VOLUME LXXX – N. 2 APRILE – GIUGNO 2026

RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA

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

FRAGILITY IN ITALIAN MUNICIPAL TERRITORIES: A SPATIAL ANALYSIS BASED ON OFFICIAL STATISTICS[1]

Simona Ballabio, Alberto Vitalini

Abstract. This study analyses social fragility in Italian municipalities, aiming to provide a structural and spatial interpretation of vulnerability at the local level. Fragility is assessed through the Municipal Fragility Index (IFC), developed by Istat, a multidimensional and non-compensatory tool that integrates twelve elementary indicators related to demographic, social, economic, environmental, and territorial dimensions. Fragility is conceptualized as a lack of territorial resilience, shaped by weak human capital, limited service infrastructure, and environmental exposure. The methodological approach combines spatial analysis using Local Indicators of Spatial Association (LISA), based on Local Moran's I statistics, with unsupervised learning (k-median clustering), allowing for the identification of five distinct profiles for each dimension and a synthetic classification of municipalities into four types (T1–T4) based on cumulative fragility patterns. The study supports policy design aimed at reducing spatial inequalities.

1. Introduction and Conceptual Framework

Territorial fragility is a significant dimension for understanding inequalities in Italy. The country is indeed marked by strong spatial disparities that manifest at various levels: socio-economic, demographic, environmental, and infrastructural. These imbalances are not only historically rooted but also tend to intensify in times of crisis—be they economic, health-related, or environmental—revealing the differing capacities of territories to absorb shocks and respond to change (Benassi *et al.*, 2022; Frigerio and De Amicis, 2016; Frigerio *et al.*, 2018).

The notion of territorial fragility represents an important conceptual perspective, as it allows us to grasp the multidimensionality of vulnerabilities at the local level. Here, it is understood as a structural condition that expresses the exposure of a municipality to natural and anthropogenic risks, in combination with socio-demographic weaknesses and economic vulnerabilities. This condition can

[1] The work is the joint responsibility of the authors. Paragraph 1 and 2 is attributed to Alberto Vitalini, paragraphs 3 and 4 are attributed to Simona Ballabio.

undermine the territory's ability to ensure well-being, resilience, and sustainable development for the resident population (Istat, 2024).

Territorial fragility can be distinguished from related concepts such as resilience, marginality, and deprivation, although it shares some analytical dimensions with them. Unlike resilience, which refers to a territory’s capacity to react and adapt to shocks and changes (Mentges *et al*., 2023), fragility highlights the structural conditions that hinder such adaptation (OECD, 2022). It includes, but is not limited to, elements of marginality, such as isolation and limited accessibility to essential services. However, while marginality often denotes a peripheral condition, either spatial or relational (UVAL, 2014), fragility adopts a systemic perspective, integrating demographic, social, economic, and environmental factors. Similarly, in contrast to deprivation, which primarily concerns a lack of resources at the individual or household level, fragility operates at the collective and territorial scale, offering a useful framework for designing integrated cohesion and development policies.

Within the theoretical framework outlined above, this study aims to offer a spatial reading of municipal territorial fragility in Italy, using an index produced within the framework of official statistics that adopts an integrated methodological approach (Istat, 2024). In particular, the use of a spatial approach—through tools of geographic autocorrelation (LISA)—makes it possible to highlight relationships between neighboring municipalities and to capture phenomena of systemic fragility. The study then focuses on the most fragile municipalities in the country, in order to construct an internal typology of this subgroup and distinguish between different forms and degrees of fragility.

The reflection proposed here fits within the broader framework of analyses on territorial cohesion and inequalities, providing useful evidence for guiding intervention strategies that are capable of responding to the complexity and diversity of local situations.

2. Methodology

The analysis is based on the use of the Municipal Fragility Index (IFC), developed by Istat, which provides a synthetic measure of the structural vulnerability of Italian municipalities (Istat, 2024). The index is designed to identify territories most exposed to risks and criticalities and to support analyses that are comparable across space and time. Its structure is multidimensional and is based on twelve elementary indicators divided into two main domains: territorial-environmental and socio-economic.

2.1 Information Sources and Territorial Scope

The data underlying the index used come from official Istat sources, including: the demographic balance of the resident population; the permanent census of population and housing; territorial, environmental, and economic indicators available at the municipal level[2].

The unit of analysis is the Italian municipality, with a reference year of 2021, the most recent available at the time of the analysis. In an initial exploratory phase, a LISA spatial analysis was conducted on all Italian municipalities, aimed at identifying spatial clusters of fragility.

