.

References

ADAM S P., ALEXANDROPOULOS S. A. N., PARDALOS P. M., VRAHATIS M. N. 2019 No free lunch theorem: A review. *Approximation and optimization: Algorithms, complexity and applications*, pp. 57-82.

BREIMAN L. 2001. Random forests. *Machine learning*, Vol. 45, pp. 5-32.

DE FAUSTI, F., DI ZIO M., FILIPPINI R., TOTI S., ZARDETTO, D. 2023. A study of MLP for the imputation of the "Attained Level of Education" in Base Register of Individuals. In: *WORKSHOP ON METHODOLOGIES FOR OFFICIAL STATISTICS*. p. 69.

DEY R., SALEM F. M. 2017. Gate-variants of gated recurrent unit (GRU) neural networks. In *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)* (pp. 1597-1600), IEEE.

GUO G., WANG H., BELL D., B, Y., GREER K. 2003. KNN model-based approach in classification. In *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, Catania, Sicily, Italy, November 3-7, 2003. Proceedings* (pp. 986-996). Springer Berlin Heidelberg.

HOCHREITER S., SCHMIDHUBER J. 1997. Long short-term memory, *Neural Computation,* Vol. 9, No. 8, pp. 1735-1780.

HONGHAI, F, GUOSHUN, C., CHENG, Y., BINGRU, Y., & YUMEI, C. 2005. A SVM regression based approach to filling in missing values. In: *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*. Springer Berlin Heidelberg, 2005. p. 581-587.

ISTAT. 2022. *Indagine Aspetti della vita quotidiana 2021.*

ISTAT 2024. *Rapporto BES 2023.*

KALTON G., KISH L. 1984. Some efficient random imputation methods, *Communications in Statistics-Theory and Methods*, Vol. 13, No.16, pp. 1919-1939.

MITCHELL R., FRANK E. 2017. Accelerating the XGBoost algorithm using GPU computing. *PeerJ Computer Science*, Vol 3: e127.

MONTGOMERY D. C., PECK E. A., VINING G. G. 2021. *Introduction to linear regression analysis*. John Wiley & Sons.

NIKFALAZAR S., YEH C. H., BEDINGFIELD S., KHORSHIDI H. A. 2020. Missing data imputation using decision trees and fuzzy clustering with iterative learning, *Knowledge and Information Systems*, Vol. 62, pp. 2419-2437.

RIGO A. 2022. *Programmazione e innovazione: il percorso verso l'efficienza interna delle Pubbliche Amministrazioni*.

SUN Y, LI J., XU Y., ZHANG T., WANG X. 2023. Deep learning versus conventional methods for missing data imputation: A review and comparative study. *Expert Systems with Applications*, Vol. 227: 120201.

TANG F., ISHWARAN H. 2017. *Random forest missing data algorithms*. Statistical Analysis and Data Mining: ASA Data Science Journal, Vol. 10, No. 6, pp. 363-377.

Simona CAFIERI, Istat, cafieri@istat.it
Francesco PUGLIESE, Istat, frpuglie@istat.it
Mauro SODANI, Istat, sodani@istat.it

ENVIRONMENTAL CONCERN: A CULTURE OF TRUST

Thomas Bassetti, Filippo Pavesi, Massimo Scotti

Abstract. In this paper, we test the hypothesis that participation in cultural associations improves social capital, and more specifically generalized trust, which in turn fosters environmental concern. Using a dataset combining two large cross-national socioeconomic surveys and an instrumental variable mediation approach, we find that a measure of generalized trust fully mediates the relationship between people's participation in cultural organizations and their environmental concerns. This result suggests that governments can increase citizens' environmental concern – a prerequisite for pro-environmental behaviors – by encouraging their participation in cultural associations.

1. Introduction

The 2005 Faro Convention and the EU Green Deal recognize the importance of cultural heritage and cultural activities in promoting sustainable development and environmental protection. A growing body of research underscores the role of cultural capital, specifically individual engagement in cultural activities, in fostering social capital and civic engagement (Hutter, 1996; Stolle and Rochon, 1998; Sacco et al., 2012; Campagna et al., 2020). Social capital, in turn, has been shown to influence pro-environmental behaviors (Macias and Williams, 2016).

Furthermore, recent studies suggest that cultural consumption can drive specific pro-environmental actions. Crociata *et al.* (2015) found a significant positive relationship between consuming cultural goods and following waste recycling instructions, as well as between social capital and recycling attitudes. Agovino *et al.* (2017) showed that the positive impact of cultural participation on organic food consumption increases with the social orientation of cultural activities. Quaglione *et al.* (2017) found that socially oriented cultural activities lead to reduced electricity consumption, while solipsistic cultural goods consumption increases it.

Despite this growing interest, there is limited understanding of the mechanisms through which participation in cultural activities affects environmental concerns and behaviors. This paper aims to analyze the relationship between cultural capital, social capital, and environmental outcomes. It hypothesizes that the impact of cultural

participation on environmental outcomes is mediated by social capital, specifically generalized trust.[1]

Using two large-scale cross-national repeated surveys, we estimate a model where social capital (namely generalized trust) mediates the relationship between cultural participation and environmental concern. To address potential endogeneity, we use the single instrumental variable (IV) mediation approach proposed by Dippel *et al.* (2020). More precisely, we instrument participation in cultural activities with the level of importance respondents place on cultivating children's imagination. The importance of nurturing children's imagination reflects the concerted cultivation approach: a parenting strategy in which children participate in a wide range of family activities characterized by a wide breadth of cultural activities (Lareau 2003; Doepke and Zilibotti 2017; Wheeler, 2018; Borra and Sevilla 2019). Our findings suggest that this IV is not only relevant but also conditionally exogenous in terms of environmental concerns.

Our results indicate that both the relationship between participation in cultural organizations and generalized trust, as well as the relationship between cultural participation and environmental concern are endogenous. When accounting for endogeneity, a strong mediation effect of generalized trust emerges, fully mediating the impact of cultural participation on environmental concern.

Our study adds to the literature by explicitly investigating the mediating role of social capital in the relationship between cultural capital and environmental concern, disentangling the direct effect from the mediated one.

Our findings suggest that promoting civic engagement in cultural activities enhances environmental concern through the development of generalized trust thereby providing valuable insights for policies aimed at fostering environmental sustainability and social cohesion.

2. Data and Methodology

2.1. Data

Our data derive from a joint project involving two large-scale cross-national survey research programs: the European Value Study (EVS) and the World Value Survey (WVS). These two repeated surveys were initiated in the 1980s to measure human beliefs and values including occupational, social, political, moral, and family values. The EVS surveyed people living in European countries, while the WVS surveyed people living outside of Europe. Five European countries took part in both

[1] Generalized trust can be defined as the trust accorded to individuals we do not know (Bac, 2009). According to the existing literature, generalized trust is one of the main components of social capital (see, e.g., Durlauf, 2002).

Table 1 – *Descriptive Statistics (41,160 obs.)*

Variable	Mean	SD	1st quartile	Median	3rd quartile	Min	Max
Dependent Variable							
Environmental concern	0.634	0.296	0.564	0.564	0.846	0	1.128
Main Explanatory Variables							
Social capital	0.000	0.918	-0.545	-0.039	0.703	-2.320	2.648
Culture	0.247	0.431	0	0	0	0	1
Control Variables							
Age	43.275	16.380	29	41	56	16	82
Female	0.503	0.500	0	1	1	0	1
Religious	1.519	0.670	1	2	2	0	2
Republican leaning	5.675	2.470	4	5	7	1	10
Competition	5.935	2.674	4	6	8	0	9
Demonstrations	0.738	0.729	0	1	1	0	2
Happiness	2.164	0.699	2	2	3	0	3
Health status	2.824	0.847	2	3	3	0	4
Education	3.674	1.997	2	3	6	0	8
Self-reported income	4.918	2.080	4	5	6	1	10
Mills	0.067	0.031	0.045	0.062	0.084	0.006	0.250
Excluded Instrument							
Imagination	0.225	0.418	0	0	0	0	1

surveys: Germany, Romania, Russia, Serbia, and Ukraine. In this study, we utilize the first joint dataset covering the period 2017-2022 (EVS/WVS, 2024).

Table 1 provides the main descriptive statistics for the variables entering our study.[2] Our final sample consists of 41,160 observations with non-missing values in the dependent and independent variables.

We use Item Response Theory (IRT) to construct our dependent variable, the *Environmental Concern*. In particular, *Environmental Concern* is the latent variable obtained from a graded response model (GRM) in which categories are not constrained to be the same across items (Samejima, 1969). Moreover, the strength of the relationship between an item and the measured construct (i.e., the

[2] A description of the variables entering the analysis can be found in the online supplementary material.

discrimination parameter) may differ across items. This variable captures the subject's concerns regarding environmental protection and trust in environmental protection movements.[3]

The graded response model is also employed to create the variable *Social Capital* which is meant to capture generalized trust. Indeed, this variable is constructed using several items indicating whether the respondent trusts most people, family, neighbors, acquaintances, strangers, people of another religion, or people of another nationality.[4]

The variable *Culture* measures cultural capital. Broadly defined, cultural capital includes art and literature, as well as lifestyles, community-building practices, value systems, traditions, and beliefs (UNESCO, 2001). Despite the broad academic and political acceptance of this definition, operational challenges (Romainville, 2015) have led influential empirical research to adopt a more practical approach based on participation in cultural activities. Following this approach, we operationalize cultural capital through involvement in voluntary organizations that promote education, arts, and music. Accordingly, *Culture* is a binary variable, assigned a value of 1 if the respondent is a member of such organizations. Table 1 indicates that about 25 percent of respondents belong to a cultural organization.

Following Stern (2000), we control for socio-demographic characteristics such as age, gender, religiosity, political preferences, education, and self-reported income.

Wright et al. (2003) report a negative association between age and some sustainability concerns. Zelezny *et al*. (2000) show that more educated individuals and women are more likely to exhibit pro-environmental behaviors. Although the empirical literature on the relationship between income and environmental concern is mixed, some studies suggest that high-income individuals are more likely to participate in green electricity programs (Clark et al., 2003). Therefore, we control for self-declared income status. In line with Aldy et al. (2012) and Emiru and Waktola (2018), we consider political preferences and religiosity to account for ideological positions on ecological issues. For the same reason, we include subjects' attitudes towards competition and their propensity to attend lawful demonstrations. We also account for whether the subject felt happy and healthy during the interview, as depressed or sick individuals may be less concerned with social aspects like

[3] The two questions entering the construction of the dependent variable are: i) what is more important: protecting the environment or economic growth? 1=protecting the environment, 2=economic growth, creating jobs, and other answers; ii) how much confidence you have in the Environmental Protection Movement (EPM)? 0=a great deal, 1= quite a lot, 2= not very much, 3= none. The dataset also contained a third question about the subject's participation in environmental organizations. However, we preferred to exclude it because of potential omitted factors influencing the decision to participate in many social groups, including cultural ones.

[4] The coefficients of the graded response models as well as the item information functions associated with *Environmental Concern* and *Social Capital* can be found in the Supplementary Material available online.

environmental quality. Furthermore, to address self-selection issues in answering environmental questions, we computed the Mills' ratio.

Finally, *Imagination* is the variable that we use to instrument participation in cultural activities. This variable measures the level of importance respondents place on cultivating children's imagination. We expect *Imagination* to be highly correlated with respondents' likelihood to participate in cultural activities (relevance of the instrument). Indeed, according to the existing literature, this variable is indicative of the concerted cultivation approach (Lareau 2003, Doepke and Zilibotti 2017; Borra and Sevilla, 2019). This parenting strategy consists of involving children in a broad range of family activities characterized by a wide breadth of cultural activities (Wheeler, 2018). Thus, individuals who place a greater emphasis on children's imagination are also more likely to participate in cultural activities. At the same time, there are no apparent reasons to believe there will be a correlation between beliefs regarding the importance of imagination and concerns about environmental protection (instrument's conditional exogeneity). Nevertheless, we conduct a test and a sensitivity analysis to determine the validity of the exclusion restriction.

2.2. Methodology

To test the hypothesis that social capital mediates the relationship between participation in cultural activities and environmental concern, we first conduct a mediation analysis à la Baron and Kenny (1986), and then we use an instrumental variable (IV) approach to address potential endogeneity issues related to participation in cultural activities. More precisely, we adopt the methodology proposed in Dippel *et al.* (2020) to identify the causal chain between intermediate and final outcomes in a standard IV model without disregarding the endogeneity of cultural participation with respect to social capital and environmental concern.

Based on Baron and Kenny (1986), we separately estimate the following equations:

$$S = \beta_0 + \beta_S^C C + \sum_{i=1}^{n} \delta_i W_i + \varepsilon_S, \quad (1)$$

$$E = \beta_1 + \beta_E^S S + \beta_E^C C + \sum_{i=1}^{n} \rho_i W_i + \varepsilon_E, \quad (2)$$

where C denotes the treatment variable (*Culture*), S is the mediator (*Soc. capital*), E is the final outcome (*Env. concern*), W_i for $i=1,\ldots,n$ is a set of individual socio-demographic characteristics and country fixed effects, and ε are the error terms. Therefore, we can estimate the indirect effect of *Culture* on *Env. concern* by simply multiplying β_S^C with β_E^S, whereas the direct impact corresponds to coefficient β_E^C. Notice that, if we exclude the mediator (S) from Equation (2), the coefficient of C will represent the total impact of *Culture* on *Env. concern*.

However, traditional mediation analysis does not consider the possibility that both treatment and mediator may be endogenous. We must therefore change our identification strategy to address this important issue and claim a causal mediation effect of social capital. Dippel *et al.* (2020) show that a single instrumental variable suffices to identify both the direct and the indirect effects when the same unobserved confounders jointly influence the relationship between *C* and *S* and the relationship between *C* and *E*. This is the only identifying assumption that is needed to apply their methodology.

Keeping the linearity assumption and given the availability of an instrument Ω, we can estimate a causal mediation model by using the following procedure:

1) Considering that *C* is a binary variable, parameter β_S^C can be identified through the following system:

$$Z = \Phi(b\Omega + W\omega), \tag{3}$$
$$C = \alpha_0 + \alpha_1 \mathrm{Z} + \sum_{i=1}^{n} \delta_i W_i + \varepsilon_C, \tag{4}$$
$$S = \beta_0 + \beta_S^{\hat{C}} \hat{C} + \sum_{i=1}^{n} \rho_i W_i + \varepsilon_S, \tag{5}$$

where $Z \equiv \Pr(C = 1|W, \Omega)$ and $\Phi(.)$ is the cumulative distribution function of the standard normal distribution.[5]

2) We then exploit the fact that the identifying assumption implies a new exclusion restriction (Dippel *et al.*, 2020), using *Z* as an IV for *S* when conditioned on *C*. Formally, we compute parameter β_E^S as follows:

$$S = \beta_0 + \gamma_S^C C + \gamma_S^Z Z + \sum_{i=1}^{n} \delta_i W_i + \epsilon_S, \tag{6}$$
$$E = \beta_1 + \beta_E^S \hat{S} + \beta_E^C C + \sum_{i=1}^{n} \rho_i W_i + \varepsilon_E. \tag{7}$$

We test the exclusion of *Z* in Equations (4)-(5) and (6)-(7) by using the methodology described in D'Haultfoeuille *et al.* (2021).[6]

By denoting with *X* the potentially endogenous variable and with *Y* the outcome, we can write the exclusion restriction as $Y = g(X, Z, \varepsilon) = g(X, \varepsilon)$.[7] D'Haultfoeuille *et al.* (2021) demonstrate that if there exists a point (x^*, z, z') such that $F_{X|Z}(x^*|z) = F_{X|Z}(x^*|z') \in (0, 1)$, where $F_{X|Z}$ is the conditional distribution of X, then we can

[5] Following Wooldridge (2010), given the binary nature of our cultural variable, we first estimated a Probit model for Culture, using as excluded instrument the dummy variable indicating whether respondents consider imagination an important dimension for children's education, and then we used the predicted treatment probability as instrument in a linear two-stage least square estimator. As a robustness check, we repeat the analysis by considering the two variables entering our environmental score separately. Results can be found in the online supplementary material.

[6] Because in Equation (7) it is crucial to control for *Culture*, we first partial out the effects of our controls and then apply the test procedure proposed by D'Haultfoeuille *et al*. (2021).

[7] Notice that D'Haultfoeuille *et al.* (2021) keep the outcome function as general as possible, since their test relies on a control function method.

test the following exclusion restriction: $Y(x^*, z, \varepsilon) = Y(x^*, z', \varepsilon)$. The null hypothesis can be written as:

$$H_0: F_{Y|X,Z}(\cdot\,|x^*, 0) = F_{Y|X,Z}(\cdot\,|x^*, 1), \tag{8}$$

where the conditional distribution of Y (i.e., $F_{Y|X,Z}$) is estimated using a kernel function approach and a bandwidth parameter. D'Haultfoeuille *et al.* (2021) prove that the Kolmogorov–Smirnov statistic is a powerful test against a fixed alternative hypothesis of the form: $F_{Y|X,Z}(y|x^*, 0) \neq F_{Y|X,Z}(y|x^*, 1)$ for some y.

Finally, we also conduct a sensitivity analysis aiming to relax the assumption that our instrument is correctly excluded from the mediation equation (5) and the outcome equation (7) (i.e., $Z \perp \varepsilon_C, \varepsilon_S$ and $Z \perp \epsilon_S, \varepsilon_E$). More precisely, following Conley *et al.* (2012), we construct a confidence region for any possible correlation level between the excluded instrument and the error terms. This allows us to assess the robustness of our results with respect to violations of the exclusion restrictions.[8]

3. Results

Table 2 shows the estimates of Equations (1) and (2) with and without the inclusion of social capital. According to Baron and Kenny (1986), the coefficient of *Culture* reported in Column 1 represents the total effect of participating in cultural activities on the environmental concern. This effect explains approximately 10 percent of the standard deviation of the environmental concern. Columns 2 and 3 allow us to distinguish between the direct and indirect effects of cultural activities on the environmental concern. Column 3 provides the direct effect of Culture on Env. concern, which is about 86 percent of the total effect (i.e., 0.032 out of 0.037). Therefore, the mediation effect only accounts for a small part of the total effect. Indeed, we can compute the impact of Culture on Env. concern passing through Soc. capital by multiplying the coefficient of Culture in Column 2 with that of Soc. capital in Column 3.

According to our findings, those who participate in cultural activities are more likely to trust other people. At the same time, those who have a higher level of trust in other people are also more concerned about the environment. This supports the hypothesis that, since environmental quality is a common good that requires coordination to be maintained and participation in cultural activities facilitates the development of generalized trust, individuals who participate in cultural activities may be more prone to concern about the environment if they feel that coordination is easier because they can trust others.

[8] The results of both the test and the sensitivity analysis for the exclusion restriction are reported in the online supplementary material.

Table 2 – *Environmental Concern (41,160 obs.). This table reports the linear regression coefficients of our mediation model, including country fixed effects.*

	Env. concern	Soc. capital	Env. concern
	(1)	(2)	(3)
Culture	0.033***	0.091***	0.027***
	(0.006)	(0.025)	(0.006)
Soc. capital			0.064***
			(0.004)
Age	-0.001***	0.005***	-0.001***
	(0.000)	(0.001)	(0.000)
Female	0.022**	-0.035*	0.024***
	(0.009)	(0.020)	(0.008)
Religious	0.006	0.053***	0.002
	(0.007)	(0.019)	(0.007)
Republican leaning	-0.011**	-0.015**	-0.010*
	(0.005)	(0.006)	(0.005)
Competition	0.002	0.011***	0.001
	(0.001)	(0.004)	(0.001)
Demonstrations	0.049***	0.111***	0.042***
	(0.009)	(0.017)	(0.008)
Happiness	0.030***	0.099***	0.024***
	(0.004)	(0.014)	(0.004)
Health	0.008*	0.058***	0.004
	(0.004)	(0.011)	(0.005)
Education	-0.001	0.020***	-0.003
	(0.002)	(0.007)	(0.002)
Self-reported income	-0.002	0.013**	-0.003
	(0.002)	(0.006)	(0.002)
Mills	-0.370	-1.149	-0.297
	(0.334)	(0.870)	(0.342)
Constant	0.611***	-0.808***	0.663***
	(0.057)	(0.111)	(0.062)
Country FE	Yes	Yes	Yes
R-sq (within)	0.036	0.046	0.067
R-sq (between)	0.103	0.293	0.107

Notes: Robust standard errors in parentheses. Significance: *p<10%, **p<5%, ***p<1%.

Table 3 – *Causal mediation analysis for linear IV models (41,160 obs.). This table reports the estimates of Equations (5) and (7), as well as the computation of the total, direct, and indirect effects of culture on the environmental concern.*

A. Regression Coefficients		
	Social Capital	Env. Concern
Culture (C)	1.992***	0.007
	(0.300)	(0.006)
Social Capital (S)		0.285***
		(0.047)
B. Effects of culture on the environmental concern		
Total		0.574***
		(0.097)
Direct		0.007
		(0.006)
Indirect		0.567***
		(0.126)
Additional controls	Yes	Yes
Country FE	Yes	Yes
First stage F-statistic (C on Z) =81.161		
First stage F-statistic (S on Z|C) =80.060		

Notes: Standard errors in parentheses. Significance: *p<10%, **p<5%, ***p<1%.