Subsequently, the analysis was restricted to the municipalities falling within the last three deciles (8th, 9th, and 10th) of the IFC distribution, representing the most fragile territories at the national level, totalling 2,019 municipalities.

2.2 Municipal Fragility Index (IFC)

The IFC is a non-compensatory composite index, calculated using the Adjusted Mazziotta-Pareto Index (AMPI+), which corrects the average effect of the arithmetic mean through a penalty linked to indicator variability. It assumes that the dimensions of fragility are not (or only partially) substitutable, meaning a disadvantage in one cannot be offset by an advantage in another (Mazziotta & Pareto, 2024).

All indicators are normalized with respect to the 2018 national value, set equal to 100, using a linear transformation based on specific goalposts. The final value of the index, calculated for each municipality, reflects both the average level of fragility and the internal consistency among the dimensions considered.

The two thematic areas of the index include:

- Territorial and environmental indicators:
  - Landslide hazard (percentage of municipal area at risk)
  - Incidence of protected natural areas (protected surface as a percentage of total municipal area)
  - Land consumption (percentage of urbanized land)
  - Accessibility to essential services (average travel time to services)
  - High-emission motorization (Euro 0–3 vehicles per 100 inhabitants)
  - Non-recyclable waste collection (kg per inhabitant)
- Economic and social indicators:
  - Workers in low-productivity units (percentage share in industry and services)

[2] Available at https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,Z0930TER,1.0/CFI_MUN

- Density of local production units (units per 1,000 inhabitants)
- Employment rate (age 20–64, employed over working-age population)
- Population growth rate (net migration balance)
- Adjusted demographic dependency ratio (youth + elderly relative to population aged 20–64)
- Population with low educational attainment (aged 25–64 with at most lower secondary education)

2.3 Clustering Analysis

To deepen the understanding of the different dimensions of fragility, an unsupervised classification of the most fragile municipalities was carried out through two separate clustering exercises: one focused on the territorial and environmental domain, and the other on the economic and social domain. The goal is to identify homogeneous groups of municipalities that share similar structural characteristics within each set of variables.

In both cases, the k-median clustering algorithm was used (which selects as the centroid of each cluster an actual data point: the median). This method is particularly robust in the presence of skewed distributions and outliers. It proved well-suited to the heterogeneous and highly variable nature of municipal-level territorial data.

The distance metric used to calculate similarity between municipalities was the Manhattan distance (also known as city block distance), which is more appropriate than Euclidean distance for standardized and multidimensional data. Prior to applying the algorithm, all variables were standardized to eliminate the influence of measurement scales and to ensure equal weight among the indicators.

The number of clusters was fixed in advance at 5 for each domain, based on empirical considerations, result stability, and interpretability from a policy perspective. This choice also ensured symmetry between the two readings and facilitated the subsequent construction of an integrated fragility typology.

2.4 Construction of a Fragility Typology

The interpretative and operational aim of the study required the development of a synthesis capable of integrating two distinct analytical perspectives—socio-economic and territorial-environmental—into a single typological variable. To this end, after conducting the two independent clustering analyses on the socio-economic variables (CL_soc) and the territorial-environmental variables (CL_terr), a combined four-class typology was constructed. This step was not carried out through an additional automated statistical procedure, but rather as a logical-interpretive operation based on the cross-referencing of the clusters obtained in the two domains, without assuming any implicit hierarchical order.

The intersection of the two classifications initially generated an attribute space structured as a 5x5 matrix, resulting in 25 theoretical combinations. These combinations were then aggregated into a synthetic typology, based on reduction criteria that considered the overlap or divergence of fragility dimensions, the intensity of the phenomenon, and operational clarity for policy-making purposes. This allowed a transition from the attribute space (based on the original variables to the space of interdimensional configurations (CL_terr × CL_soc), constructing an explanatory typological variable.

The resulting typology makes it possible to go beyond the separate analysis of individual domains, offering a synthetic yet informed interpretation of vulnerability combinations. This is useful for setting intervention priorities and differentiating territorial policies. The results have been assigned to each municipality and serve as the basis for the typological and territorial analysis presented in the following section of the paper.

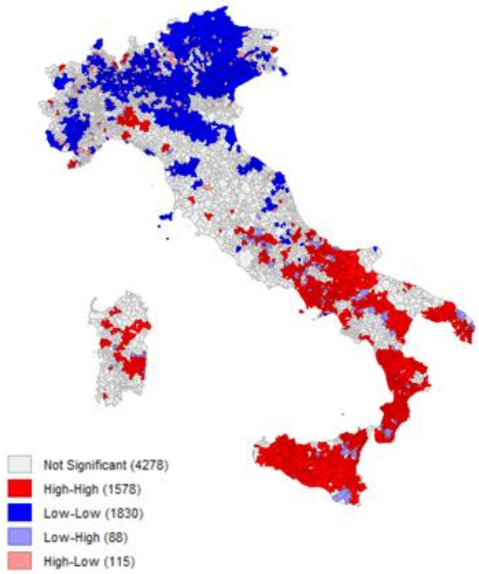
3. Results

3.1 Spatial Distribution of Fragility: LISA Analysis

The exploratory analysis carried out using the Local Indicators of Spatial Association (LISA) made it possible to identify significant patterns of homogeneous fragility among neighboring municipalities. Clusters of the High-High type emerged—groups of municipalities with high levels of fragility surrounded by others with similarly high fragility—as well as Low-Low clusters, representing "islands of resilience" within structurally robust areas (Figure 1).
The LISA cluster map highlights a well-known and marked territorial polarization:

- Northern Italy: A predominance of Low-Low clusters, particularly in urban and flat areas of the North-West and North-East, confirming the stronger structural and infrastructural capacity of these regions.
- Central Italy: A mix of areas, with High-High clusters mainly located in the inland Apennine regions (Abruzzo, Molise, Lazio) and Low-Low clusters in metropolitan cities and along the Tyrrhenian coast.
- Southern Italy and the Islands: A widespread concentration of High-High clusters, especially in Calabria, Sicilia, Campania, Basilicata, and inland Sardegna. These areas exhibit systemic fragility, evenly distributed across large portions of the territory.

Figure 1 – *Italian municipalities by IFC. LISA values, year 2021.*



3.2 Environmental and Territorial Typologies (Clusters A1–A5)

As mentioned in the previous section, the clustering analyses and the subsequent construction of the typology were limited to municipalities falling within the last three deciles (8th, 9th, and 10th) of the IFC distribution. These deciles identify municipalities characterized by high, very high, and extreme levels of fragility, representing the most structurally vulnerable territories at the national level.

The analysis of the physical and infrastructural characteristics of municipalities, conducted through clustering techniques, made it possible to identify five environmental and territorial profiles. Each cluster reflects a distinct configuration of vulnerability, resulting from the combination of variables related to accessibility, environmental protection, and anthropogenic pressure. This classification provides a clearer representation of the geography of environmental and infrastructural fragility, highlighting the heterogeneity of conditions in more peripheral and marginal contexts (Figure 3).

Cluster A1 – Accessible municipalities with low environmental protection. This represents the least critical configuration. It mainly includes flat and urban areas in the North and Centre of Italy, characterized by good infrastructure but limited coverage of protected natural areas. In this case, fragility is linked to potential risks from land consumption and a reduced presence of ecological buffers.

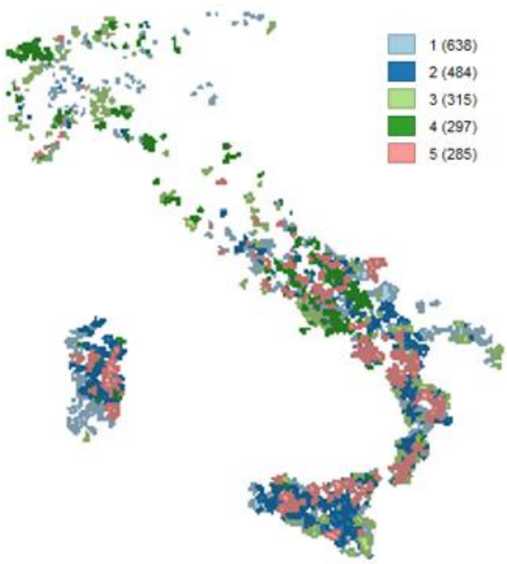
Cluster A2 – Peripheral areas under high anthropogenic pressure. Primarily concentrated in the South and the Islands, these areas suffer from poor accessibility to essential services and a highly polluting vehicle fleet. Infrastructural marginality combines with high environmental pressure, outlining a multidimensional vulnerability scenario.

Cluster A3 – Municipalities with multiple environmental issues. Mainly located in mountainous and hilly areas, these municipalities are marked by high exposure to hydrogeological risks (landslides, instability) combined with inefficiencies in waste management systems. These are internal and marginal territories, often penalized by persistent infrastructural weaknesses.

Cluster A4 – Areas with partial environmental protection and difficult access. Typical of Alpine and Apennine systems, these areas have some presence of protected zones but suffer from poor accessibility and vulnerability to natural hazards. Partial environmental protection is not enough to offset territorial isolation.