In Table 3, we re-estimated our mediation model controlling for endogeneity. The estimates of Equations (5) and (7) are presented in Panel A, whereas the corresponding direct effect, indirect (i.e., mediated) effect and total effect of Culture on the Environmental concern are presented in Panel B The direct effect is statistically insignificant and numerically negligible, whereas the indirect effect is positive and statistically significant. Thus, once we account for the possibility of omitted variables confounding the relationship between Culture and (intermediate and final) outcomes, the positive relationship between Culture and the Environmental concern is fully mediated by Social capital. Multiplying $\beta_S^{\hat{C}} = 1.992$ with $\beta_E^S = 0.285$ gives us the indirect effect.

The reason why IV coefficients (in Table 3) are larger than the OLS coefficients (in Table 2) is that the former capture a local average treatment effect (LATE). Specifically, whereas the OLS refers to the average impact of a unitary increase in cultural participation on environmental concern across the entire population, the IV

captures the average impact of a unitary increase in cultural participation on environmental concern for only those who consider imagination important to children. Therefore, the IV estimates are larger given the heterogeneous population we are considering. In other words, the full mediation result only applies to individuals who indicated that imagination is an important quality for children. Since they represent 22.54 percent of the sample, a partial mediation effect can be observed among individuals who do not view imagination as being so important but who participate in cultural activities.

More importantly, the IV estimates demonstrate a causal relationship between cultural participation and environmental concern passing through social capital.

4. Conclusions

This paper examines the relationship between participation in voluntary cultural associations and environmental concerns, focusing on whether generalized trust, a key component of social capital (Putnam, 1995), mediates this relationship. While a traditional mediation model found only a small partial mediation effect, an IV approach revealed that social capital fully mediates the impact of cultural participation on environmental concerns. Thus, public policies promoting cultural activities can enhance environmental concern and other socio-economic dimensions related to social capital.

The study has two key limitations for future research to address. First, it lacks variables capturing environmental behaviors like energy conservation or recycling. According to Bamberg (2003), an environmental concern resulting from a situation-specific cognition may influence certain pro-environmental behaviors without affecting others. Second, the dataset does not account for informal and unstructured cultural activities that contribute to cultural capital. Future research could also use social reforms promoting cultural activities to further explore this mediation analysis.

References

AGOVINO M., CROCIATA A., QUAGLIONE D., SACCO P., SARRA A. 2017. Good taste tastes good. Cultural capital as a determinant of organic food purchase by Italian consumers: Evidence and policy implications. *Ecological Economics*, 141, 66-75.

ALDY J. E., KOTCHEN M. J., LEISEROWITZ A. A. 2012. Willingness to pay and political support for a US national clean energy standard. *Nature Climate Change*, Vol. 2, No.8, pp. 596-599.

BAC M. 2009. Generalized trust and wealth. *International Review of Law and Economics*, Vol. 29, No.1, pp. 46-56.

BAMBERG, S. 2003. How does environmental concern influence specific environmentally related behaviors? A new answer to an old question. *Journal of environmental psychology*, Vol. 23, No. 1, pp. 21-32.

BARON R. M., KENNY, D. A. 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, Vol. 51, No. 6, pp. 1173-1182.

BORRA C., SEVILLA A. 2019. Competition for university places and parental time investments: Evidence from the United Kingdom. *Economic Inquiry,* Vol. 57, No. 3, pp. 1460-1479.

CAMPAGNA D., CAPERNA G., MONTALTO V. 2020. Does culture make a better citizen? Exploring the relationship between cultural and civic participation in Italy. *Social Indicators Research*, Vol. 149, No. 2, pp. 657-686.

CLARK C. F., KOTCHEN M. J., MOORE M. R. 2003. Internal and external influences on pro-environmental behavior: Participation in a green electricity program. *Journal of environmental psychology*, Vol. 23, No. 3, pp. 237-246.

CONLEY T. G., HANSEN C. B., ROSSI P. E. 2012. Plausibly Exogenous. *The Review of Economics and Statistics*, Vol. 94, No. 1, pp. 260-272.

CROCIATA A., AGOVINO M., SACCO P. L. 2015. Recycling waste: Does culture matter?. *Journal of Behavioral and Experimental Economics*, 55, pp. 40-47.

D'HAULTFŒUILLE X., HODERLEIN S., SASAKI Y. 2021. Testing and relaxing the exclusion restriction in the control function approach. *Journal of Econometrics*, Vol. 24, No. 2.

DIPPEL C., FERRARA A., HEBLICH S. 2020. Causal mediation analysis in instrumental-variables regressions. *The Stata Journal*, Vol. 20, No.3, pp. 613-626.

DOEPKE M., ZILIBOTTI M. 2017. Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission. *Econometrica*, Vol. 85, No. 5, pp. 1331–1371.

DURLAUF S. N. 2002. On the empirics of social capital. *The Economic Journal*, Vol. 112, No. 483, pp. 459-479.

EMIRU T. S., WAKTOLA D. K. 2018. The environmental awareness of higher education students and the implications for the Paris Climate Agreement: empirical evidences from Ethiopia and USA. *International Research in Geographical and Environmental Education*, Vol. 27, No. 3, pp. 216-233.

EVS/WVS (2024). Joint EVS/WVS 2017-2022 Dataset (Joint EVS/WVS). *GESIS, Cologne. ZA7505 Data file Version 5.0.0, https://doi.org/10.4232/1.14320.*

HUTTER M., 1996. The impact of cultural economics on economic theory. *Journal of Cultural Economics*, 20, pp. 263–268.

LAREAU A. 2003. *Unequal Childhoods*. Berkeley: University of California Press.

MACIAS T., WILLIAMS K. 2016. Know your neighbors, save the planet: Social capital and the widening wedge of pro-environmental outcomes. *Environment and Behavior*, Vol. 48, No. 3, pp. 391-420.

PUTNAM R. D. 1995. Tuning in, tuning out: The strange disappearance of social capital in America. *PS: Political science & politics*, Vol. 28, No. 4, pp. 664-683.

QUAGLIONE D., CASSETTA E., CROCIATA A., SARRA A. 2017. Exploring additional determinants of energy-saving behaviour: The influence of individuals' participation in cultural activities. *Energy Policy*, 108, pp. 503-511.

SACCO P.L, FERILLI G., TAVANO BLESSI G. 2012. Culture 3.0: A new perspective for the EU active citizenship and social and economic cohesion policy. In: EUROPEAN HOUSE FOR CULTURE (Ed.) *The cultural component of citizenship: An inventory of challenge*, Access to the culture platform pp.195–213.

SAMEJIMA, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika monograph supplement*, Vol. 34, No. 4, Pt. 2, 100.

STOLLE D., ROCHON T. R. 1998. Are all associations alike? Member diversity, associational type, and the creation of social capital. *American Behavioral Scientist,* Vol. 42, No.1, pp. 47-65.

UNESCO 2001. UNESCO Universal Declaration on Cultural Diversity. Paris: UNESCO.

WHEELER S. 2018. 'Essential assistance' versus 'concerted cultivation': theorising class-based patterns of parenting in Britain. *Pedagogy, Culture & Society*, 26.3, pp. 327-344.

WOOLDRIDGE J. M. 2010. *Econometric analysis of cross section and panel data*. The MIT press.

WRIGHT S. D., CASERTA M., LUND D. A. 2003. Older adults' attitudes, concerns, and support for environmental issues in the "new West". *The International Journal of Aging and Human Development*, Vol. 57, No. 2, pp. 151-179.

ZELEZNY L. C., CHUA P. P., ALDRICH C. 2000. Elaborating on gender differences in environmentalism. *Journal of Social issues*, Vol. 56, No. 3, pp. 443-458.

Thomas BASSETTI, Department of Economics and Management, University of Padua, thomas.bassetti@unipd.it

Filippo PAVESI, School of Economics and Management, LIUC (Carlo Cattaneo University), Stevens Institute of Technology, School of Business, Hoboken, New Jersey, USA, fpavesi@liuc.it;

Massimo SCOTTI, School of Economics and Management, LIUC (Carlo Cattaneo University), mscotti@liuc.it

TECHNO-ECONOMIC IMPACTS OF ENERGY TRANSITION: THE STATE OF THE ART

Marina Albanese, Monica Varlese

Abstract. Achieving net-zero carbon emissions goals requires tangible actions that regions, countries, sectors, and organizations should take to accelerate the energy transition along with a growing trend toward renewable energy sources. Digitalization, a phenomenon introduced by the transformative power of digital technologies, plays a key role in the energy transition process. Providing advancements in technology leads to significant changes in the way energy is produced, transmitted, and consumed. In light of that, the paper discusses the impacts of digitalization on energy transition by highlighting its benefits and crucial incentives favoring investment processes. In terms of policy messages, the paper suggests that policies based on a more integrated assessment of the links between the transforming energy sector and the overall economy would be needed. Moreover, regulations promoting financial incentives to support investments would be recommended to satisfy the novel paradigm of energy transition.

4 Introduction

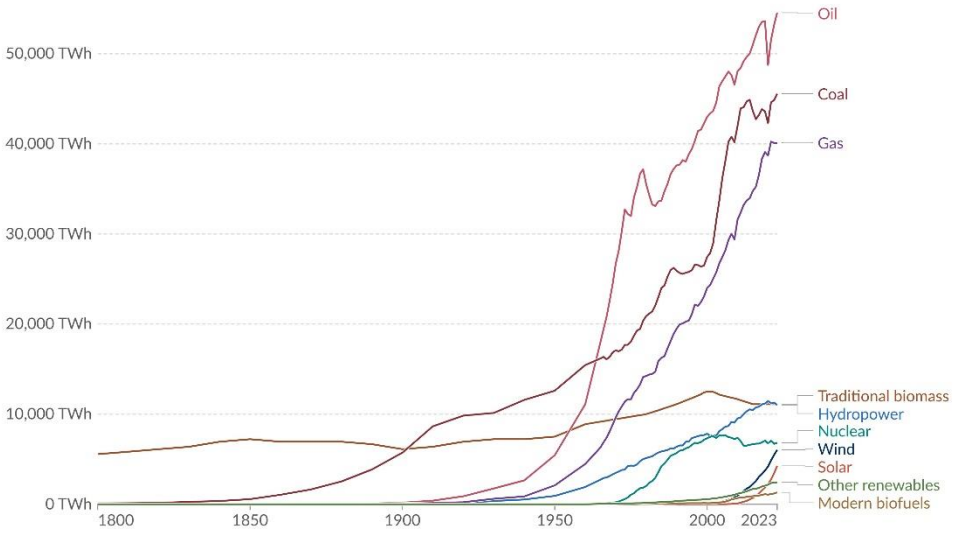
The threats posed by climate change and the need to achieve global sustainable development goals (SDGs) necessitate a fundamental transformation of the global energy system. The SDGs report highlights the risks of climate change, including its impact on inequality and food security, which may undermine progress toward other SDGs. Given that energy use is the primary driver of global greenhouse gas emissions, transitioning to a sustainable global energy system is critical (Bogdanov et al., 2021).

Up until the mid-19th century, traditional biomass—such as wood, crop residues, and charcoal—was the dominant global energy source. The Industrial Revolution led to a major change, with coal becoming the leading energy source, followed by oil, gas, and hydropower by the early 1900s. Nuclear energy only became part of the mix in the 1960s, while "modern renewables" like solar and wind began to play a significant role in the 1980s (Figure 1).

Today, renewable energy sources are growing rapidly, with one-sixth of the world's primary energy coming from low-carbon sources like nuclear, hydropower, wind, solar, bioenergy, geothermal, and tidal power. However, oil remains the largest

global energy source, followed by coal, gas, and hydropower. Fossil fuel combustion accounts for about 75% of global greenhouse gas emissions, highlighting the need for a shift to low-carbon sources to achieve decarbonization. This transition is essential for transforming the global energy sector from fossil fuels to zero carbon emissions by the second half of the 21st century (Ritchie and Rosado, 2020).

Figure 1 – *Global primary energy consumption by source.*



Source: Our World in data available at Global primary energy consumption by source (ourworldindata.org).

The transition towards a more efficient energy system implies access to higher shares of renewable energy (RE) by reducing GHG emissions and water scarcity.

Such process has already started with renewables providing more than 27% of the global electricity generation by the end of 2019, including about 11% generated by new renewable energy technologies, mainly wind turbines and solar photovoltaics (PV) (Bogdanov et al., 2021).

The future transition of energy systems largely depends on a set of assumptions, methods, and objectives that should be extensively applied from a global perspective. In light of that, studying energy systems from a multi-level perspective appears relevant to understanding the implications that the transition path might have. Especially the link between technological innovation and social aspects represents a key drive to ensuring the sustainability of energy systems (Hansen et al, 2019).

In that context, digital transformation reveals a top priority for both businesses and society.

Digital transformation is a phenomenon regarding the transformative power of digital technologies which have become integral to our daily lives, influencing households' work, communication, and behaviour as consumers. Specifically, digital technologies by connecting the physical and digital worlds, can satisfy individualized needs of consumers. Historically, digitalization has been helping to improve the energy sector. Indeed, the energy system has long been at the forefront of embracing novel technologies to produce assets, including reservoirs and pipelines, and, to improve quality and yields by minimizing energy use (Nazari and Musilek, 2023). The main goal of this technologically-driven dimension of the energy transition is the creation of Smart Energy Systems (SES, hereafter), i.e., providing inter-linkages between energy sectors to improve the efficiency and sustainability of future energy systems (Cantarero 2020). The rapid digital transformation in the energy sector, including smart grids and the energy internet, promotes a path toward more resiliency and flexibility in sustainable energy systems. Indeed digitalization addresses multi-energy connectivity and complementarity across regions through instant data collection, analysis, and sharing (Zhang et al, 2023).

However, the energy system's technological transformation requires consumer awareness, commitment, and active participation to ensure a fair and equitable distribution of its benefits (Cantarero 2020).

In light of the significance of digital transformation in shaping the future, it is relevant to study the current status and impacts of this trend on energy sector and anticipate future developments and their potential consequences.

This paper first reviews the literature on the economic growth-energy transition nexus to investigate the relationship between economic growth and renewable energy consumption. Second, it studies the impacts of digital transformation on the energy sector through a review of existing studies focusing on the technology-energy transition nexus. Moreover, it offers an overview of digital transformation at the EU countries level by assessing the impact of digital trends both on the demand and on the supply side.

This study emphasizes the importance of coordinating energy storage and digitalization in the energy transition, highlighting the need for global attention to the relationship between energy and society. Policies should integrate the energy sector with the broader economy, promoting innovation in distribution systems and offering financial incentives to support smart grid investments in line with the energy transition.

The remainder of the paper is organized as follows. Section 2 provides an overview of existing studies focusing on energy transition and its digital transformation. Section 3 discusses the main implications of digitalization in the

energy sector both on the demand and the supply side. Section 5 presents the conclusions and policy implications.

5 Literature review

A growing body of studies focuses on the energy transition economic growth nexus. Khan et al. (2021) conducted an econometric analysis across thirty-eight IEA countries and found a positive relationship between economic growth and renewable energy consumption. They highlighted that the shift from non-renewable to renewable energy has significant economic externalities, with energy use acting as both a cause and facilitator of growth. Indeed, renewable power generation technologies accelerate new job creation. In this regard, Ram et al. (2020) estimate that global direct employment in the electricity sector raises from about 21 million in 2015 to nearly 35 million in 2050. Tzeremes et al. (2023) observed that higher economic activity in BRICS countries leads to increased carbon emissions, driving environmental degradation and emphasizing the need for a renewable energy transition. Economic growth, therefore, facilitates renewable energy consumption and accelerates the transition process.

Another strand of the literature, instead, reveals a negative relationship between economic growth and renewable energy. Tsagkari et al. (2021) analyze how renewable energy can support both economic growth and degrowth through local energy projects on islands, using case studies from El Hierro (Spain) and Tilos (Greece). They suggest that these projects' goals—such as energy democracy, self-sufficiency, and revitalization of local economies—align with degrowth principles. The study demonstrates that transitioning to renewable energy can enhance local control over resources and energy independence, leading to either economic growth or degrowth depending on community priorities. Similarly, Muazu et al. (2022), using a panel threshold regression model in 54 African countries, find a non-linear, negative relationship between renewable energy consumption and economic growth, with the impact varying at different stages of development. Bhattacharya et al. (2016) also report that while renewable energy generally promotes long-term growth, it may hinder short-term growth in countries with underdeveloped infrastructure or where investments in renewables divert resources from other sectors. Indeed, the energy transition process calls for higher and more diversified investment both by states and private investors. In this regard, Dong et al. (2022) highlight that renewable energy faces structural challenges in promoting sustainable development due to high capital and technology costs. They argue that financial incentives can improve factor allocation efficiency and enhance renewable energy technologies, positively impacting carbon emission efficiency. Tian et al. (2022) suggest that green stimulus plans are crucial for economic recovery and energy

transition, playing a key role in restarting renewable energy development. To address financing gaps, Qadir et al. (2021) propose involving financial institutions to support renewable energy investments through soft loans or crowdfunding platforms. Bayulgen (2020) examines the political and socio-economic factors influencing US local governments' energy transition policies, emphasizing that citizen concerns about the unpredictability of local reforms highlight the importance of a well-coordinated policy champion configuration for successful clean energy implementation.

In light of that, capacity for successful of energy transition depends from several aspects of a nation such as energy mix, potential for diffusing renewable energy, infrastructure, technological innovation and capacity to transform; as well as societal values and political ambition. Deshmukh et al. (2023) stress the need for international collaboration to address the unequal distribution of pollution's impacts, with countries reliant on fossil fuels, particularly coal, facing higher pollution and health consequences, while those shifting to renewable energy generally experience lower pollution levels. Neofytou et al. (2020) find that Sweden has the most favorable conditions for energy transition, followed by Western Europe and Canada. Developing countries, which are more dependent on fossil fuels, play a critical role in achieving a low-carbon economy (Cantarero, 2020; Murshed et al., 2021). Nam and Jin (2021) advocate for energy efficiency regulations as an effective and low-cost tool for carbon mitigation.

Starting from the key role that innovation technology plays in the sustainable energy process, another strand of literature studies the link between digital technology and sustainable energy development. Wang et al. (2022) develop a measurement index system for digital technology levels and use a VAR model to demonstrate its role in reducing energy-related pollution in China. Similarly, Tzeremes et al. (2023) find that innovation and communication technology significantly influence energy transition in BRICS countries. Green technology innovation helps lower pollution from fossil energy, boosts clean energy use, and increases productivity, while also controlling waste emissions in production processes (Du et al., 2021). Renewable energy technologies not only mitigate environmental impacts but also generate new wealth and jobs (Ram et al., 2020).

However, the effects of green technology on carbon productivity in less developed countries remain unclear. Du and Li (2019) show that technological progress enhances carbon productivity in high-income countries, but its impact in other economies is limited due to high costs. Similarly, Shahbaz et al. (2022) find that technology innovations contribute to the energy transition in high-income countries, but their impact varies regionally.

The global trend towards digitalization has created substantial demand for advanced technologies in energy storage and other emerging sectors. Digitalization,

by promoting technological innovation, supports the transition to a low-carbon economy, especially in energy storage (Zhang et al., 2023). Look (2020) notes that digitalization fosters sustainable energy transitions by driving business model innovation. Bergman and Foxon (2023) show that digitalization influences energy demand, with its effects on energy consumption being shaped by individual and collective choices. Policymakers should focus on social well-being and environmental protection when promoting digitalization. Finally, Hansen et al. (2019) emphasize that the interaction of social and technical elements in models can vary significantly.

6 Digitalization of energy system

Digital technology plays a crucial role in the energy transition by driving decentralization and increasing connections between devices, producers, distributors, and users. This supports investments in renewable and distributed energy resources such as solar photovoltaics, energy storage, and electric mobility. The European Green Deal and the Digital Decade Policy Program 2030 outline goals for digitalizing the EU energy system, including installing solar panels on all commercial and public buildings by 2027, deploying 10 million heat pumps, and replacing 30 million fossil fuel-powered cars with zero-emission vehicles by 2030. Achieving these targets requires the adoption of digital technologies like smart IoT devices and 5G/6G connectivity (EU Commission, 2023).

The term “Digitalization” refers to the increasing use of Information and Communication Technology (ICT) across various sectors, particularly in energy. It includes innovations like smartphones, mobile internet, GPS for location tracking, and real-time traffic data. Digitalization has been enhancing the energy sector for decades by integrating IT systems to manage assets such as reservoirs and pipelines, while also improving quality and efficiency by reducing energy consumption (IEA, 2017).