Cluster A5 – Isolated natural areas with high emission levels. These are sparsely populated areas with significant natural features but burdened by a high incidence of polluting vehicles. This configuration is common in Sardinia and some inland areas of Sicilia, where isolation coexists with latent vulnerabilities.

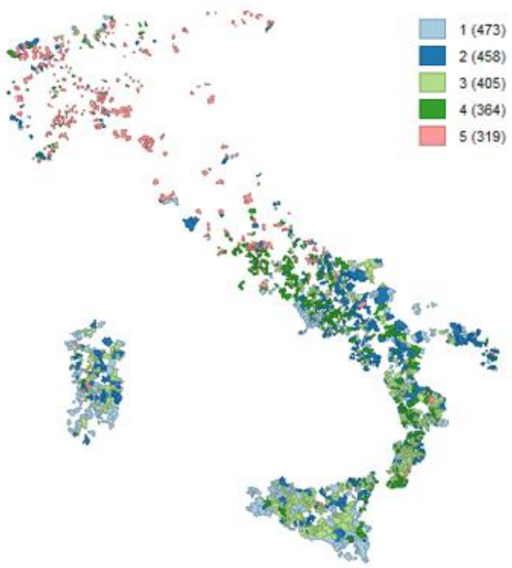
Figure 3 – *Choropleth map of municipal fragility typology – environmental and territorial dimension (Cluster A), year 2021*



3.3 Socio-Economic Typologies (Clusters S1–S5)

The analysis of social and economic dimensions led to the identification of five municipal profiles, resulting from a classification based on human capital composition, demographic dynamics, social cohesion, and productive structure. The resulting taxonomy captures differences in levels of social fragility across territories, with particular attention to the quality of human resources and the resilience of local communities. The first three clusters represent the most vulnerable configurations, while the last two reflect relatively more favorable conditions, especially from a social standpoint (Figure 4).

Figure 4 – *Choropleth map of municipal fragility typology – social and economic dimension (Cluster S), year 2021*



Cluster S1 – Socially vulnerable municipalities with partial economic stability. Includes territories with negative demographic trends, low education levels, and fragile social structures. It is widespread in various inland areas of the South and the Islands, as well as in some marginal zones of Central Italy.

Cluster S2 – Productive areas with marked social fragilities. These municipalities show a relatively dense productive fabric, accompanied by weak social conditions: low educational attainment, high demographic dependency, and low employment. Found primarily in Southern Italy and Sardegna, but also in some peri-urban areas of the Centre.

Cluster S3 – Marginal territories with structural social deprivation. This is the most critical configuration in social terms. It is marked by low employment rates, demographic unattractiveness, a depleted human capital base, and significant social vulnerability. It is widely present in the South, the Apennine hinterlands, and much of Sicilia, outlining a geography of marginality that aligns with the country's historical divides.

Cluster S4 – Municipalities with relatively strong human capital but economic vulnerabilities. Represents an intermediate profile, where the presence of social resources—such as higher educational levels or a more balanced demographic structure—partially offsets economic difficulties. These territories are distributed unevenly, with notable concentrations in Central Italy (Tuscan-Umbrian-Marchigian Apennines), parts of Sardegna, and the North-East.

Cluster S5 – Socially resilient contexts with strong cohesion. This is the most solid typology. It includes municipalities with good education levels, positive migration balances, demographic attractiveness, and a more balanced social structure. It is largely prevalent in Northern Italy (particularly in Lombardia, Emilia-Romagna, Veneto, and Trentino-Alto Adige), with smaller clusters also found in urban centers of the Centre.

3.4 Types of Combined Fragility (T1–T4) and Territorial Analysis by Macro-Area

The integration of the environmental-territorial and socio-economic classifications led to the definition of a four-class synthetic typology, aimed at representing the main interdimensional configurations of fragility among the selected municipalities (Tab. 1). The resulting profiles reflect significant differences in the intensity and nature of vulnerabilities and offer a useful basis for guiding differentiated intervention policies.

T1 – High combined fragility. Represents the most exposed municipalities, where socio-economic and environmental -infrastructural vulnerabilities overlap. These areas face marginality, poor service access, low human capital and employment, and high physical risks.

T2 – Predominantly social fragility. Includes municipalities with mainly social vulnerabilities, marked by low education, employment challenges, demographic dependency, and weak cohesion, while environmental conditions are less critical.