As for energy demand, digital technologies play a predominant role for sectors of transport, buildings and industry. For transport, connectivity allows entailing new mobility-sharing services by making the transport system more intelligent, efficient, reliable, and sustainable. Advances in vehicle automation and electrification, driven by digitalization, could transform transportation, with uncertain effects on energy use and emissions (Noussan and Tagliapietra, 2020). In the building sector, digital technologies can help reduce energy use by around 10% by 2040, especially in heating and cooling through smart thermostats and sensors (Asif et al., 2024). Digitalization also offers potential energy savings in the industrial sector, improving

process evaluations and leading to significant energy reductions with short payback periods (Erdogan, 2021).

On the supply side, digital technologies have been used for years to enhance fossil resource recovery, improve production processes, reduce costs, and increase safety. The widespread adoption of digital technologies could lower oil and gas production costs by 10% to 20% (IEA, 2017). Additionally, it could improve geological modeling and optimize automation in the coal sector. In the power sector, digitalization may result in annual savings of approximately USD 80 billion by reducing operation and maintenance costs, thus minimizing unplanned outages and downtime (IEA, 2017).

Table 1 describes the main contributions of the most important digital technologies in the EU energy sector. Based on the literature, Światowiec-Szczepańska and Stępień (2022) highlight that digitalization enhances the energy sector by improving system stability, environmental protection, energy demand reduction, revenue, cost efficiency, and customer satisfaction. These benefits are driven by digital applications such as smart grids, optimized procedures, flexible systems, anomaly detection, and improved process efficiency, trust, and transparency.

Table 1 – *Advantages of digital transformation in the energy sector.*

Main advantages of digital technologies	1. System protection and stability	2.Environmental safeguard	3.Minimized cost	4.Customer satisfaction
Applications in energy sectors	- Smart grid and optimized operations - Smart market and flexibility combination - Anomaly detection and prediction - Trust and transparency	- Smart grid and optimized operations - Smart market and flexibility combination - Anomaly detection and prediction -Process efficiency	- Smart grid and optimized operations - Smart market and flexibility combination - Smart home	- Smart home - Trust and transparency
Types of digital technologies	Blockchain, Artificial neural networks (ANN), Internet of Things (IoT), Artificial intelligence (AI), Big data, among many others.			

[1]*Source: Our elaboration from Światowiec-Szczepańska and Stępień (2022)*

As well, digitalization through the increasing application of information- and communication technology (ICT), also impacts energy efficiency. Specifically, ICT technologies can help reduce energy demand through two main factors:

2. Energy efficiency: ICT technologies optimize processes and improve efficiency, reducing energy consumption.
3. Sectoral change: ICT can shift economic dynamics, reducing energy use in certain sectors or introducing more energy-efficient business models.

In this regard, using analytical methodology, Lange et al. (2020) find that digitalization enhances energy efficiency and reduces consumption, although the extent of these reductions is uncertain due to "rebound effects." These effects occur when efficiency improvements lead to increased energy use, either through re-spending savings or substituting other production factors with energy, limiting overall reductions in demand. Estimates of rebound effects vary, but they are generally significant enough to prevent substantial energy savings. Similarly, Xu et al. (2022) show that in high-income countries, digital transformation has driven innovation and system changes, resulting in smarter energy distribution and lower consumption. They also highlight how digitalization fosters low-carbon transitions in other regions through technology spillovers, promoting reduced energy intensity and improved efficiency. Overall, digitalization entails three main cross-cutting threats to the energy system including cybersecurity, privacy, and economic turmoil. Specifically, the digitalization of energy systems can make them more vulnerable to digital troubles, namely geomagnetic storms, and cyber-risks. Moreover, privacy and data ownership represent a major concern since more detailed information is collected, especially data about household energy use. Eventually, digitalization is influencing jobs and skill requirements in the energy sector by changing job patterns and tasks, causing work losses in some areas and creating new opportunities in others (IEA, 2017).

4 Conclusions

The threats of climate change require a fundamental transformation of the global energy system since energy use is the major responsible for global greenhouse gas (GHG) emissions. The rapid digital transformation in the energy sector, including smart grids and the energy internet, promotes a path toward more resiliency and flexibility in sustainable energy systems. Indeed, there is now widespread awareness that to achieve national and European objectives it is essential to increasingly integrate non-renewable sources programmable - a key element of the energy transition - safely and reliably in electrical systems, modernizing and strengthening them with the support of information and communication technologies (Valenti and Graditi, 2020).

In light of that, assessing the impacts of digital transformation on the energy sector appears relevant. Accordingly, the study investigates the role of technology innovation in the energy market by studying the techno-economic impacts of the digital trend on energy sectors.

Given the relevant impacts that digitalization in energy transition entails throughout the economy, the paper suggests that such transformation cannot be considered far away from the broader socio-economic context. In light of that, policies based on a more integrated assessment of the links between the transforming energy sector and the overall economy would be needed. Finally, regulation should promote innovation in the distribution system and provide financial incentives to support investments in technology innovation projects and satisfy the novel paradigm of energy transition. In terms of future perspective, an integrated and multi-carrier energy networks based on digital logic, advanced and more complex that will allow integrated and coordinated management of energy needs (thermal and electrical) will be needed. This, in turn, will favour an even more effective exploitation of generation from renewable sources and ensure adequate levels of resilience and flexibility in the energy system by facilitating the achievement of the targets set to pursue energy transition at the global level.

Acknowledgments

This work is supported by the "Network 4 Energy Sustainable Transition - NEST" project (MUR project code PE000021, Concession Decree No. 1561 of 11.10.2022), in the framework of the NextGenerationEu PNRR plan (CUP E63C22002160007).

We would like to thank the participants of the SIEDS 2024 conference held in Varese, as well as the participants of the COBENARE and IAFEP 2024 conferences held in Naples, for their valuable comments and suggestions, which have been useful in improving this work.

References

ASIF M., NAEEM G., KHALID M. 2024. Digitalization for sustainable buildings: Technologies, applications, potential, and challenges. *Journal of Cleaner Production*, 141814.

BHATTACHARYA M., PARAMATI S. R., OZTURK I., BHATTACHARYA S. 2016. The effect of renewable energy consumption on economic growth: Evidence from top 38 countries. Applied energy, Vol. 162, pp. 733-741.

BAYULGEN O. 2020. Localizing the energy transition: Town-level political and socio-economic drivers of clean energy in the United States, *Energy Research & Social Science*, Vol. 62, 101376.

BERGMAN, N., FOXON, T. J. 2023. Drivers and effects of digitalization on energy demand in low-carbon scenarios, *Climate Policy*, Vol. 23, No. 3, pp. 329-342.

BOGDANOV D., RAM M., AGHAHOSSEINI A., GULAGI A., OYEWO A. S., CHILD M., …, BREYER C. 2021. Low-cost renewable electricity as the key driver of the global energy transition towards sustainability, *Energy*, Vol. 227, 120467.

CANTARERO M. M. V. 2020. Of renewable energy, energy democracy, and sustainable development: A roadmap to accelerate the energy transition in developing countries, *Energy Research & Social Science*, Vol. 70, 101716.

DESHMUKH M. K. G., SAMEERODDIN M., ABDUL D., SATTAR M. A. 2023. Renewable energy in the 21st century: A review. *Materials Today: Proceedings,* No. 80, pp. 1756–1759.

DONG F., LI Y., GAO Y., ZHU J., QIN C., ZHANG X. 2022. Energy transition and carbon neutrality: Exploring the non-linear impact of renewable energy development on carbon emission efficiency in developed countries, *Resources, Conservation and Recycling*, Vol. 177, 106002.

DU K., CHENG Y., YAO X. 2021. Environmental regulation, green technology innovation, and industrial structure upgrading: The road to the green transformation of Chinese cities, *Energy Economics,* Vol. 98, 105247.

DU K., LI J. 2019. Towards a green world: How do green technology innovations affect total-factor carbon productivity, *Energy Policy*, Vol. 131, pp. 240-250.

ERDOGAN S. 2021. Dynamic nexus between technological innovation and building sector carbon emissions in the BRICS countries, *Journal of Environmental Management*, Vol. 293, 112780.

EUROPEAN COMMISSION, 2023. Digitalization of the European Energy System, available at https://digital-strategy.ec.europa.eu/en/policies/digitalisation-energy

HANSEN P., LIU X., MORRISON G. M. 2019. Agent-based modelling and socio-technical energy transitions: A systematic literature review, *Energy Research & Social Science*, Vol. 49, pp. 41-52.

IEA 2017, Digitalisation and Energy, *IEA*, Paris https://www.iea.org/reports/digitalisation-and-energy, License: CC BY 4.0

KHAN I., HOU F., ZAKARI A., TAWIAH V. K. 2021. The dynamic links among energy transitions, energy consumption, and sustainable economic growth: A novel framework for IEA countries, *Energy*, Vol. 222, 119935.

LANGE S., POHL J., SANTARIUS T. 2020. Digitalization and energy consumption. Does ICT reduce energy demand?. *Ecological economics*, Vo. 176, 106760.

LOOCK M. 2020. Unlocking the value of digitalization for the European energy transition: A typology of innovative business models, *Energy research & social science,* Vol. 69, 101740.

MUAZU A, YU Q, LIU Q. 2023. Does renewable energy consumption promote economic growth? an em- pirical analysis of panel threshold based on 54 african countries. International Journal of Energy Sector Management, No. 17, pp. 106–127.

MURSHED M., AHMED Z., ALAM M. S., MAHMOOD H., REHMAN A., DAGAR V. 2021. Reinvigorating the role of clean energy transition for achieving a low-carbon economy: evidence from Bangladesh, *Environmental Science and Pollution Research*, Vol. 28, pp: 67689-67710.

NAM E., JIN T. 2021. Mitigating carbon emissions by energy transition, energy efficiency, and electrification: Difference between regulation indicators and empirical data, *Journal of Cleaner Production*, Vol. 300, 126962.

NOUSSAN M., TAGLIAPIETRA S. 2020. The effect of digitalization in the energy consumption of passenger transport: An analysis of future scenarios for Europe, *Journal of Cleaner Production,* Vol. 258, 120926.

NAZARI Z., MUSILEK P. 2023. Impact of Digital Transformation on the Energy Sector: A Review, *Algorithms*, Vol. 16, No. 4, p. 211.

NEOFYTOU H., NIKAS A., DOUKAS H. 2020. Sustainable energy transition readiness: A multicriteria assessment index, *Renewable and Sustainable Energy Reviews*, Vol. 131, 109988.

QADIR S. A., AL-MOTAIRI H., TAHIR F., AL-FAGIH L. 2021. Incentives and strategies for financing the renewable energy transition: A review, *Energy Reports*, Vol. 7, 3590-3606.

RAM M., AGHAHOSSEINI A., BREYER C. 2020. Job creation during the global energy transition towards 100% renewable power system by 2050, *Technological Forecasting and Social Change*, Vol. 151, 119682.

RITCHIE H., ROSADO P. 2020. *Energy mix*. OurWorldInData.org. Retrieved from https://ourworldindata.org/energy-mix

SHAHBAZ M., WANG J., DONG K., ZHAO J. 2022. The impact of digital economy on energy transition across the globe: The mediating role of government governance, *Renewable and Sustainable Energy Reviews*, Vol. 166, 112620.

SWIATOWIEC- SZCZEPAŃSKA J.; STĘPIEŃ B. 2022. Drivers of Digitalization in the Energy Sector—The Managerial Perspective from the Catching Up Economy, *Energies*, Vol. 15, 1437. https://doi.org/10.3390/ en15041437

TIAN J., YU L., XUE R., ZHUANG S., SHAN Y. 2022. Global low-carbon energy transition in the post-COVID-19 era, *Applied energy*, Vol. 307, 118205.

TSAGKARI M, ROCA J, KALLIS G. 2021. From local island energy to degrowth? exploring democracy, self- sufficiency, and renewable energy production in greece and spain. *Energy Research & Social Science* No. 81: 102288.

TZEREMES P., DOGAN E., ALAVIJEH N. K. 2023. Analyzing the nexus between energy transition, environment and ICT: a step towards COP26 targets, *Journal of Environmental Management,* Vol. 326, 116598.

VALENTI M., GRADITI G., 2020. Le Smart Grid per un futuro energetico sostenibile e sicuro, focus ENEA, 2/20 Energia, ambiente e innovazione. DOI 10.12910/EAI2020-043

WANG J., MA X., ZHANG J., ZHAO X. 2022. Impacts of digital technology on energy sustainability: China case study, *Applied Energy*, Vol. 323, 119329.

XU Q., ZHONG M., LI X. 2022. How does digitalization affect energy? International evidence. *Energy Economics*, Vol. 107, 105879

ZHANG H., GAO S., ZHOU P. 2023. Role of digitalization in energy storage technological innovation: Evidence from China, *Renewable and Sustainable Energy Reviews,* Vol. 171, 113014.

Websites

https://www.italiadomani.gov.it/en/Interventi/investimenti/rafforzamento-smart-grid.html

https://www.opencup.gov.it/portale/web/opencup/elenco-progetti?p_p_id=ricercaliberaportlet_WAR_OpenCupPortlets&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&_ricercaliberaportlet_WAR_OpenCupPortlets_action=ricerca

https://www.e-distribuzione.it/progetti-e-innovazioni/PAN/il-progetto-pan.html

https://finanza.lastampa.it/News/2023/09/08/e-distribuzione-avvia-il-progetto-edge-gli-utenti-diventano-fornitori-della-smart-grid/MTY1XzIwMjMtMDktMDhfVExC

https://www.actea-project.it/il-progetto.html

https://datacellarproject.eu/uncategorized/public-energy-data-space/

https://www.gruppoiren.it/it/everyday/energie-per-domani/2022/smart-grid-cosa-sono-e-come-funzionano.html

---

Marina ALBANESE, University of Naples Federico II, albanese@unina.it
Monica VARLESE, University of Naples Federico II, monica.varlese@unina.it

TRAVEL ANALYTICS FOR BUSINESS INTELLIGENCE: THE CASE OF ITALIAN INSTITUTE OF STATISTICS

Alessandra Dentini

Abstract. In this paper, an administrative data base for internal use is treated to show the power of the Business Intelligence for increasing the efficiency of the management system. It is not just a tool for understanding the "status quo" of Istat human resources but it is a set of statistical analysis to construct a strategy for improving the "idea" of modern Public Administration. As part of the Business Intelligence (BI) process, organizations collect data from internal IT systems and external sources, prepare it for analysis, run queries against the data and create data visualizations, BI dashboards and reports to make the analytics results available to business users for operational decision-making and strategic planning.
The duty travel information selected from administrative Istat data base are made up of a several variables which include employee identification data (registration number, profile, staffing plan), duty travel administrative data (number of assignments, duration, destination, object, and institution visited, cost center) and finally, the detail of the cost items for single duty travel expenditure (type of transport, board and lodging).

1. Introduction

In this paper, an administrative database for internal use is constructed to show the strength of Business Intelligence for increasing the efficiency of the management system. It is not just a tool for understanding the "status quo" of Istat Human Resources (HR) but it is a set of statistical analysis to construct a strategy for improving the "idea" of modern Public Administration.

The Business Intelligence (BI) is a technology-driven process for analyzing data and delivering actionable information that helps executives, managers and workers make informed business decisions (De Vivo, Polzonetti, Tapanelli, 2011). As part of the BI process, organizations collect data from internal IT systems and external sources, prepare it for analysis, run queries against the data and create data visualizations, BI dashboards and reports to make the analytics results available to business users for operational decision-making and strategic planning. In particular, statical analysis for BI plays a crucial role in:

- data-driven Decisions: Statistics provides the tools to analyze vast amounts of data, uncovering patterns, trends, and relationships that might not be

obvious otherwise. Your research can contribute to this process by identifying new insights and correlations that can fuel innovative ideas;

- quantifying Impact: Innovation often thrives on taking risks. Statistics helps assess the potential impact and viability of those risks. Your research findings can inform these calculations, providing a data-backed foundation for groundbreaking solutions;
- benchmarking and Improvement: Statistics allows us to compare new innovations to existing solutions. Your research can create new benchmarks, measuring the effectiveness of emerging technologies or processes, and pushing the boundaries of what's possible;
- hypothesis Testing and Experimentation: Innovation requires testing and refinement. Statistical methods enable the creation of robust experiments to validate new ideas and optimize their performance. Your research can contribute to this process by designing statistically sound experiments that yield reliable results.

The duty travel information selected from administrative Istat database are made up of a several variables which include employee identification data (registration number, profile, staffing plan), duty travel administrative data (number of assignments, duration, destination, object, and institution visited, cost center) and finally, the detail of the cost items for single duty travel expenditure (type of transport, board and lodging) (Dentini, Mazziotta, Zeppieri, 2022).

The data analysis covers the years 2009-2023 (Fourteen years of time series!) So that it is possible to assess the situation before, during and after the pandemic crisis.

The first phase towards the construction of the database involved the transposition of the raw matrix to obtain the single duty travel for each record (row). The statistical analyzes used have extracted from a very large matrix of unstructured administrative data the information necessary to design strategic company policies aimed at saving and investing towards new strategies for enhancing professional skills. In fact, the ability to reduce the costs of missions through the signing of agreements with service providers is a goal that Istat will pursue in the coming years.

The paper is one of the first example of Business Intelligence (BI) in Istat, in which data purely used for administrative matters are used to obtain useful information for the management of activities, economic savings policies, relations with other institutional organizations and much more. We are talking about a wealth of information to be exploited to increase the effectiveness of some internal and external processes of the Istat.

The paper is structured as follows. In the second section, the pandemic effect on the tourism sector is presented; in the third section, the results of the anlysis are focused; in the fourth the concluding remarks are highlighted.

2. The pandemic effect on the tourism sector

Following the pandemic, the tourism market experienced an unprecedented and extraordinary season. The geopolitical scenario further influenced the international environment. The rising cost of oil and energy, in general, led to widespread price increases in all sectors, including tourism, driving up prices for hotel and transportation services.

Table 1 – *Cost variation (%) for flights and rail transport.*

	2021 - 2022	2022 – 2023
Domestic flights	21%	44%
European flights	128%	43%
Intercontinental flights	46%	16%
Rail transport	14%	13%

Istat

Due to the economic crisis, the entire tourism industry has also suffered considerable hardship as a result of severe staff shortages mainly due to the large number of layoffs made between the year 2020 and 2021. During 2022, in the airline industry, the increase in airline ticket prices for international and domestic flights was due to both the increase in the price of fuel and a general increase in inflation, which led to a truly staggering increase in airline ticket costs. Istat has estimated, compared to 2021 that airline ticket prices have increased by 90 percen. This is a really significant and unprecedented price change, and the most significant increases are mainly in international flights. In 2022, airline tickets for international flights increased by 128% compared to the year 2021 and 92% compared to 2020.

As for domestic flights, the change in air ticket prices in 2022 is smaller, in fact there was a 21% increase compared to 2021.

The increase in airline ticket prices is also evident in 2023, in fact, the high fuel prices and other difficulties that have affected the industry have shown an increase in flight prices by an average of 40% compared to 2022 (Istat data referring to May). Note that as of May 2023, overall inflation was 8.3%.

The largest increase compared to the year 2022 is in domestic flights, which go up by 44 %; while intercontinental flights show a smaller increase (16 %) compared to May 2022. The price increase related to passenger travel does not only affect air transportation; in fact, maritime transportation has also been mainly affected by the situation related to the increase in fuel costs. In 2022, sea transportation fares were increased by 19% compared to 2021, although looking at the trend of fares, there was already a 21% increase last year compared to 2020 (data not in tableIl costo

degli spostamenti ferroviari, come evidenziato in Tabella 1, mostra un aumento inferiore rispetto ai voli e al trasporto marittimo. In fact, prices increase 14 % from the year 2021 to the year 2022 and 13 % from 2022 to 2023.

The hotel industry has been hard hit by the Covid-19 pandemic. The forced closure of many accommodations, staff reductions, and increased costs have had a significant impact on hotel supply and prices, with major consequences for business travel as well. According to Federalberghi's estimate, the closure of 180 hotels in Rome, as in other cities, has posed a real problem for those who have to go on duty travel. This reduction in supply, combined with increased post-pandemic demand, has made the sector less competitive, especially during peak season.

The increase in flight and hotel prices, along with other critical issues in the tourism sector, has generated a significant impact on government business travel expenses.

Entering into agreements with transportation provider companies and hotel facilities is an effective strategy to contain costs and protect the Public Administration from excessive outlays. This strategy demonstrates the Administration's focus on savings and efficiency, strengthening its image to employees and citizens. Entering into agreements requires careful evaluation and planning, which is already present in Istat's strategic policy (Dentini e Zeppieri, 2023).

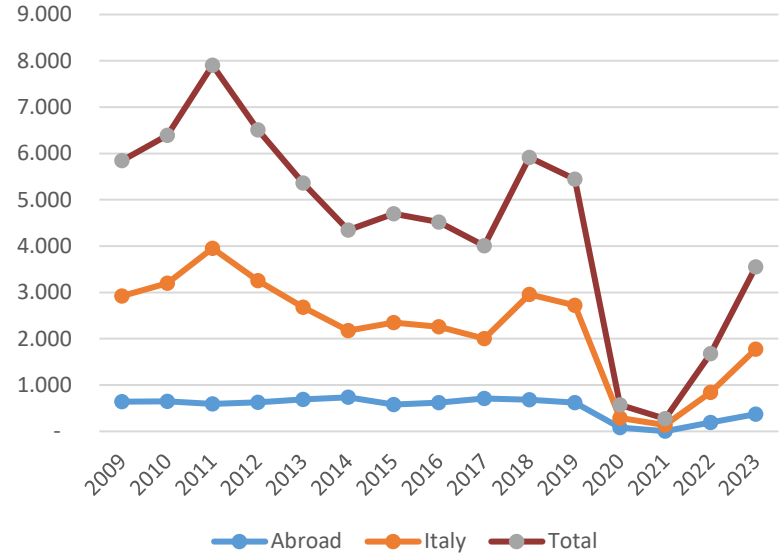
3. Analysis of the results

The descriptive analyses in this section allow us to focus on time-series data to study trends and increase awareness of appropriate strategic decisions. The goal of the analyses and to bring out evidence that can guide the decision maker in making optimal choices toward strategic business travel planning. In short, Business Intelligence becomes an irreplaceable tool for Istat's internal management.

In the figure 1, Duty travel by year and destination is presented. From Figure 1 we can see how the number of duty travel is distributed over the time period considered. It is easy to see how the duty travel peaks occur at the traditional censuses in 2010 and 2011 and the beginning of the continuous population census in 2018 and 2019. It is evident that, during the Covid period, duty travel are drastically reduced. The number of duty travel abroad seems to be constant except for the Covid time.

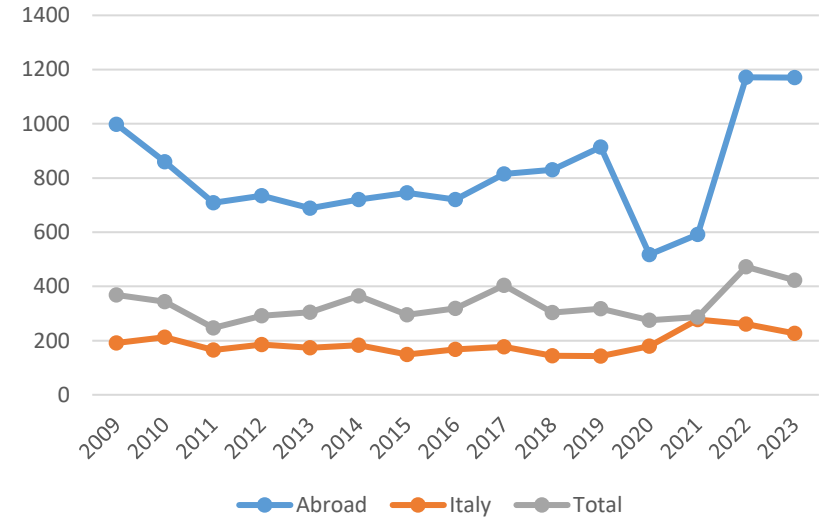
The average cost per year and destination is presented in the figure 2. Average costs show a constant trend except for the post covid time which is driven by a strong inflation.

Figure 1 – *Duty travel by year and destination-*



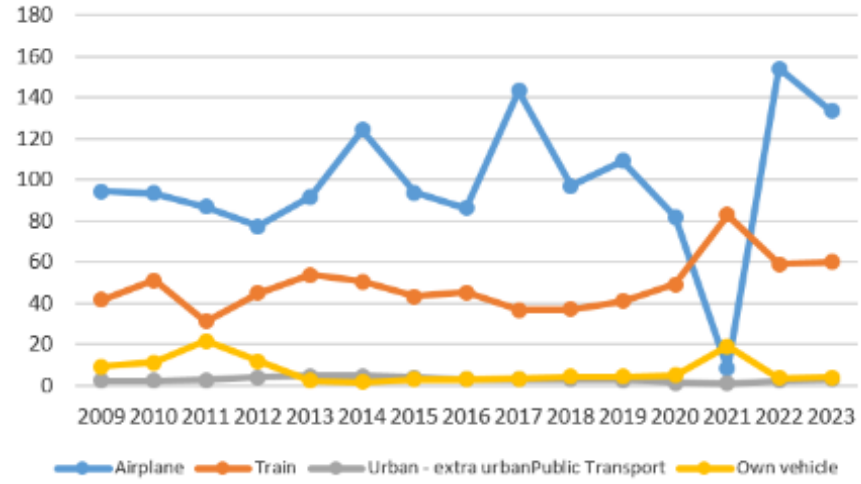
Our elaborations on duty travel data base

Figure 2 – *Average cost by year and destination.*



Our elaborations on duty travel data base

Figure 3 – *Average cost by year and transport supplier.*

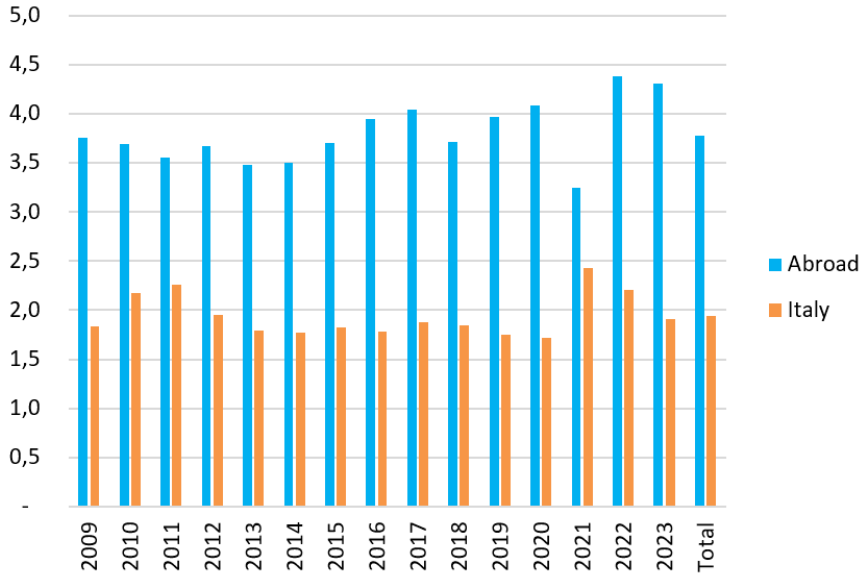


Our elaborations on duty travel data base

In Figure 3, the average cost per year and transport supplier is presented. The average cost of extra urban transport is constant throughout the time considered. The cost of using one's personal vehicle is constant over all time except in 2021 when there is a spike due to an inflationary push which drove up fuel costs. Morever, a conseguence of the pandemic time is the use of own vehicle in order to avoid infection. The average train cost is fairly constant over the years except for 2021, which shows a slight spike due to rising energy costs. The average airplane cost has a very variable trend, plummeting in 2021 and then rising to a peak in 2022.

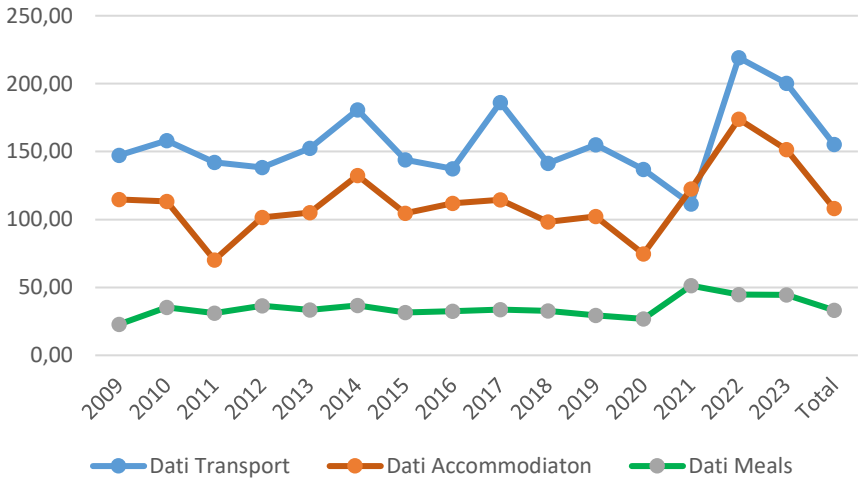
In Figure 4, we can see that the average days' duration of a mission is constant for both Italy and abroad. An exception is the value occurring during the pandemic (2021), in fact the average duration of duty travel (expressed in days) increases significantly. One possible explanation could be related to the use of own vehicle (as shown in Figure 3). The driver of the vehicle is the employee who performs the duty travel and therefore needs more recovery time. Another possible hypothesis could be the concentration of several activities in one duty travel.

Figure 4 – *Average day for duty travel.*



Our elaborations on duty travel data base

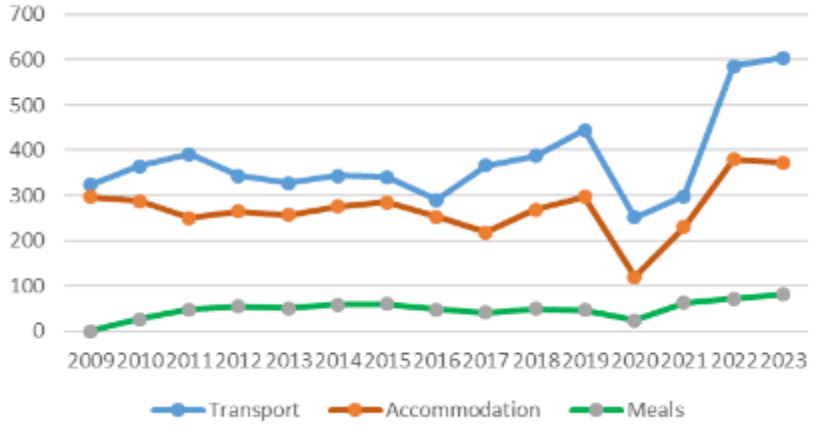
Figure 5 – *Average cost by exepense items (Italy).*



Our elaborations on duty travel data base

In Figure 5, we can see that the average cost per item for transport and accommodation is remarkably similar over time. Interestingly, the curves of transport and accomodation engage after 2020 and instead declines after 2022, all as a result of the inflationary push and the following relapse.

Figure 6 – *Average cost by exepense items (Abroad).*



Our elaborations on duty travel data base

In Figure 6, the average costs per item of expenditure are presented. The breakdowns for transport and accomodation show more or less the same trend and, as in Figure 5, we can clearly see the inflationary pressure of the last years.

Statistical analyses have shown us that the situation has become much more variable in recent years than at the beginning of the first decade. The pandemic certainly upset many market laws by causing inflation to rise and fall very abruptly. The analyses show us that in order to fight inflation, it is necessary to make agreements with suppliers in order to keep duty travel costs down.

4. Conclusions

In an economic phase dominated by an abundance of data and limited time to make decisions, it is important to exploit information to be competitive and generate business actions. Business Intelligence (BI) is a broad term that encompasses the strategies and technologies used by organizations to collect, analyze, and interpret data.The goal of BI is to provide businesses with insights that can be used to improve decision-making, optimize operations, and achieve strategic goals. BI can be used to

analyze a wide range of data, including administrative data. Applying Business Intelligence and Analytics increase the productivity of your institution and Strategic choices become easier.

Istat is starting a process of managerial growth in order to undertake activities for the coordination and use of strategies in order to provide technical-organizational support to the government structures of the Institute of Statistics.

This research experience must be understood as inserted within the context of great innovation that statistics are going through from the point of view of the use of administrative sources in order to represent complex realities with increasingly clear images that can assist the stakeholders in strategic choices. Statistical analysis plays a crucial role in driving innovation in management field. Overall, the use of administrative data in BI techniques is a powerful trend that allows organizations to gain deeper insights, make better decisions, and drive innovation at a lower cost.

Business Intelligence, traditionally associated with the private sector, is acquiring an increasingly central role in the public sector as well. Data analysis, in fact, can become a powerful tool for improving the efficiency, transparency and impact of administrative actions, significantly contributing to the creation of public value.

The expression "Public Value" means the multidimensional well-being (social, economic, environmental, health, institutional, scientific, etc.) generated for the benefit of citizens, businesses and stakeholders (External Public Value) by leveraging organizational health, professional, infrastructural, digital, economic-financial, Administration (Internal Public Value).

The concept of public value is constantly evolving and refers to the ability of public administrations to generate and sustain benefits for the community. In this context, Business Intelligence emerges as a strategic tool for measuring, analyzing and optimizing the impact of public policies.

The creation of external Public Value depends on the improvement of performance which, in turn, depends on the improvement of the health of the administration's resources: the creation of internal Public Value is, therefore, the prerequisite for the creation of external Public Value.

Thanks to data analysis, it is possible to identify inefficiencies and optimize internal processes, leading to resource savings.

The objective of this contribution is to use internal Istat administrative sources for exploiting statistical information to rationalize internal processes and carry out economies of scale for improving the organizational structure of the Institute.

Besides, the goal of this scientific paper has been achieved: in fact, agreements were made with transport and hotel providers to reduce mission costs and to invest the savings in other research activities..

The results of the statistical analyzes demonstrate that the number of duty travel has been constant over time, therefore allowing forecasts to be made on the future

amount. Furthermore, the trend of the last year is strongly increasing; therefore, we are back to pre-pandemic situation. The analyzes agree in designing a scenario in which the agreements already made are already creating considerable advantages in the organizational and financial management of Istat.

Istat is acting in a new direction, exploiting internal administrative sources to increase process efficiency and financial savings. the Italian Institute of Statistics, stands to gain significant advantages by combining its knowledge of BI statistical techniques with unstructured data for a modern working paradigm.

References

DE VIVO M.C., POLZONETTI P., TAPANELLI P. 2011. Open Data, Business Intelligence e Governance nella Pubblica Amministrazione, *Informatica e diritto*, XXXVII Annata, Vol. XX, 2011, No. 1-2, pp. 239-262.

DENTINI A., MAZZIOTTA M., ZEPPIERI I. 2022. The Destinations of Research Activity in Istat: Statistical Methods of business Intelligence for Enhancing Territorial Information, *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXXVII, No. 1 January-March 2023, pp 125-136.

CAMISASCA M., PIETRANTONIO E., MAGRO R., ARBOREA A., FABIANI A., GIORDANO C., PIETRANTONIO M., SPARANO A. 2022. A Pandemic Health Risk Management Model for the Protection of Workers: The Istat Experience, *Rivista Italiana di Economia Demografia e Statistica* - Vol. LXXVII, No. 1 January-March 2023, pp 53-64.

DENTINI A., ZEPPIERI I. 2023. Business Intelligence for Analysing Human Resources Data Base, *Rivista Italiana di Economia Demografia e Statistica* Vol. LXXVIII, No.2 Aprile-Giugno 2024, pp 7-18.

Alessandra DENTINI, Istat, aldentin@istat.it

TERRITORIAL INEQUALITIES BETWEEN DIGITAL INNOVATION AND THE LABOUR MARKET[1]

Francesca Carta, Ilaria Straccamore

Abstract. Digitalisation has redefined the strategic choices of companies and the skills required by the labour market. In Italy, digital innovation continues to suffer from significant shortcomings, which are accentuated in specific territories. On this basis, the present study aims to contribute by analysing the relationship between the digital innovation of businesses and the digital skills of workers from a complementary perspective, with the ultimate objective of outlining the digital profile of the Italian regions. The analysis begins with the selection of labour market digitalisation indicators on the demand and supply side, processed by ISTAT and Unioncamere's Excelsior Information System. The relationship between the digital innovation of enterprises and the digital skills of workers is analysed through three composite indices to capture the degree of digitalisation of enterprises, the skill level of the workforce and the expected demand for digital skills in the labour market. The synthesis methodology used here is the Mazziotta-Pareto Index (MPI). The findings indicate significant regional differences in the digitalisation paths of firms and the workforce.

1. Introduction

The contemporary economy is undergoing profound transformations driven by digitalisation, globalisation and flexible specialisation (Semenza, 2022). Digitalisation has posed many changes, such as skill composition and increasing demand for advanced digital skills (Lorraine *et al.*, 2022), making digital literacy a prerequisite for employment and economic inclusion. The demand for digital skills is projected to continue rising significantly, underscoring their critical role in enhancing business performance (OECD, 2023). Consequently, organisations are adapting by flattening hierarchies and emphasising broader competencies of human resources to meet the challenges of rapid technological advancement (Semenza, 2022). Overall, these developments highlight the imperative for both individuals and institutions to embrace digital transformation and its implications on economic competitiveness and social progress. Digitalisation has the power to boost both

[1] This article expresses only the opinions of the authors. Although the paper is the result of joint work, the sections are attributed as follows: paragraphs 1, 2 and 3 to Francesca Carta and paragraphs 4, 5 and 6 to Ilaria Straccamore.

efficiency and well-being. However, it could widen inequality if it leaves some individuals or regions behind (OECD, 2023). In Italy, signs of digital territorial inequalities emerged from an analysis that, through the measurement of the European composite indicator Digital Economy and Society Index (Desi) at the regional level (rDesi), highlighted wide regional disparities in the levels of digitalisation throughout Italy—particularly between North and South (Benecchi *et al*., 2021). Although the gap between the northern and southern regions of Italy has not narrowed over the years, there has been a general improvement in digitalisation levels across all regions (Giannini *et al*. 2022). This finding is also confirmed by a specific analysis of the digital divide at the regional level conducted in 2023. The study, using a reduced version of the Desi index, shows a higher degree of digital penetration and technological development in northern regions (Bruno *et al*., 2023).

In recent decades, the process of internal convergence has slowed, and economic crises such as the 2008 financial downturn and the COVID-19 pandemic have further exacerbated existing territorial disparities. Individuals from disadvantaged areas are more likely to have fewer opportunities to acquire the skills needed for the ongoing digital and green transitions (OECD, 2023). Consequently, European policies and Italian strategies, including the National Recovery and Resilience Plan (NRRP), are crucial for promoting digital development and reducing territorial inequalities. Based on the results of the previous literature and context analysis, this contribution aims to examine the degree of digitalisation in the labour market at the local level. It focuses on the relationship between the digital innovation of enterprises and the digital skills of the workforce, identifying common trends and regional specificities. More precisely, we analysed the digitalisation of labour supply and demand to compare the level of business digitalisation, the digital competencies available in different regions, and the digital skills required by companies. To explore these aspects, we have built three composite indices. These integrated analyses will help to identify the effects of digitalisation on the Italian labour market and measure the digital skill mismatch within it.

2. The European approach to a sustainable digital transition

The digitalisation process must be regulated to ensure that individuals without adequate resources are not excluded from its benefits (OECD, 2019). In 2023, recognising the strategic importance of digital technologies, the European Parliament declared 2021–2030 as the 'Digital Decade'[2] and set specific objectives to be achieved by 2030. This initiative aims to empower businesses and individuals

[2] Decision (EU) 2022/2481 of the European Parliament and of the Council

for a human-centred, sustainable digital future, with progress monitored annually through the Desi. With this in mind, national programmes funded by the EU's Next Generation Fund have a minimum 20% share of funding for digital transformation. In Italy, more than a quarter of NRRP resources have been earmarked for this purpose, and the ambitious initiative 'Digital Italy 2026' has been defined. This intervention focuses on two main axes: digitalisation of public administration investment and an ultrafast network. This initiative joins that of the 'Digital Republic', a national strategy that aims to reduce the digital divide and promote education on the emerging technologies. Key goals for 2026 include reaching 80% of the population with at least basic digital skills, doubling the population with advanced digital skills and increasing the number of Small and Medium Enterprises (SMEs) using ICT specialists by 50%. The EU Council recently adopted conclusions on 'The future of Europe's digital policy'[3] that aim to identify the main priorities of EU digital policy, which must be enhanced by attracting and maintaining a digitally skilled workforce and bridging the digital divide.

3. Digital innovation state of the art: enterprises and labour force

In Italy, there is a significant institutional debate on the impact of digitalisation on the labour market and the necessary policy responses. While Italian companies have made progress in adopting digital technologies, challenges persist in both upskilling workers and enhancing companies' capacity for innovation and digitalisation. According to the Excelsior Information System survey (2023), there is a growing discrepancy between the demand for digital skills and their availability in the workforce. In 2023, 58.1% of companies reported difficulties in finding professionals with the required digital competencies, marking a notable increase from previous years (41.8% in 2022 and 37.8% in 2021). Looking ahead, forecast 2024–2028 emphasises that digital skills are expected to be increasingly in demand across various sectors (Excelsior Information System, 2024).

3.1 Enterprises and the state of innovation

The innovation of enterprises has had a strong acceleration with the COVID-19 pandemic crisis. In the first epidemic phase (2020), digitalisation proved to be an essential ally in managing the crisis and mitigating its consequences in economic terms. The monthly surveys of the Excelsior Information System in May/June 2020 on critical issues and the abilities of businesses to respond during the COVID-19 lockdown period highlighted the centrality of digitalisation in riding out the crisis,

[3] Council of the European Union 9957/24 'The Future of EU Digital Policy - Council Conclusions (21 May 2024)'

mitigating its consequences and nurturing positive prospects for emerging from it in the following months. Comparing the firms that had adopted integrated digitalisation plans before the COVID-19 crisis with those that had not yet adopted them, the first group maintained business regimes unchanged from the pre-emergency ones in 36.1% of cases, compared to 28.2% in the second group. Only 8.2% of digitalised companies considered suspending or closing down their activities, compared to 12.1% of non-digitalised enterprises. The Italian government has enacted various measures to face the emergency and facilitate ongoing transformation, such as the establishment of the Digital and Technological Innovation Fund and New Skills Fund in 2020 (Relaunch decree n.34/2020). The COVID-2019 emergency has thus been translated into an innovative push of unexpected proportions that has helped our country to emerge from the bottom of the European and world rankings in terms of digital and innovation. In 2023, according to the second Digital Decade report (2024), most Italian SMEs have a basic level of digital intensity in line with the EU average (60.7% compared to the EU average of 57.7%). Progress has been particularly strong in the use of electronic invoices (97.5% in 2023 compared to the EU27 average of 38.6%) because the Italian government has put very strong regulations in place from 2014, with constant updates (ISTAT, 2023b). Moreover, in 2023, among Italian companies with at least 10 employees, 61.4% confirmed the use of cloud computing (a leading indicator of technological innovation) against 45.2% of EU27 companies. Despite this, in Italy, only 8 out of 100 companies use artificial intelligence tools. In 55.1% of businesses, it is precisely the lack of digital skills that hinders the adoption of these technologies. Notably, the difficulties are most evident for SMEs. For example, activities that require more specialised skills, such as data analysis, are carried out by 25.7% of SMEs, compared to 74.1% of large enterprises (ISTAT, 2023b).

3.2 Labour force and skills for the digital transition

Digital transformation requires specific skills for the labour force and the population as a whole. Italy has major shortcomings in this area: in 2023, only 45.8% of people have at least basic digital skills (below the EU average of 55.6%), with gaps across all age groups and with a limited dynamic in recent years (European Commission, 2024). The second European objective of the Digital Decade in the axis of action relating to human capital is to reach 20 million specialist employees in ICT by 2030 (from approximately 9 million in 2022), with convergence between men and women (the latter currently accounts for less than 20% of the total). As in the rest of Europe, in Italy, the trend of employment in ICT professions has been more favourable than in the overall labour market. However, the growth of this segment in Italy between 2012 and 2022 was 27%, compared to 58% for the EU27 as a whole. In 2023, employees in ICT professions represented 3.9% of the total in

Italy, and 4.6% in the EU27. In 2022, the number of graduates in ICT in Italy remains significantly below the forecasts of the EU Digital Decade (1.5% against 4.2% of the EU average), which has led to a shortage of qualified personnel for Italian companies. Additionally, there is a negative impact on the use of human capital and the availability of skills, especially digital skills, due to an increase in older workers (aged 65–69).

4. The methodological approach

As previously mentioned, the analysis focuses on measuring the degree of digitalisation in the labour market at the local level. From the literature review, we have drawn inspiration from the Desi and its Italian variant, the rDesi, and from the reduced composite indicator for measuring the digital divide at the regional level (Benecchi *et al*., 2021; Bruno *et al*., 2023; Giannini *et al*., 2022). The theoretical framework of these three studies has been partially reproduced; moreover, we have adopted different choices, motivated by the data at our disposal and the purpose of the work. Specifically, unlike the aforementioned works, the present study highlights two main innovations. Unlike previous studies characterised by the construction of a single composite index, our work involves the establishment of three different composite indices aimed at distinctly highlighting three aspects of labour market digitalisation: the transformation of enterprises, the skills of the workforce and the skills needs of businesses. Additionally, we have introduced the dimension of the digital skills needs of businesses as the link between the level of enterprises' digitalisation and that of the workforce's digital skills.

4.1. Conceptual framework: indices, dimensions and selected indicators

Three composite indices have been built: the *Business Digitalisation Index*, the *Labour Force Index* and the *Digital Skills Prediction Index*. Each index reflects different dimensions of analysis, and each dimension is measured by different elementary indicators, as shown in Table 1.

The *Business Digitalisation Index* reflects the state of the digital transformation on the demand side and is structured along three dimensions: digital infrastructure, the integration of digital technology, and business e-commerce. This index is measured by 7 indicators (from 'a' to 'f' in Table 1) of the ISTAT survey "The Use of ICT by Businesses", with reference to enterprises with at least 10 employees (ISTAT, 2023b). The *Labour Force Index* measures the level of digital skills on the

supply side[4] and is structured along three dimensions: households' connectivity, use of internet services, and digital skills. This index is measured by 12 indicators (from 'g' to 'r' in Table 1) of the ISTAT survey 'Aspects of Daily Life', with an ad hoc module on the use of digital technologies by citizens (ISTAT, 2023a), and the ISTAT survey on "The Use of ICT by Businesses" (ISTAT, 2023b). The *Digital Skills Prediction Index* provides another measure on the demand side regarding the companies' forecasting needs for digital skills. This index is structured along the digital requirements for the professions dimension, consisting of indicators 's' to 'u' in Table 1, which originate from the Unioncamere and ANPAL survey 'Demand Analysis of Digital Skills in Enterprises' (Excelsior Information System, 2023).

The first two indices utilise stock data, whereas the third index relies on flow data. Indicators were selected based on their relevance, validity, timeliness and accessibility, aiming to strike a balance between redundancy and information loss.

4.2. Statistical method

Measuring a digital economy and society is a multidimensional phenomenon that requires the aggregation of many indicators. A composite index is formed when individual indicators are compiled into a single index based on an underlying model. It should ideally measure multidimensional concepts that cannot be captured by a single indicator (OECD, 2008).

The synthesis methodology used in this context is the Mazziotta-Pareto Index, in the MPI version (Mazziotta and Pareto, 2020). It is a composite 'non-compensatory' indicator based on the hypothesis of 'non-replaceability' of the value of one elementary indicator with another. It also enables simplicity and transparency of calculation, robustness, and the immediate use and interpretation of results. These characteristics have made it particularly suitable for our purpose.

All indicators are normalised into z-score, with an average of 100 and an average square error of 10. Therefore, the transformed values will be in the range of approximately 70–130. Based on the arithmetic mean of z-scores, the aggregation function is corrected by a penalty coefficient that penalises units with the greatest imbalances between the individual indicators while having the same average.

All selected indicators have been selected with positive polarity; that is, as indicators in accordance with the analysed phenomenon. For the weighting system, we decided to assign the same weight to all indicators.

[4] To measure the digital skills of the workforce, we used indicators related to citizens' digital skills. Since these indicators could also refer to the so-called 'no workforce' (such as inactive or retired people), we will use them as a proxy for the digital skills of the workforce.

Table 1 – *Indices, dimensions and selected indicators.*

Index	Dimension	Indicator
Business digitalisation	Digital infrastructure	a. enterprises connected to internet through fixed broadband
	Integration of digital technology	b. enterprises with a website or a homepage c. enterprises having Enterprise Resource Planning (ERP)[5] software d. enterprises using at least one social media e. enterprises buying cloud computing services
	Business e-commerce	*f.* enterprises active in e-commerce in the previous year in on-line sales, selling at least 1% of total sales
	Households' connectivity	g. households with internet at home
Labour Force	Use of internet services	People aged 14 and over who have used the internet in the last 12 months to access public administration online services to: h. download or print official forms i. make an appointment/make a reservation (e.g. clinics, etc.) j. get information about services, benefits, etc. k. request for certificates/documents (e.g. birth, residence, etc.) l. enrolment in school or university m. application for social security benefits (e.g. pension, etc.) n. persons aged 14 and over who have used the internet in the last 12 months for the use of digital identity (SPID/CIE) o. persons aged 14 and over who have used the internet in the last 12 months to purchase goods/services for private use
	Digital skills	p. persons employed using computers with access to internet at least once a week q. persons aged 16-74 who have used the internet in the last 3 months and have at least basic digital skills *r.* 16-74 year old people who have used the internet in the last 3 months and have digital skills above the basic level
Digital Skills Prediction	Digital requirements for professions	Predicted new entrants by companies for 2023 according to the ability to use: s. mathematical languages and methods t. internet technologies/multimedia communication tools u. Industry 4.0 technologies to innovate processes

5. Results

5.1. Business Digitalisation Index

The *Business Digitalisation Index* shows that all the northern regions, except for Liguria and Valle d'Aosta, and Lazio in the centre, record the highest level of

[5] Software to manage resources by sharing information between different functional areas (e.g. accounting, planning, production, marketing, etc.). It can be a standard, customized or a self-created software.

digitalisation, while the regions of the South lag behind (Table 2). The components of this composite index contribute differently to the results. The *digital infrastructure of companies*, measured by the fixed broadband connection, is highly and uniformly distributed throughout the territory (Italian average of 97.2% and standard deviation of 2.8%). The *integration of digital technology* by enterprises is measured by the number of companies that have a website/home page, ERP software and social media, and by enterprises that purchase cloud computing services. The use of websites is the most widespread technology (average of 69.7%); however, this is characterised by the most heterogeneous distribution between Italian regions (standard deviation of 9.9%). A technology less commonly used is ERP software (average of 37.7% and standard deviation of 7.4%). For all these components, the regions with the highest scores are mainly concentrated in the North, with the exception of Sardinia and Puglia for the use of social media and Trentino Alto Adige for a low percentage of companies using cloud computing services. *E-commerce* remains rare within Italian companies (average of 13.9%). In the regional ranking, Sardinia, Trentino and Valle d'Aosta are at the top, followed by Calabria and Campania, while Molise, Umbria and Liguria are the last three. Therefore, e-commerce partially reduces the disparity between the North and the South of the country. This result is in line with the general evidence suggesting that e-commerce is relatively more widespread among regions with a higher concentration in the service economy.

5.2. Labour Force Index

The *Labour Force Index* shows a large disparity between the Centre-North and the South, with the exception of Sardinia (Table 2). The components of this index contribute differently to the results. *Households' connectivity* has a wide and homogeneous spread in all the regions (average of 83.2% and standard deviation of 3.4%). The *use of internet services* is measured by persons who use digital identity, online government services and websites, and who order or purchase goods or services online. The latter is the most widespread competence (on average 49.3%), while the use of public administration websites for the application of social security benefits, enrolment in school or university, and requests for certificates are the least used (average of 10.0%, 11.9%, 12.3% respectively). These components confirm the North-South gap, with Liguria being the only northern region below the Italian average, while Sardinia is the only southern region above the Italian average. *Digital skills* are measured by the employees of companies using the internet at work, by people who have at least basic digital skills or digital skills above the basic level. The former is the most widespread competence (on average, 51.6% against 23.4% for basic digital skills and 21.4% for digital skills above the basic level). All of these

indicators highlight the North-South gap, with the Lazio region at the top of the ranking for workers' digital skills and citizens' advanced digital skills.

5.3. Digital Skills Prediction Index

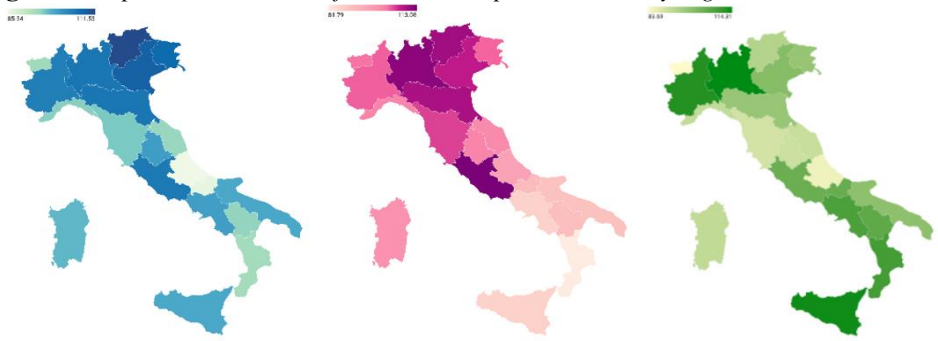
The *Digital Skills Prediction Index* measures companies' predictive demand for digital skills required for the profession. With the other two composite indices, this index allows a better understanding of territorial dynamics, completing the picture of businesses' digitalisation levels and the digital skills available in the territory, with the digital competency needs of companies. The ability to use internet technologies (e.g. desktop PCs, internet browsers, etc.) and manage visual and multimedia communication tools is the most widespread skill required by companies (on average, 63.4% of the predicted new entrants in 2023), while the use of mathematical languages and methods is the second (on average 50.6%). The use of Industry 4.0 technologies (e.g. artificial intelligence, internet of things, data analytics, big data, etc.) is required for 37.1% of the predicted new entrants. The *Digital Skills Prediction Index* displays the highest variance among all three composite indices, showing a heterogeneous distribution across the territory. Lombardy and Piedmont are the regions with the greatest volume of requests for digital competencies, followed by Sicily, Calabria and Campania (Table 2). Valle d'Aosta and Abruzzo are at the bottom of this ranking. In particular, Lombardy and Piedmont, followed by the southern regions, are at the top of the ranking for the search for the ability to use internet technologies, multimedia communication tools, and mathematical languages. Southern regions, particularly Calabria and Sicily, ranked highest in terms of the search for the ability to use Industry 4.0 technologies.

5.4. Regional profiles: an integrated overview

Comparing the three indices used in this study reveals significant territorial divides in the digital skills of the workforce. Both the demand for digital skills by firms (*Digital Skills Prediction Index* results) and the supply of these skills (*Labour Force Index* results) show greater regional variability than the adoption of digital technologies by companies. Moreover, by comparing the blue map with the pink one in Figure 1, we observe that the spread of business digitalisation aligns with the distribution of digital skills available in the territories, indicating clear spatial patterns. Northern regions and Lazio show the most innovative digital systems and higher digital skills, while the rest of the country lags behind. This evidence could highlight a mutual relationship according to which the digital skills in the territory could not only contribute to the construction of a digital industrial system but also be fed by it (e.g. through training programmes being promoted by companies). Another consideration emerges by comparing the green map with the blue and pink

maps in Figure 1, where we can see that the demand for digital skills by enterprises sometimes reflects both business digitalisation and the dissemination of digital skills (e.g. Piedmont and Lombardy), but not always (e.g. Sicily, Calabria and Campania). This result could have several possible explanations.

Figure 1 − *Spatial distribution of the three composite indices by region.*



Source: ISTAT and Excelsior Information System

The first one is that the business demand for digital skills is based on forecasting needs and, therefore, on provisional flow data, which implies that regions with low numbers of new entrants expected in 2023 could have acquired the necessary skills before. Another aspect is the stimulus provided by the European Cohesion Policy and the NRRP, which, with substantial funding, could have encouraged a greater demand for digital skills, especially for companies in the southern regions. In 2023, for example, as part of the National Strategy of Smart Specialisation, funding was allocated to companies in the southern regions. Another explanation may lie in the lower level of professional attractiveness of the South.

6. Conclusions

Digital innovation changes production processes and requires the adaptation of workers' skills. From this perspective, our paper analyses the relationship between enterprises' digital innovation and workers' digital skills by using a complementarity approach. We focus on measuring digitalisation at the regional level, comparing three aspects of the digital economy: the transformation of enterprises, workforce skills and business needs. This integrated analysis reveals regional disparities in the digital innovation paths of companies and the workforce, and highlights the skill mismatch in the labour market, which is pronounced in southern Italy. Moreover, for Italy, which lags far behind other EU countries in the digital transformation, the development of digital skills in human capital continues to be a major challenge.

Table 2 – *Business Digitalisation Index, Labour Force Index and Digital Skills Prediction Index at the regional level. Year 2023, Italy.*

REGIONS	Business's digitalisation Index	Labour Force Index	Digital Skills Prediction Index
Piedmont	104.6	103.8	112.6
Valle D'Aosta/Vallée d'Aoste	94.3	102.7	83.7
Lombardy	105.8	111.6	114.3
Trentino Alto Adige/Sudtirol	111.5	110.1	94.4
Veneto	108.5	107.7	100.6
Friuli Venezia Giulia	107.2	103.5	98.5
Liguria	95.9	101.7	92.0
Emilia Romagna	105.7	109.4	98.2
Tuscany	96.4	105.2	90.0
Umbria	100.6	101.7	90.9
Marche	95.1	100.6	91.5
Lazio	105.4	113.1	104.4
Abruzzo	85.5	96.5	85.8
Molise	86.5	91.0	105.0
Campania	100.3	86.5	108.4
Puglia	99.0	89.6	99.7
Basilicata	95.5	90.4	106.2
Calabria	93.9	81.8	109.4
Sicily	99.0	86.6	113.8
Sardinia	97.9	99.5	92.4

Source: ISTAT and Excelsior Information System

We believe that the methodological framework of this study and its results could be useful to monitor these issues over time and attempt to measure the effectiveness of policies, which will be fully assessed close to the deadlines of the European and national targets.

References

BENECCHI A., BOTTONI C., CIAPANNA E., FRIGO A., MILAN A., SCARINZI E. 2021. Digitalisation in Italy: evidence from a new regional index. *Questioni di Economia e Finanza (Occasional Papers),* No. 662, Banca d'Italia Eurosistema.

BRUNO G., DIGLIO A., PICCOLO C., PIPICELLI E. 2023. A reduced Composite Indicator for Digital Divide measurement at the regional level: An application to the Digital Economy and Society Index (DESI), *Technological Forecasting and Social Change*, Vol. 190, No.122461, pp. 1-16.

EUROPEAN COMMISSION. 2024. *2030 Digital Decade - Report on the state of the Digital Decade 2023*. Luxembourg: Publications Office of the European Union.

GIANNINI V., IACOBUCCI D., PERUGINI F. 2022. I divari regionali nella digitalizzazione. In CORÒ G., DE CASTRIS M. and SCALERA D. (Eds.) *PNRR Italia - Il difficile equilibrio tra i territori*, Roma: Donzelli Editore.
ISTAT. 2023a. Cittadini e Ict, anno 2023. *Statistica Report*, ISTAT.
ISTAT. 2023b. Imprese e Ict, anno 2023. *Statistica Report*, ISTAT.
LORRAINE C., SHUTING X., COUTTS A.P. 2022. *Digitalization and Employment - A Review*. Geneva: Publications of the ILO.
MAZZIOTTA M., PARETO A. 2020. *Gli indici sintetici*. Torino: Giappichelli.
OECD. 2023. *OECD Skills Outlook 2023: Skills for a Resilient Green and Digital Transition*. Paris: OECD.
OECD. 2019. *OECD Employment Outlook 2019: The Future of Work*. Paris: OECD
OECD. 2008. *Handbook on constructing composite indicators. Methodology and user guide*. Paris: OECD.
SEMENZA R. 2022. *Manuale di Sociologia del lavoro*. Torino: Utet Editore.
SISTEMA INFORMATIVO EXCELSIOR. 2020. *Impatti dell'emergenza COVID-19 sulle imprese e prospettive di ripresa rilevati attraverso il sistema informativo Excelsior*. Roma: Unioncamere e ANPAL.
SISTEMA INFORMATIVO EXCELSIOR. 2024. *Previsioni dei fabbisogni occupazionali e professionali in Italia a medio termine (2024-2028)*. Roma: Unioncamere e ANPAL.
SISTEMA INFORMATIVO EXCELSIOR. 2023. *Le competenze digitali. Analisi della domanda di competenze digitali nelle imprese, indagine 2023*. Roma: Unioncamere e ANPAL.

Francesca CARTA, Istat, francesca.carta@istat.it
Ilaria STRACCAMORE, Istat, straccamore@istat.it

EXPLORING THE CONCEPT OF SOCIAL INNOVATION USING OFFICIAL STATISTICS[1]

Stefania Della Queva, Manuela Nicosia, Sabrina Stoppiello

Abstract. The concept of social innovation is linked to a new way of producing goods and services able to enhance the quality of life by developing new processes of integration in the labour market, creating new relations, meeting new needs (Mulgan, 2007). Moreover, the relevance of social innovation grows in the local development models as factor of change of territorial context (Moulaert *et al*., 2017). The agents of this process are different: public institutions, enterprises, nonprofit institutions. Within the Istat Research Project "The nonprofit sector as factor of local development and social innovation", the concept of social innovation has been decomposed into its main dimensions in order to identify specific indicators able to describe the socioeconomic actor's initiatives and to explore their role in the regional development.
The Research Project applied a methodology based on the following steps: study of literature; breakdown of social innovation concept and definition of the dimensions to study; identification of indicators and the best territorial scale according to the research aim and to available data. A multivariate analysis was performed on data collected by Istat surveys on socio-economic units. In particular, the data sources were: Permanent Census of NPI, Permanent Census of Enterprises, Permanent Census of Public Institutions, Survey on Research and Development in enterprises and Community Innovation Survey (CIS).
The paper describes the main results about the different steps of analysis, from the operational definition of the concept to the multidimensional analysis performed in order to explore the possibility of studying the social innovation by applying official statistics in territorial analysis.

1. The Research Project

The paper describes the results achieved with the Research Project "The nonprofit sector as driver of local development and social innovation", approved by the Scientific Committee for the Innovation and scientific research Laboratories set up by Istat (2017/2018).

[1] Although the contribution is the joint responsibility of the authors, sections 1 and 3.3 are attributed to Sabrina Stoppiello, sections 2.1 and 3.2 are attributed to Manuela Nicosia, section 2.2 and 3.1 are attributed to Stefania Della Queva. Final remarks and future steps (section 4) is the result of joint work by all the authors.

The aim of the Project was to analyse the role of nonprofit sector as factor of local development and social innovation, within the socio-economic context. Based on data analysis, a secondary objective was enhancing data availability, in order to create indicators able to measure social innovation at territorial level. The Research Project applied a methodology based on the following steps: study of literature; breakdown of social innovation concept and definition of the dimensions to study; identification of indicators and the best territorial scale according to the research aim and to available data; multivariate analysis with explorative aim. In order to improve the measurement of social innovation within official statistics, new contents have been defined to be included in the surveys on socio-economic units.

During the three years-Project, the concept of social innovation has been studied and explorative analysis have been performed by applying indicators from official statistics in order to enhance the statistical sources available. It is important to stress that the research have been carried out thanks also the cooperation with experts from University and the Third sector, whom contributed to enrich each step with their reflections and allowed to come up with a shared vision of results.

2. Methodology and data

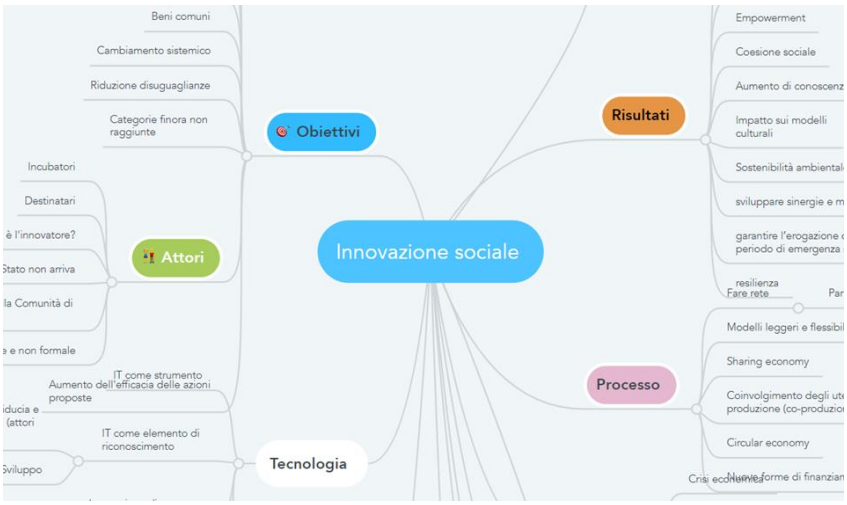
2.1 Breakdown of social innovation concept and definition of data analysis dimensions

The concept of 'social innovation' is closely linked to the economic and socio-cultural changes that have arisen since the Second World War as a result of globalisation, international financial crises, as well as climate change and the increase and ageing of the world population. There was the need for European welfare systems to face new demographic, social and cultural challenges and, at the same time, to involve different actors, whom are in charge of community services in the definition of public policies. What emerges from the study of the literature is the polysemic nature of the concept, which has led scholars to identify different connotations and construct taxonomies to define its theoretical boundaries. We can find different theoretical approaches, which highlight different aspect of social innovation. One the most important comes from Mulgan, who defines social innovation as "new ideas that work in meeting social goals … Innovative activities and services that are motivated by the goal of meeting a social need and that are predominantly developed and diffused through organisations whose primary purposes are social" (Mulgan, 2007, p. 8). Among others, Moulaert (2017) underlines the connection between social innovation and the new models of local development. Social innovation takes play as factor of change of socio-economic

context. The new models of economics and way to produce goods and services, which lead to the concept of 'social innovation', refer to the development of new services that improve the quality of life of individuals and communities, the development of new labour market integration processes and new skills and professions, and the creation of new forms of participation (Murray et al, 2010). The actors in these processes are different in nature: public institutions, business organizations and nonprofit institutions.

During the first step of research, the teamwork performed a brainstorming on social innovation definition, in order to compare different ideas and to develop new visions. Each definition was argued and each concept enunciated went into a shared picture, enriching the group's reflections. The ideas that emerged were reordered and grouped around four key-concepts, such as: Process, Objectives, Actors, and Results. The method of mind mapping (Buzan, 1995) was applied to the result of brainstorming. The mind map presents hierarchies and associations between concepts more clearly. The team worked with a digital mode during the lockdown period (in July 2020) addressing specific tasks: better specify the concepts already present; add more "branches" and sub-branches, i.e., other related concepts; add associations between concepts and branches (Figure 1).

Figure 1 – *Mind mapping on social innovation concept implemented by the teamwork and experts.*



Source: Istat – Research Project (Innovation and scientific research Laboratories)

The final step allowed to identify the dimensions of social innovation to be studied. Social innovation concerns: 1) the pursued objective, which is linked to social needs, and focuses on community well-being; 2) the result achieved, in terms of services, products or creation of new relations; 3) the process, in terms of collective initiative able to change the community conditions; 4) the resources (human, economic and social) which social innovation involve when it takes action. Based on these four dimensions, single indicators were selected in order to measure the different aspects of social innovation.

2.2. The selection of indicators and data analysis

After the definition of conceptual model through the identification of the dimensions and sub-dimensions that characterize the concept of social innovation, the study involved into the operationalization phase by the association of the elementary indicators with the sub-dimensions. The variables, included in the analysis, were identified and selected by various statistical sources, which allowed us to obtain a very rich informative dataset.

In particular, the data sources were the Permanent Census of NPI (2015), the Permanent Census of Enterprises (2018), the Permanent Census of Public Institutions (2017), the Survey on Research and Development in enterprises (2017) and the Community Innovation Survey-CIS (2018). Although the reference year of the data is different, it is necessary to underline that the sources have been selected to represent the information in a limited period and thus allow a comparison between the different data. Even though, in some cases, more up-to-date data exists, it was preferred to use those with a time reference closer to that of the census results, which play a central role in this analysis process.

All selected variables, on which we built the set of indicators useful for investigating the phenomenon, are at the regional level so that the first indicators built refer to the presence/diffusion of public, profit and nonprofit institutions in relation to the resident population and of particular types/sectors of activity.

The indicators relating to NPIs, which allow assumptions on the capacity of NPIs to generate social innovation, have been outlined based on some significant objectives, which can also be read in the analysis in terms of results (where produced). The objectives outlined are the reduction of inequalities and the social inclusion of disadvantaged and vulnerable people; the ability to reach new social categories; community wellbeing; the care of the commons; the economic development of the territorial context of reference; innovation in fundraising activities. Furthermore, another pillar on which social innovation is based is the creation of new relationships, therefore a fundamental indicator included in the

analysis is the spread of multi-stakeholder nonprofit institutions, capable of structuring significant relationships with different stakeholders (with different ways of involving them). Considering the subjects with which the NPIs structure significant relationships, two indicators were included in the analysis: the incidence of the NPIs which have beneficiaries as stakeholders and the NPIs that have public institutions as stakeholders.

As regards public institutions, the main indicators are the services provided in the sectors of social assistance and personal services, environmental protection and sustainability, and the research and development activities (by direct or indirect management). The survey on R&D activities in public institutions made it possible to select indicators on the propensity to implement innovation-oriented policies and on the relevance of R&D public spending on (considered in terms of impact on GDP) which represents a strategic variable for measuring the competitiveness of economic systems. In the questionnaire for the multi-purpose survey of enterprises was included a new section to investigate social and environmental responsibility as well as the role they play in the implementation of initiatives of collective interest. Thanks to this section, many indicators were selected precisely following the articulation of the questions. However, as regards the dimension that refers to R&D activities, the indicators were selected from the information bases relating to the surveys on Research and Development and on Innovation in companies.

A multivariate analysis was performed on data collected. In the principal component analysis carried out for exploratory purposes, a total of 46 numerical variables were taken into consideration, synthesized through linear combinations, in two synthetic indices (principal components), which allowed to outline the regional contexts. For the purposes of providing a summary framework, the first two components were selected which reproduce 46% of the variance and which define the different dimensions of social innovation linked to the presence of different peculiarities of the non-profit sector, businesses and public institutions on the territories. As is known, the first principal component extracted is the one that reproduces the maximum percentage of variance (in our case 27.8%) and the second, orthogonal to the first, presents a slightly lower percentage of variance extracted (18.7%). Given the small amount of variance explained by the subsequent components, it was decided not to proceed with the interpretation of other components[2].

[2] The third component extracted reproduces the 10.8% of the variance. The study of the graphic representation of the components' eigenvalues according to their order of extraction shows how the line presents a strong inclination in correspondence with the second component after which an inflection is recorded.

3. Results

3.1 Profit and nonprofit sectors for community wellbeing

The first component represents the role of nonprofit institutions for community wellbeing and their spread across the territories (Figure 2).

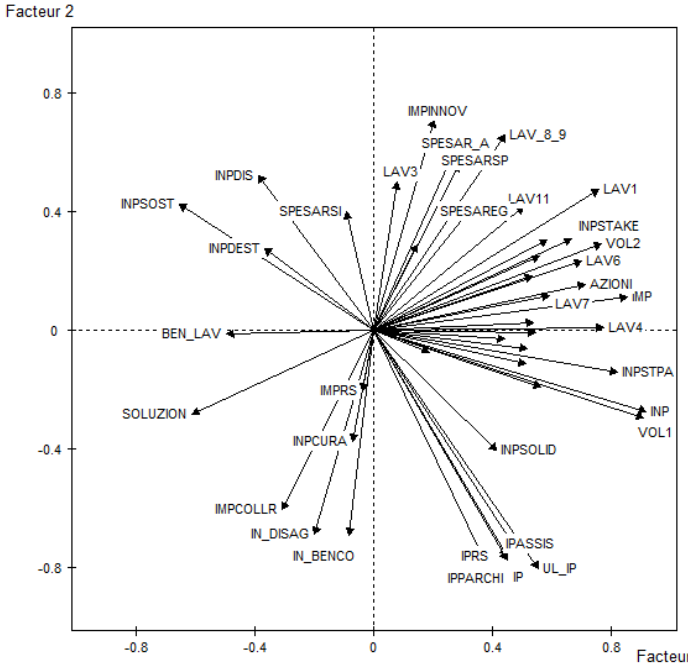
The positive values of the index, related to the first principal component, are associated with the significant presence of nonprofit institutions (Inp in Figure 2[3]) and enterprises in relation to the resident population (Imp). The public sector emerges as key-stakeholder of nonprofit institutions (Inpstpa), which build relations with different actors (Inpstake) as well. The index describes the presence of employees of nonprofit institutions engaged in the provision of services of social assistance and civil protection (Lav6). The identified component is also characterized by the presence of NPIs and enterprises dedicated to the economic and social wellbeing. On the one hand, the component synthesizes the relevant presence of paid workers in the sectors of Culture, sports and recreation (Lav1), Economic development and social cohesion (Lav6), Law, advocacy and politics (Lav7). On the other hand, the component describes the role of enterprises that have undertaken at least one action to reduce the consumption of natural resources and sustainably manage them (Azioni), by reflecting the commitment of these economic institutions in the future development of territories that is also achieved through attention to the environment. The indicators related to the presence of nonprofit institutions dedicated to the support of weak or struggling individuals (Inpsost) and oriented to fragile groups (Inpdis), with a role therefore in supporting social vulnerability, are associated with the negative semi-axis of the index. This typology is flanked by nonprofit institutions that recognize as stakeholders their beneficiaries (Inpdest), who are “at the center” of both planning and carrying out various social activities. The business sector is characterized by the presence of enterprises that have adopted at least one solution to promote environmental sustainability (Soluzion) or initiatives in favour of labour welfare and the development of equal opportunities within the company (Ben_lavo). The action of these economic actors also reaches outward, paying particular attention to combating poverty and social hardship (In_disag). The dense network of institutions is linked to the properly innovative component, which is that of research and development. Indeed, the variables referring to R&D activity in collaboration with external entities (Impcollr) and the high incidence of public R&D spending on GDP (Spesarsi) are found here.

[3] Figure 2 shows the representation of the variables on the first two components; variables’ labels are included in brackets when each single indicator is described.

3.2 The role of different actors for the local development

The second component provides a synthetic measure of the role which different actors (public, private for profit and nonprofit) play in the local development, by considering their offered services, the initiatives carried out in addressing community wellbeing, the care of common goods and driving to innovation (Figure 2).

Figure 2 – *Graphical representation of variables on the first two components.*



Source: Our data processing from Istat – Research Project (Innovation and scientific research Laboratories)

Negative values of the index are associated with a strong prevalence of public institutions (Ip; Ul_ip) in the territory and related services, especially those for environmental protection (Ipparchi) and assistance, public charity and personal care (Ipassis). The for-profit sector is characterized by enterprises that carry out initiatives for the public interest (In_benco) and against poverty and social hardship (In_disag). Drive for innovation at local level is well represented in the index, both in the public and private sector. In fact, the component synthesizes the presence of R&D services provided by public institutions (directly managed by the PA and/or entrusted to third

parties) (Iprs) and the presence of enterprises that have carried out R&D activities in cooperation with other organizations (Impcollr). As for nonprofit sector, negative values of the Component are associated with NPIs oriented to solidarity (Inpsolid), NPIs which mission is caring of commons (Inpcura) and voluntarism on Culture, sport and recreation (Vol1), which represents the most relevant part of the sector.

Positive values of the component represent a different contribution from actors to the local development. The index synthesizes the for profit sector investment on R&S, through the presence of enterprises that have introduced product and/or process innovations (Impinnov), the incidence of public (Spesarsi) and private R&D expenditure on GDP (Spesarsp), the incidence of business R&D expenditure (Spesar_a) and the average of regional expenditure on business innovation (Spesareg). Associated with these kind of economic actors are NPIs active in sectors more closely linked to hardship and social emergency. In fact, the component synthesizes the prevalence of NPIs focusing on vulnerability and the higher share of employment in sectors such as Health (Lav3), Philanthropy and promotion of voluntary work and International cooperation (Lav_8_9).

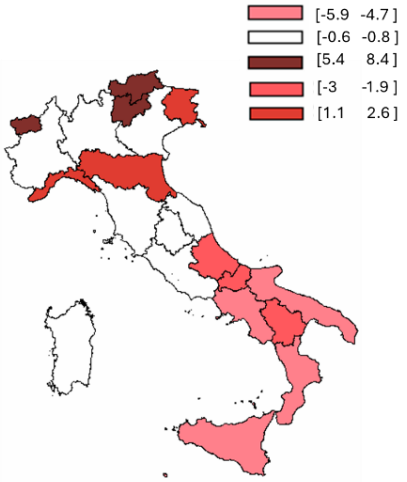
3.3 The regional contexts through the two components

The analysis of the Regions' scores on the first two components allows to describe the territories by considering the dimensions and to identify some homogeneities. Looking at the role of economic actors for the community wellbeing, Southern Italy shows a greater presence of NPIs oriented towards hardship and vulnerability, along with the presence of enterprises aimed at community wellbeing and at social responsibility. In the Autonomous Provinces of Trento and Bolzano, Valle d'Aosta, Emilia-Romagna, Liguria and Friuli Venezia Giulia there is a greater presence of volunteering and nonprofit institutions, of enterprises and networks between the various economic actors, as well (Figure 3).

Emilia-Romagna, followed by Lazio, Lombardia, Toscana and Piemonte shows the highest scores on the second component outlined, which describe the role of different actors for the local development. These regions are characterised by investments in research and development, nonprofit employment in healthcare, NPIs dedicated to the disadvantaged and fragile categories. In the regions of Southern Italy, in the Autonomous Provinces of Trento and Bolzano and in Valle d'Aosta, there is high presence of public institutions active in the management of services on the territory, businesses oriented towards social responsibility and NPIs dedicated to the community and common goods (Figure 4).

Figure 3 – *Graphical representation of regional scores on the first principal component*

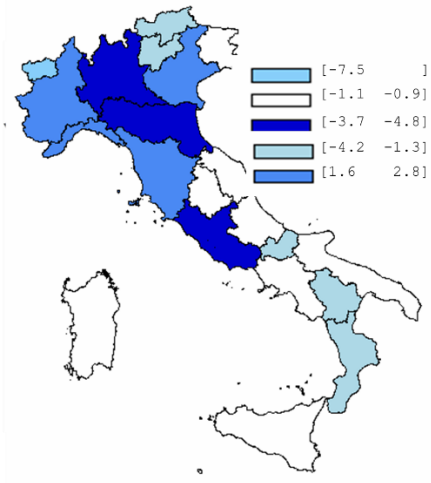
Regions	First component
Campania	-5,9
Calabria	-5,1
Puglia	-4,9
Sicilia	-4,7
Basilicata	-3,0
Abruzzo	-2,2
Molise	-1,9
Sardegna	-0,6
Veneto	0,3
Lombardia	0,4
Marche	0,4
Umbria	0,5
Toscana	0,5
Piemonte	0,7
Lazio	0,8
Liguria	1,1
Emilia-Romagna	1,4
Friuli-Venezia Giulia	2,6
Valle d'Aosta / Vallée d'Aoste	5,4
Provincia Autonoma Trento	5,8
Provincia Autonoma Bolzano	8,4



Source: Our data processing from Istat – Research Project (Innovation and scientific research Laboratories)

Figure 4 – *Graphical representation of regional scores on the second principal component.*

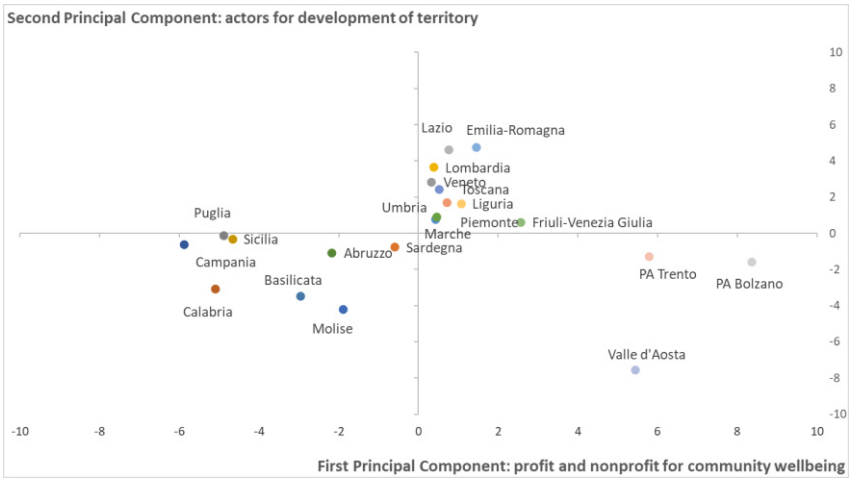
Regions	Second component
Valle d'Aosta / Vallée d'Aoste	-7,5
Molise	-4,2
Basilicata	-3,5
Calabria	-3,1
Provincia Autonoma Bolzano/ Bozen	-1,6
Provincia Autonoma Trento	-1,3
Abruzzo	-1,1
Sardegna	-0,8
Campania	-0,6
Sicilia	-0,3
Puglia	-0,1
Friuli-Venezia Giulia	0,6
Marche	0,8
Umbria	0,9
Liguria	1,6
Piemonte	1,7
Toscana	2,4
Veneto	2,8
Lombardia	3,7
Lazio	4,6
Emilia-Romagna	4,8



Source: Our data processing from Istat – Research Project (Innovation and scientific research Laboratories).

In order to summarize the results of multivariate analysis, it is interesting looking at the position of Italian regions in the two component simultaneously (Figure 5).

Figure 5 – *Regional scores on principal components.*



Source: Our data processing from Istat – Research Project (Innovation and scientific research Laboratories)

In the second quadrant, at the top right (given by the intersection of the positive semi-axes of the first and second components), we find the regions in the Centre-North of Italy, including Lombardia, Emilia-Romagna, Toscana and Veneto, characterised by the forms of social innovation identified by the second component. In these territories we find a capillary infrastructural endowment of enterprises and NPIs (among which those active in the cultural, sports and recreational sector and employing paid staff emerge). Crucial factors of development, such as private sector investments in research & development, high average of regional expenditure for business innovation and presence of enterprises that innovate and collaborate with other organizations are associated to those indicators. The region of Lazio stands out in this group because its socio-economic system, which is mainly characterised by the presence of NPIs dedicated to the various social emergencies and voluntarism committed to assisting vulnerable people.

In the third quadrant of Figure 5, we find the Autonomous Provinces of Bolzano and Trento and Valle d'Aosta, whose socio-economic system is characterised by the high presence of the nonprofit sector oriented to the wellbeing of the community and to the care of commons and of volunteering linked to expressive, cultural and socialisation needs. In these contexts, the propensity of NPIs towards multi-

stakeholder relations also prevails, especially with PA subjects, which is present to a capillary extent (to resident population ratio), offering specific services linked to environmental care, assistance and public charity, and research and development services. As for the business sector, we find the same high incidence of units to the resident population ratio, connected to enterprises that have taken at least one action to reduce the consumption of natural resources and apply the sustainability to the business management.

Finally, in the fourth quadrant, the regions of Southern Italy are characterized by the presence of enterprises oriented to social responsibility, which have undertaken initiatives of community wellbeing, to fight poverty and social hardship, the incidence of which to the total number of enterprises is particularly evident in Calabria, Campania and Basilicata. In these territories, enterprises that have adopted solutions to promote environmental sustainability, such as the acquisition of voluntary environmental product or process certifications, the drafting of sustainability reports, and the designation of an internal contact person, are also widespread. The incidence of this kind of enterprises (out of the rest of enterprises in the region) is also important in Puglia and Sicily. The attention of for-profit sector to the human resources, through the adoption of measures to promote working wellbeing and guarantee equal opportunities within the company and thus favour the professional growth and empowerment of employees, is considerable especially in Calabria and Sardegna. As for the nonprofit sector, the presence of human resources involved in the social assistance and civil protection sector (both volunteers and paid workers) is significant in these territories. The presence of enterprises that have set up forms of collaboration with other actors (particularly relevant in the regions of Calabria, Sardegna, Puglia, Campania and Molise) can also be observed in these territories. Networking also characterises the nonprofit sector, consisting of institutions that involve their beneficiaries in the design and implementation of the various initiatives and nonprofit institutions providing services aimed at the community.

4. Final remarks and future steps

Although the results described come from an exploratory analysis, some interesting issues can be highlighted. The multi-actor aspect of social innovation: we can thus confirm the existence of a virtuous link between factors of social cohesion - fostered by the presence of a high number of nonprofit institutions focusing on vulnerability - and the presence of for-profit sector oriented to community wellbeing and to environmental protection (Venturi and Santuari, 2023). Therefore, it is necessary to study social innovation through the interdependence between economic

and institutional actors. A need to produce new statistics that are able to describe more accurately the social innovation in connection with the quality of life, especially in terms of resilient responses during crisis period, emerged from the analysis carried out. For this purpose, a new section was included in the questionnaire of the Permanent Census of Non-Profit Institutions 2021 in order to collect information about social innovation projects or initiatives carried out by NPI, i.e. the constituent elements of the project, the partnerships set up, as well as the territorial level of the projects. These data will allow to deepen the social innovation analysis by applying new statistics in a perspective of territorial analysis and comparability trough time.

References

BUZAN T. 1995. *The mindmap book*. London: BBC Books.

MOULAERT F., VAN DER BROECK P., MANGANELLI A. 2017. Innovazione sociale e sviluppo territoriale, *Rivista Impresa sociale,* No.10 dicembre 2017.

MULGAN G. 2007. Social Innovation: What it is, why it matters and how it can be accelerated. *Working Paper Skoll Centre for Social Entrepreneurship*.

MURRAY R., CAULIER-GRICE J., MULGAN G. 2010. *The open book of social innovation*. Social Innovator Series: Ways to design, develop and grow social innovation. London: The Young Foundation.

VENTURI P., SANTUARI A. 2023. Co-progettazione e valutazione d'impatto sociale. Profili giuridici e prospettive per una nuova stagione di politiche orientate all'innovazione sociale. In BONAGA, ECCHIA, PRANDINI, VENTURI. *Finanza d'impatto sociale*, Milano: Franco Angeli.

Stefania DELLA QUEVA, Istat, dellaqueva@istat.it
Manuela NICOSIA, Istat, mnicosia@istat.it
Sabrina STOPPIELLO, Istat, stoppiel@istat.it

WHAT TOOLS FOR A GOVERNANCE OF AI SOLUTIONS TO SUPPORT OFFICIAL STATISTICS?[1]

Massimo Fedeli, Sara Letardi, Simona Pace

Abstract. Artificial Intelligence (AI) refers to the ability of computers to perform repetitive tasks that would usually require the presence of human intelligence. AI can become a powerful tool in support of Official Statistics, innovating traditional approaches to statistical information production processes throughout the entire data life cycle. To fully exploit the potential of AI initiatives, it is not enough to adopt advanced technology solutions but it becomes strategic to integrate them into a reliable, transparent and consistent framework.
While the speed at which AI solutions are being developed augurs an increasingly rich future of opportunities, it must also find producers of official statistics ready not only to embrace the innovative stimuli but also to know how to guide and govern them to take into account the ethical demands that such a challenge entails. This translates into the need to adequately increase the skills required in the various areas affected by AI while stimulating a reflection on how to equip oneself with a standardized approach and processes to develop AI-based solutions that are in line with the agency's strategic objectives.
Istat has activated several experimental projects for the use of artificial intelligence to support the production of statistical information.
In parallel, through dedicated workshops, training moments and monitoring of the initiatives, it is pursuing a path to build a specific framework for the introduction of AI in the Institute, with the aim of aligning the ethical priorities of AI, ensuring its inclusiveness and sustainability.
In this paper, we present the Institute's experience in the conception, development, and early AI-based use cases, highlighting its challenges and most significant milestones.
We also illustrate how, through the integration of use cases into a reliable and effective monitoring and governance ecosystem, AI applications can be overseen to mitigate their risks without reducing their innovative scope.

1. Introduction

In line with the three-year plan proposed by the Agency for Digital Italy for information technology in public administration (AgID) 2024-2026 (AgID, 2023), the National Institute of Statistics, Istat, has started to adopt Artificial Intelligence

[1] This article is the joint work of the authors, however paragraph 1 is written by Simona Pace, paragraphs 2, 3 and 4.1, by Sara Letardi, paragraph 5 is written by Massimo Fedeli.

(AI) in order to promote innovation in methods and technologies for the production and dissemination of statistical information and the increase of administrative efficiency while respecting the principles of transparency, reliability and sustainability.

In general, public administrations (PAs), with regard to their role in the country's digitization process and their task of protecting the public interest face particular challenges in both the management and proposed regulation of artificial intelligence. Indeed, on the one hand, they play the role of a user, adopting AI technologies to more efficiently perform tasks related to their institutional mission; on the other hand, they directly support the technology, such as through the creation of infrastructure services or joining research programs or through the imposition of procurement criteria (Djefall, 2022).

In its role as a public research organisation, Istat therefore adopts a responsible approach to managing the deployment path of the AI solutions it intends to develop or acquire. AI, if implemented without proper controls and assessments, potentially entails several risks, for example related to the way decisions are processed, privacy violations or security issues.

The main tool to face this new challenge is the creation of a reliable, transparent ecosystem in line with international regulations and best practices, i.e. an AI governance framework that is in line with Istat's own characteristics and values.

In the literature, a framework is defined as a set of tools, rules, processes, procedures and values aimed at ensuring regulatory compliance and alignment with ethical values of the AI solutions to be adopted (Mäntymäki et al., 2022).

Following a path of innovation, the Institute has activated several experimental projects on the use of artificial intelligence to support the production of statistical information. In this context, it is pursuing the construction of a specific governance framework for the introduction of AI in the Institute.

The aim of this article is to outline the path for the construction of the framework, starting with the context in which it is to be developed, defining its scope of action and outlining how it will be activated in the Institute, as well as possible future developments.

The construction of the framework within the Institute will be as "agile" as possible, providing for constant and continuous updates to keep pace with the rapid technological developments in AI and the evolution of national and international regulations.

By integrating planned or developing use cases into a reliable and effective monitoring and governance ecosystem, AI applications can then be overseen to mitigate their risks without stifling their innovative scope.

2. Opportunities and challenges for official statistics

Several definitions of Artificial Intelligence can be found in the literature, in this paper we will refer to the definition contained in the AI ACT, published in the European Official Journal on 12 July 2024, at Art. 3, point 1:

" AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments" (European Parliament, Council of the European Union, 2024).

For example, AI makes it possible to analyse and process large amounts of data in real time, automate complex processes, improve operational efficiency and ultimately make more informed and timely decisions. In addition, by assisting researchers in their work, AI can bring innovation to the entire process of statistical data production, from data collection, through respondent support tools such as dedicated chat bots, to data dissemination, such as semantic searches of statistical information or the use of virtual assistants to respond to user queries (HLG-MOS White Paper, 2023). In particular, Large Language Models (LLMs), with their ability to automatically sort textual data into predefined categories, can be used to code and classify collected information under the supervision of researchers, integrating both current methodologies and accumulated experience in organisations. Or they can help to simplify data cleaning and pre-processing activities by identifying and correcting errors, missing values or inconsistencies (HLG-MOS White Paper, 2023).

Finally, as part of the Italian AI strategy, it is planned to promote the definition of approaches for generating synthetic datasets that can be used in specific application contexts. It will be essential to support the use of Privacy Enhancing Technologies (PETs), i.e. digital solutions that allow the collection, processing, analysis and sharing of information while protecting the confidentiality and privacy of the data, thus encouraging participation (AgID, 2024).

At the same time, the adoption of AI raises specific ethical issues (Benanti, 2022)

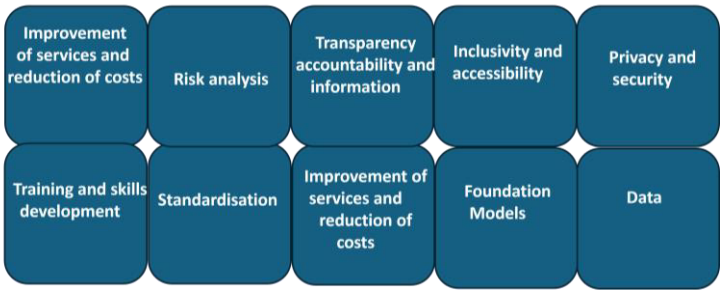
different stakeholders throughout the cycle of an AI solution, be it a product or a service (De Almeida et al., 2021).

In general, trustworthy AI relies on three essential and closely related elements: legality, which ensures consistency with applicable regulations; ethics, which ensures adherence to ethical values and principles; and robustness, which involves careful monitoring and mitigation of the risk of causing intentional or unintentional harm (AI HLEG, 2019).

3. Tools and methods for AI governance.

With the publication of the Artificial Intelligence Regulation (AI Regulation) in the Official Journal of the European Union on 12 July 2024, Europe has set the course to ensure that AI in the European context is trustworthy and respects fundamental rights and the basic principles of democracy and the rule of law. Specifically, the Regulation sets out a number of principles and regulatory solutions that will have an impact in public administration in several areas, from respect for fundamental rights to ethical principles, from methods of legal risk analysis to the need for new skills and organisational models (Fondazione ICSC, 2024).

Figure 1 - *AgID Decalogue, abstract.*



As mentioned in the introduction, currently in Italy, the AgID Three-Year Plan for Information Technology in Public Administration 2024-2026 outlines a general framework for the introduction of AI in public administrations and, in the light of the European regulation, has been supplemented with a decalogue of points of attention[2] for the adoption of AI in PA, a brief summary of which we present in Figure 1.

In addition to the regulatory framework, the development and adoption of AI solutions must necessarily take into account both ethical principles, which are not always promptly translated into legal norms, and the interests of users and, more generally, the requirements expressed by stakeholders.

The set of tools, rules, processes, procedures and values that collectively aim to ensure ongoing alignment with ethical principles, applicable legislation and, more generally, an organisation's own values and goals is called AI governance (Mäntymäki *et al*., 2022). However, although the importance of AI governance has

[2]Cfr. https://www.agid.gov.it/sites/default/files/repository_files/documentazione/Degalogo%20AgID%20IA%20nella%20PA.pdf

been widely emphasised, there are relatively few practical systems for understanding and implementing these principles, i.e. defined AI governance frameworks, including guidelines and best practices for organisations (Mäntymäki *et al*., 2022).

In order to develop a framework at the Institute, we analysed what exists in the international landscape and drew on the experience of other international public institutes and private IT or IT consulting companies.

4. Istat's experience

The definition of the governance of AI applications in the Institute is based on the analysis of the best practices present in the national and international landscape, taking into account the indications of the AgID and integrating the different methodologies used in both the public and private spheres, adapting them to the specific context of a research institution such as the National Institute of Statistics. The aim is to provide practical guidelines and organisational tools for the Institute's units that develop, apply or use artificial intelligence technologies.

From the comparison with the elements described, three pillars have been identified which, taken together, form the backbone of AI governance at Istat:

- Innovation
- Data and rules
- People

Each pillar, as illustrated below, is in turn made up of a series of activities that define and characterise it and which, as a whole, translate the ethical principles and standards that inspire them into lines of action through the identification of specific objectives.

The pillars influence each other through mutual interaction and integration, facilitating communication flows between the various structures involved. For example, regulatory compliance involves the implementation of the "rules" defined in the development of AI solutions throughout their lifecycle. The adoption of new technologies to bring about "innovation" is based on the growth of skills and continuous professional updating, i.e. the training of "people", whether they are developers or users of the applications adopted.

In order to illustrate the main points of the framework to be set up, the following is an illustrative and non-exhaustive summary of the actions identified to implement the AgID indications contained in the Decalogue, shown in Table 1.

4.1. Innovation

The first pillar, Innovation, is characterised by the analysis and evaluation of existing and developing AI applications at the Institute and the definition of tools to ensure that AI solutions are developed, commissioned and monitored in accordance with AgID recommendations and the Institute's values and strategic objectives..

Research and advisory firm Gartner (Ramos *et al.*, 2023), has observed that in the early stages of testing generative AI, organisations focus their attention on technical pilots that deliver incremental improvements, making it difficult to identify and prioritise the potentially most impactful generative AI use cases. Instead, there is a need to manage the governance of requests to implement new AI solutions and identify which use cases are prioritised over others.

In the use case discovery phase, AI solution proposals are identified and catalogued. Then, for each use case, the teams involved, the objectives, the timeframe for implementation, the data used to develop or deploy the AI solution, and any costs to be incurred are identified. It also identifies any risks and/or criticalities that have been or may be encountered in the course of development or deployment, and what mitigation measures can be put in place in relation to the risks and criticalities identified. Finally, the indicators that will be used to measure the effectiveness of the results obtained will be defined.

In summary, based on the analysis of the needs expressed and the innovation plan of the Institute, AI use cases are analysed in terms of defining the objectives, activities, data and technologies used in AI applications.

The path of identification, monitoring and evaluation of the use cases is in line with the indications provided by AgID for the governance of AI, as summarised below:

Table 1 – *Innovation.*

AgID Indication	Requirement	Actions
Improvement of services and cost reduction	Approve new use cases against criteria shared with management	Define and apply priority criteria for use case adoption
Improvement of services and cost reduction	Control and monitor approved use cases	Create and update the Use Case Catalogue
Sustainability	Ensure the sustainable use of resources that are dedicated to AI solutions	Promote the reuse of IA solutions

As an example, here is an extract from the catalogue of use cases used at the Institute for the exploratory phase of approved projects.

Table 2 – *Catalogue of use cases in Istat – extract.*

Use cases	Goals	Benefits
Single point of contact	Implement a generative AI solution within the Single Point of Contact.	Improve the quality of information provided to users in terms of consistency and accuracy. Improve user satisfaction by simplifying access to the system. Improve the Institute's reputation: create a strong identity image, more modern and in tune with the times.
Semantic search engine	Implement a search engine on the new institutional website.	Ability to respond more quickly and accurately to internal and external requests, identification of previously unknown document insights. All this will improve the quality of the service provided to users.
Experimenting with the use of ia to access istat data	Definition of a solution that supports users in the guided navigation of IstatData, through the use of generative AI.	The project aims at enhancing the information assets produced by Istat through the controlled use of a solution based on generative AI. In this project, the main challenge is to produce results that are controlled by avoiding that in the prompting phase the AI algorithms produce uncontrolled effects.

4.2. Data and Rules

The second pillar, Data and Rules, aims to identify and define the processes, procedures and guidelines needed to implement the ethical and regulatory principles of AI governance. Particular attention will be paid to how the data needed for AI applications is collected, manipulated and subsequently disseminated.

In this context, an operational model of AI governance will be defined, identifying roles and responsibilities at both project and organisational levels.

After analysing the recently published AI ACT regulation, the risk level of AI applications will be assessed and possible mitigation measures will be identified. Finally, continuous monitoring of activities and evaluation of results will ensure various aspects, such as alignment with innovation objectives or compliance with regulations, and seize opportunities for continuous improvement, using the various governance tools developed.

For this purpose, it is necessary to continuously collect and update project documentation, in particular for those aspects that concern ethical issues, such as the

data and models used. The use of templates such as scorecards or other *ad hoc* tools makes it possible to monitor the lifecycle of AI solutions and clarify the frame of reference in which they are developed and/or used (Lu *et al*., 2024).

With reference to the AgID indications, several lines of action have been envisaged, by way of example:

Table 3 – *Data and Rules.*

AgID Indication	Requirement	Actions
Risk analysis	Identifying and mitigating risks associated with the use of AI	Defining and constantly updating risk assessment forms.
Transparency, accountability and information	Ensuring compliance with regulations	Recognition and monitoring of project documentation for AI solutions Definition of roles and responsibilities and processes
Privacy and security	Ensure users privacy	Establish training courses on the responsible use of AI

One of the first processes defined in this area concerns the adoption of new use cases, which can be summarised as follows:

- Indication of the strategic objectives or needs of the Institute.
- Creation of a transversal working group
- Experimentation of a Proof of Concept (POC) that includes the minimum functionality and technology required to validate the case.
- Comparison with top management for possible approval and allocation of resources for the development of the POC
- POC implementation
- Evaluation of the results obtained by top management and possible approval of the POC to go into production.
- Go to production

As a first step towards monitoring and managing the growing number of solution projects and the resulting AI systems, a transversal organisational structure has been put in place at the Institute, in line with international best practice. The aim is to monitor and link the initiatives of the various operational structures, ensuring alignment with institutional objectives and working closely with the IT and methodology teams for data management and the integration of AI into applications, with the communication teams to promote cultural growth, and with the legal teams to ensure compliance with the relevant regulations.

4.3. People

The third pillar, People, is dedicated to those within the Institute who use AI applications or are involved in the development of innovative solutions.

A fundamental element in building an AI framework is the design of actions to promote training and cultural growth to develop the necessary skills and awareness regarding the use and development of AI solutions.

On the one hand, training activities will be generalist in nature, to bring non-specialist personnel closer to AI issues and their responsible use, thus aiming at training future AI users. On the other hand, they will be aimed at a smaller and more experienced contingent, with the aim of enhancing and consolidating the skills of those who will be involved in the design and implementation of AI solutions.

At the same time, streams of communication with the outside world will be activated in order to strengthen the positioning of the organisation with its stakeholders and contribute to the cultural growth of the country.

According to AgID, the following are planned in this area:

Table 4 – *People.*

AgID Indication	Requirement	Actions
Training and skills development	Promote AI literacy to all staff	Set up courses, seminars Disseminate information material. Assess existing skills and align them with those required by AI.
Training and skills development	Promote Istat's positioning towards stakeholders. Participate in the cultural growth to promote a conscious use of AI systems	Participation in dissemination events on AI; activation of internal communication lines.

At Istat, more than 20 hours of training were provided to staff, both specialised and non-specialised, on various AI-related topics, both general and more specialised. AI ACT refresher courses and participation in national and international events are planned.

The 15th National Statistics Conference, held on 3-4 July 2024, was dedicated to the topic "Official Statistics in the Age of Artificial Intelligence[3]. In addition, an area dedicated to artificial intelligence has recently been activated on the Institute's intranet with the aim of disseminating information on artificial intelligence, illustrating its potential for the Institute and presenting ongoing use cases. This area

[3] https://www.istat.it/en/event/fifteenth-national-conference-of-statistics/

is a work in progress and will be updated as experiments and training events are added. To complete the picture, links to European and national legislation will be collected, as well as AgID indications, and finally, events and talks dedicated to the topic will be reported.

5. Conclusions and outlook

Several experimental projects will be started at the Institute, including participation in research programmes related to the development of the Italian LLM, through agreements with public bodies, universities, and private entities, in line with the Italian Artificial Intelligence Strategy (AgID, 2024). The framework will therefore be activated in the field and in "agile" mode, in order to seize opportunities for continuous improvement and to adapt dynamically to the different application areas of AI solutions.

The next step is to achieve ISO/IEC 42001 - ARTIFICIAL INTELLIGENCE MANAGEMENT SYSTEM CERTIFICATION, which sets out the requirements for establishing, implementing, maintaining and continually improving an artificial intelligence management system.

AI governance is not just about ensuring one-off regulatory compliance, but must maintain ethical standards over time. As such, it is a system that, by its nature, is constantly evolving to keep fe with technological advances and regulatory developments. It is therefore itself subject to continuous monitoring and evaluation to ensure that it is effective, adaptable and in line with emerging objectives and needs, and that it promotes ethical and responsible digital innovation.

References

AGID 2023. *Piano Triennale per l'Informatica nella Pubblica Amministrazione 2024-2026.*

AGID 2024. *Strategia Italiana per l'Intelligenza Artificiale 2024-2026.*

AI-HLEG HIGH-LEVEL EXPERT GROUP ON AI, 2019. Ethics guidelines for trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

BENANTI P. 2022. *Human in the loop. Decisioni umane e intelligenze artificiali.* Milano: Mondadori.

DJEFFAL C. 2022. Intelligenza artificiale e pubblica amministrazione in Germania. Riflessioni sui principi orientativi di carattere normativo a partire dall'esperienza

tedesca. *Teoria e Critica della Regolazione Sociale / Theory and Criticism of Social Regulation,* Vol 2, No 23, pp. 179-197

DE ALMEIDA P. G. R., DOS SANTOS C. D., FARIAS J. S. 2021. Artificial Intelligence Regulation: a framework for governance. *Ethics Inf Technol*, Vol. 23, pp. 505–525.

EUROPEAN PARLIAMENT, COUNCIL OF THE EUROPEAN UNION 2024. Regulation (eu) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

FONDAZIONE ICSC CENTRO NAZIONALE DI RICERCA IN HIGH PERFORMANCE COMPUTING, BIG DATA E QUANTUM COMPUTING - SOCIETAL IMPLICATIONS AND IMPACT RESEARCH GROUP - OSSERVATORIO. 2024. *AI Act: Raccomandazioni per un uso responsabile ed etico dell'Intelligenza Artificiale nella Pubblica Amministrazione*. https://www.supercomputing-icsc.it/wp-content/uploads/2024/05/AI_PA_Raccomandazioni_13maggio2024.pdf

GARTNER 2023. *How to Pilot Generative AI*. https://www.gartner.com/en/doc/797246-how-to-pilot-generative-ai

HLG-MOS White Paper 2023, Large Language Models for Official Statistics. Modernstats

LU Q., ZHU L. , XU X., WHITTLE J., ZOWGHI D., JACQUET A. 2024. Responsible AI Pattern Catalogue: A Collection of Best Practices for AI Governance and Engineering. *ACM Computing Surveys*. Vol. 56, No. 7, pp. 1-35.

MÄNTYMÄKI M., MINKKINEN M., BIRKSTEDT T., VILJANEN M. 2022. Putting AI Ethics into Practice: The Hourglass Model of Organizational AI Governance. arXiv:2206.00335

Massimo FEDELI, Istat, massimo.fedeli@istat.it
Sara LETARDI, Istat, sara.letardi@istat.it
Simona PACE, Istat, simona.pace@istat.it

TRAINING MODEL AND INTERVIEWER PERFORMANCE IN THE PERMANENT POPULATION CENSUS[1]

Nadia Mirante, Antonella Bernardini, Nunzia Balì

Abstract. Training interviewers is crucial in survey design to minimize interviewer effects, which impact non-sampling errors (partial non-response, refusals, and measurement errors) and data quality. Istat has developed expertise in training interviewer networks and continuously works on strategies to improve and harmonize its training model, adapting to societal changes and respondents' lifestyles.

This study examines the effect of interviewer training on survey performance to improve the training program. The data analyzed relate to the performance of 5,894 interviewers trained for the 2023 Italian Population Census "List" survey.

To explore the relationship between training and fieldwork performance and identify interviewer profiles, we used a statistical classification method (decision tree), considering socio-demographic and geographical factors as potential covariates.

The results represent a crucial step toward the ongoing use of statistical methods to enhance the data collection process.

1. Introduction

The quality of produced data depends significantly on the quality of the data production process. In recent decades, awareness of this connection has solidified, leading to greater attention on both upstream (design) and downstream (processing and analysis) data quality.

In the realm of Official Statistics, Eurostat has outlined quality principles for both the statistical process and the statistical product (Eurostat, 2017). These principles have been adopted by Italy through the Codice Italiano per la qualità delle Statistiche Ufficiali (Istat, 2022), recognized as a best practice for the production of Official Statistics internationally. This is complemented by the ESS Quality Assurance Framework, or ESS QAF (Eurostat, 2019), which identifies methods and tools at the institutional or process level to ensure compliance with the Code's principles, thereby promoting good practices. While the Code's principles tell us *what to do*,

[1] The authors share contents and views expressed in this paper. However N. Mirante drew up the sections 1 and 2; A. Bernardini the section 3 and N. Balì the section 4.

the GSBPM (Generic Statistical Business Process Model) guides us on *how to do it*. Developed by UNECE (2019), the GSBPM describes and defines the set of processes necessary for the production of official statistics, providing a standard framework and harmonized terminology to assist statistical agencies in modernizing their statistical production processes and sharing the methods used. Additionally, it serves as a model for evaluating and improving the quality of these processes.

Since the 1990s, Istat has systematically worked to ensure the quality of statistical information and the service provided to the community. This has involved adopting the aforementioned principles and models and drafting various quality guidelines (Istat, 2012; Istat, 2018) to illustrate the principles to be followed in the design, execution, and control of a statistical survey, as well as to describe the quality requirements that statistics must meet.

These guidelines also include principles concerning interviewers and their training, specifically:

- The principle "D.2.4. Interviewers" from the Istat guidelines for the quality of statistical processes states: "*Interviewers must be selected, trained, and monitored in such a way as to ensure that the data collected are as accurate* as *possible"* (Istat, 2012);
- The principle "C.2." from the SISTAN quality guidelines for statistics states: "*To ensure the quality and completeness of the information collected, careful attention should be paid to the selection and training of interviewers. Additionally, the data collection phase should be monitored during the process and evaluated afterward using appropriate tools and objective indicators"* (Istat, 2018). This principle identifies two monitoring indicators for interviewers:
    - Average daily number of interviews per interviewer (workload);
    - Total response rate per interviewer (number of completed interviews out of those planned).

Among the various aspects to consider in the design and execution of a direct statistical survey, one of the most important is the training of interviewers. Proper training is crucial to minimize the interviewer effect, which significantly impacts non-sampling error (e.g., partial non-response, refusals, and measurement error), and thus the quality of collected data. However, interviewer training is often an overlooked aspect in reducing interviewer effects in interviewer-administered surveys (Daikeler and Bosnjak, 2020).

Istat has extensive experience in training survey networks, which has undergone various modifications over the years in a continuous search for new training strategies to improve and harmonize the proposed training model, keeping pace with societal changes and respondents' lifestyles, thereby generating economies of scale.

Istat has consistently placed great emphasis on education, implementing new strategies to ensure a continuous and efficient training model, especially during the transition to the new census framework. This approach takes into account economic constraints and available resources. Significant innovations have been introduced in the training strategy, resulting from a thorough examination and analysis of various training processes used within the Institute. The goal is to design a permanent training model with a circular perspective.

Initially, training followed a waterfall model (e.g., central Istat, territorial Istat, municipalities). With Istat's modernization, training became direct and centralized, allowing interviewers to be trained directly by Istat. This resulted in the standardization of the training process and control over its effectiveness and efficiency. Another significant change introduced in the training of survey networks was the shift from on-site training to remote or blended training, including self-learning modules (FAD) [Istat, 2019]. Finally, during the COVID-19 pandemic, it was necessary to rely entirely on remote training.

Over the decades, Istat has developed substantial expertise in training survey networks. This knowledge is supported by various studies conducted by the Institute over the years to evaluate the effectiveness and efficiency of the different training models implemented (Istat, 2005; Istat, 2006; Istat, 2010; Balì and Federici, 2014; Balì, 2015; Balì *et al.*, 2023).

The importance of information produced by official statistics has led to constant attention to data quality in all survey processes, particularly in the Census. The Census process is characterized by complexity and multidisciplinarity, requiring diverse competencies. A significant contribution to the quality of collected data comes from individuals involved in survey networks, emphasizing the central role of network training.

To address this, the Institute has heavily invested in training survey networks, creating well-structured training programs. Notably, the redesign of network training was part of the innovative approach to the new Census strategy, incorporating interdisciplinary methods. The adoption of blended learning - an ongoing learning approach involving cross-cutting skills - was chosen.

Furthermore, the Covid-19 pandemic accelerated the introduction of innovative elements in training delivery. Traditional in-person training sessions shifted to virtual classrooms, utilizing communication and collaboration platforms. This adaptation highlighted the need for flexibility and adaptability in the training approach.

The aim of our study is to explore how the training program can influence interviewers' performance in the field and to try to profile the surveyors.

2. Methods

For this study, we utilized interviewer performance data from the survey network for the 2023 edition of the "permanent population and housing census – Lista".

Before the fieldwork, interviewers were required to complete an online self-learning course (Formazione a Distanza, FaD) on Moodle2. The FaD comprised 14 consecutive video lessons. Upon viewing each video lesson, the student had to pass a corresponding intermediate test. Only after successfully passing this test was the interviewer allowed access to the subsequent video lesson and its associated intermediate test. This process continued sequentially until all modules were completed. At the end of this educational path, the student was granted access to the final test, which required a minimum score of 90% for successful completion.

On the Moodle platform, interviewers also had access to additional materials such as manuals and guides. These resources were available throughout the survey period to support the interviewers when necessary.

We analyzed the data of 5,894 interviewers who met the following inclusion criteria: 1) at least one family assigned; 2) at least one interview conducted.

To investigate interviewer performance, we focused on two specific phases: 1) performance during training; 2) performance during fieldwork. The chosen parameters were "Score in the post-training learning test" and "Number of attempts to pass the test" for the training phase, and "Number of face-to-face interviews out of the total number of interviews conducted" for the fieldwork phase.

Other factors considered were: age class (18-29 years, 30-44 years, 45-59 years, and over 59 years), gender, educational level (middle school, high school, degree), geographical data (North, Centre, South), new/expert interviewer status, and whether the FaD was completed before or after the fieldwork.

To classify this population, we used the CHAID (Chi-squared Automatic Interaction Detection) tree classifier, which employs the chi-square association statistic to define, at each level, how to subdivide cases into subgroups, starting from the entire study dataset (Ritschard, 2013). This non-parametric method selects the independent variable (predictor) with the strongest interaction with the dependent variable at each step. Categories of each variable are merged if they do not differ significantly from the dependent variable. At each node, the algorithm splits the cases based on rules for different categories of independent variables, so the resulting subsets are most associated with the performance variable.

[2] Modular Object-Oriented Dynamic Learning Environment (Moodle) is a computer system for creating and sharing educational materials online

3. Results

The classification tree is an extremely useful tool for presenting the results of a recognition model for specific profiles, as it translates these results into easily understandable rules, unlike the often complex parameters of other predictive models. The dependent variable we chose is "the number of face-to-face interviews out of the total number of interviews conducted", which we consider a measure of the interviewer's performance.

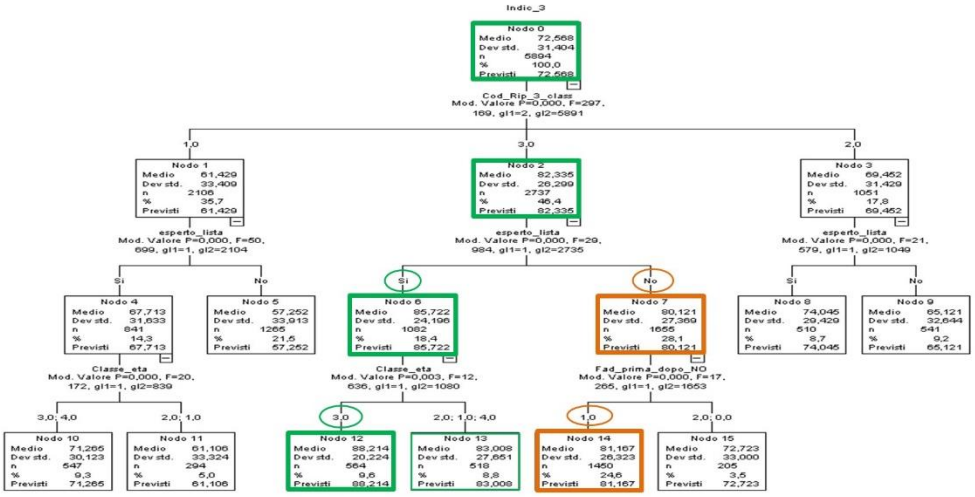
We hypothesized that certain contextual factors - geographical and spatial in nature, demo-social characteristics such as age and educational qualification, and the training behavior of the survey network members - might influence the interviewer's performance in the field. Our goal was to identify interviewer 'profiles' that could contribute to more efficient survey networks. The independent variables we considered included the interviewer's socio-demographic variables (age, gender, and educational qualification), census survey experience, and training-related variables (completion of the training course within the required timeframe, overall grade in the final test of the training course, and the number of attempts to pass the final test).

Although this first application does not provide a perfect classification of the interviewers, nor does it describe a precise profile or fully explain the relationship between training and performance, the tree is still very useful for describing individual and territorial profiles associated with different training characteristics and interviewer behaviors during fieldwork.

The results of the tree show that the first subdivision rule corresponds to territory. In the application, three levels of the territorial type variable (geographical breakdown, region, and province) were considered as independent variables. For easier interpretation, Figure 1 shows the analysis conducted with the most aggregated variable (geographical breakdown).

This results in a total of 16 nodes, with 9 terminal nodes representing the best classifications for the chosen model. At the first level, geographical breakdown constitutes the best classification step. Moving to the next level, for all breakdowns, the discriminating factor is the variable 'experience'. If the interviewer is experienced and falls within the middle age range, their field performance is the best (Node 12). Together with the younger and older age groups (Node 13), they account for 18% of the interviewers. This is crucial because a survey network with extensive Istat survey experience certainly ensures better data quality. Conversely, if there is no survey network with experience, the model shows that training is fundamental to the interviewer's work. Node 14, which includes about 25% of the observed population, represents the cluster of units that, despite having no work experience, completed the training before conducting interviews in the field.

Figure 1 - *Performance during the fieldwork: classification tree result*



The work will continue to identify the best reclassification strategy for the variables, focusing on the two phases under consideration: performance during training and performance during fieldwork.

4. Conclusions

In interviewer-administered surveys, the interviewer's role is pivotal in ensuring the collection of high-quality data. While appropriate training is a fundamental factor in enhancing field performance, our study reveals that additional elements - such as socio-demographic characteristics of interviewers, prior census survey experience, and geographical context - play a significant role in shaping outcomes.

The decision tree analysis underscored the importance of territorial differences: interviewers operating in Southern Italy outperformed their counterparts in the North and Centre. This finding could reflect regional disparities in engagement strategies, cultural attitudes toward surveys, or broader socio-economic characteristics of the interviewers. Moreover, prior experience as a census interviewer emerged as a critical determinant of achieving high rates of face-to-face interviews, highlighting the value of maintaining an experienced network of interviewers.

For new interviewers, the results indicated that completing training prior to the commencement of fieldwork was strongly associated with improved performance.

This underscores the importance of well-structured and mandatory training programs. Consequently, the evidence suggests that training should be tailored to the needs of different interviewer profiles, with comprehensive programs for new recruits and refresher courses for seasoned professionals.

Nonetheless, the findings also point to the necessity of more granular analyses to better comprehend and address performance disparities. As previous research (e.g., Bernardini et al., 2014) has noted, challenges in reaching households vary significantly even within broadly aggregated territorial classifications. Municipal-level complexities - including geographical size, population density, and socio-economic conditions - substantially influence the data collection process. In this context, the introduction of a composite index, as proposed by Bontempi et al. (2024), represents a promising advancement. Such an index could synthesize interviewer socio-demographic attributes, territorial characteristics, and training behaviors, enabling more precise profiling and targeted interventions.

This study represents a critical step toward optimizing interviewer training and resource allocation. Developing differentiated training modules based on interviewer profiles, alongside strategies tailored to specific territorial contexts, offers a promising pathway for enhancing the efficiency of survey networks and, ultimately, the overall quality of collected data.

Acknowledgements

We would like to extend our gratitude to Linda Porciani and Antonio Pitrone for their contribution to data collection. A special thanks to Matteo Mazziotta for his insightful feedback, which significantly improved this publication.

References

BALÌ N. 2015. La formazione dei rilevatori. In ISTAT (Eds.) *La nuova indagine sulle spese per consumi in Italia*. Istat, pp. 70-78 On https://www.istat.it/it/files//2016/03/Indagine-spese-per-consumi.pdf.

BALÌ N., CECCARELLI C., FIORI M.T., LUGLI A., ROSSETTI F. 2023. Impact of digital learning on the interviewer's performance, *RIEDS. Rivista Italiana di Economia Demografia e Statistica* Vol. 77, No.1, pp 65-76.

BALÌ N., FEDERICI A. 2014. Le strategie formative degli attori coinvolti nel lavoro sul campo. In FREGUJA C., ROMANO M.C. (Eds.) *La modernizzazione delle tecniche di rilevazione nelle indagini socio-economiche sulle famiglie*. Istat, pp. 91-100. On http://www.istat.it/it/archivio/145721.

BERNARDINI A., FASULO A., TERRIBILI M.D. 2014. A Model Based Categorisation of the Italian Municipalities Based on Non-Response Propensity in the 2011 Census, *RIEDS. Rivista Italiana di Economia Demografia e Statistica*, Vol. 68, No. 3, pp. 79-86.

BONTEMPI K., PIETROPAOLI S. 2024. Composite indices for measuring the "complexity of data collection" in Italian municipalities, *RIEDS. Rivista Italiana di Economia Demografia e Statistica*, Vol.78, No.1, pp. 195-206.

DAIKELER J., BOSNJAK M. 2020. Chapter 4: How to Conduct Effective Interviewer Training: A Meta-Analysis and Systematic Review. In OLSON K., SMYTH J. D., DYKEMA J., HOLBROOK A. L., KREUTER F, WEST B. T. 2020. *Interviewer Effects from a Total Survey Error Perspective*. Boca Raton: CRC Press.On https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1001&context=sociw

EUROSTAT. 2019. Quality Assurance Framework in the European Statistical System. On https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646

EUROSTAT. 2017. European Statistics Code of Practice revised edition 2017. On https://ec.europa.eu/eurostat/web/products-catalogues/-/ks-02-18-142

ISTAT. 2022. Codice italiano delle statistiche ufficiali. On https://www.sistan.it/fileadmin/redazioni/IMMAGINI/pdf/Codice-Italiano-Qualita__-Statistiche-Ufficiali.pdf

UNECE. 2019. Generic Statistical Business Process Model (GSBPM) (Ver 5.1). On https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1

ISTAT. 2019. L'approccio trasversale alla formazione delle reti di rilevazione. Letture Statistiche. Metodi. On https://www.istat.it/it/files//2020/05/approccio-trasversale-alla-formazione-delle-reti-di-rilevazione-2020.pdf

ISTAT. 2018. Linee Guida per la Qualità delle Statistiche del Sistema Statistico Nazionale. On https://www.istat.it/it/files/2018/08/Linee-Guida-2.5-agosto-2018.pdf.

ISTAT. 2012. Linee guida per la qualità dei processi statistici. Rome. On http://www.istat.it/it/files/2010/09/Linee-Guida-Qualit%C3%A0-_v.1.1_IT.pdf

ISTAT. 2010. Dalla "riunione di istruzione" all' e-learning 2.0. Un progetto di formazione continua "in rete per le reti" di rilevazione dell'Istat. Internal document.

ISTAT. 2006. Il sistema di indagini sociali multiscopo. Contenuti e metodologia delle indagini. Metodi e norme, 31. On https://www.istat.it/it/files/2014/06/met_-norme_06_31_il_sistema_di_indagini_multiscopo.pdf.

ISTAT. 2005. La rete di rilevazione Capi dell'Istat per la conduzione dell'indagine continua sulle Forze di Lavoro. Metodi e norme, 24. On

https://ebiblio.istat.it/digibib/Metodi%20e%20norme/IST0047577Ed2005N24.pdf

RITSCHARD G., 2013. CHAID and Earlier Supervised Tree Methods. In MCARDLE J.J., G. RITSCHARD (Eds.) *Contemporary Issues in Exploratory Data Mining in the Behavioral Sciences*, New York: Routledge, pp. 48-74.

Nadia MIRANTE, Italian National Institute of Statistics, nadia.mirante@istat.it
Antonella BERNARDINI, Italian National Institute of Statistics, anbernar@istat.it
Nunzia BALÌ, Italian National Institute of Statistics, bali@istat.it

SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS
ISSN ASSEGNATO: 0035-6832

Direttore Responsabile: 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

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).

LVIII Tra marginalità e sviluppo. La sfida della sostenibilità in una prospettiva mediterranea (Università LUMSA, Palermo, 26-27 maggio 2022).

LIX Aspetti economici e sociali dell'invecchiamento demografico (Università degli Studi di Napoli Federico II Napoli, 25-26 maggio 2023).

LX Salute, ambiente e disuguaglianze: istituzioni, imprese e società (Università degli studi di Milano – Bicocca & LIUC - Università Cattaneo, Milano & Castellanza (VA), 22-24 maggio 2024).