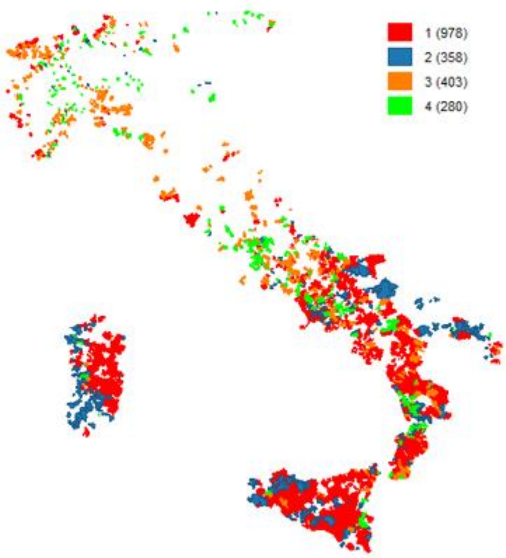
T3 – Predominantly environmental fragility. Comprises territories with infrastructural shortcomings, physical isolation, or exposure to environmental risks, yet with relatively solid social conditions.

T4 – Relatively lower fragility. It includes municipalities with more limited levels of vulnerability across both dimensions, characterized by a relatively favorable balance between social and environmental resources.

Table 1. - *Mapping between environmental and socio-economic cluster combinations and synthetic fragility typologies (T1–T4).*

	S1	S2	S3	S4	S5
A1	T2	T2	T2	T4	T4
A2	T1	T1	T1	T3	T3
A3	T1	T1	T1	T3	T3
A4	T1	T1	T1	T3	T3
A5	T1	T1	T1	T3	T3

Figure 5 – *Choropleth map of total municipal fragility typology (T), year 2021*.



The territorial distribution of the typologies reveals clear regional differences. Northern Italy is dominated by T4 and T3 typologies, indicating more selective or moderate fragility, particularly in environmental terms. Central Italy presents a heterogeneous picture, with all typologies present, but a higher incidence of T2 and T3 in inland and mountainous areas. Southern Italy and the Islands show the highest concentration of T1 and T2, reflecting widespread and multidimensional fragility, with pronounced social vulnerabilities even in areas with relatively favorable physical conditions (Figure 5).

4. Discussion, Policy Implications and Conclusions

The analysis reveals a geography of territorial fragility in Italy characterized by marked spatial and structural discontinuities. Persistent North–South divides coexist with more nuanced forms of internal marginality, particularly in mountainous and Apennine areas of the Centre and North, while more resilient dynamics emerge in major metropolitan zones.

The clustering and resulting typology (T1–T4) provide a nuanced understanding of fragility, distinguishing not only its intensity but also its nature—offering concrete implications for targeted and differentiated policy design:

*T1 – Combined fragility (environmental + social): r*equires integrated actions across infrastructure, environment, and human capital. Isolated measures would be insufficient in these highly disadvantaged contexts.

T2 – Predominantly social fragility: found in municipalities with good environmental conditions but social and demographic weaknesses. Interventions should focus on education, services, and strengthening territorial attractiveness.

T3 – Predominantly environmental fragility: calls for environmental planning, risk mitigation, and adaptive strategies to preserve physical resilience.

T4 – More resilient contexts: represent models to protect and reinforce through sustainability, prevention, and territorial cohesion strategies.

Methodologically, the study combines an institutional fragility index (IFC), clustering, and spatial analysis (LISA), enabling a systemic and localized reading of territorial fragility.

Looking ahead, this framework can evolve through time-series analysis and additional data sources, offering valuable support for territorial planning in an era of environmental, demographic, and digital transitions. Understanding differentiated forms of fragility is essential for effective, equitable, and sustainable public action.

5. Temporal limitations and sensitivity of the socio-economic component

It is important to emphasize that the analysis is based on data from the year 2021, the most recent available at the time of the study. This temporal constraint implies that the results provide a snapshot of territorial fragility in the immediate post-pandemic period. Although the Municipal Fragility Index (IFC) is designed to capture structural vulnerabilities, some dimensions—particularly those related to socio-economic conditions—may be subject to short- and medium-term variations.

In order to assess the sensitivity of the index to the pandemic phase, preliminary analyses were conducted using data from different reference years (pre- and post-pandemic). These explorations did not reveal significant changes in the overall levels of fragility, suggesting a relative stability of the IFC during the period considered. This does not exclude, however, the possibility that delayed structural effects may emerge over time, especially in relation to demographic shifts, employment conditions, or access to services. In this regard, the analysis may be appropriately updated as soon as more recent data become available, with the aim of capturing any transformations that may have occurred across territories.

References

BENASSI F., BUSETTA A., GALLO G., STRANGES M. 2022. Local Heterogeneities in Population Growth and Decline. A Spatial Analysis of Italian Municipalities. In SALVATI N., PERNA C., MARCHETTI S., CHAMBERS R. (Eds) *Studies in Theoretical and Applied Statistics. SIS 2021*. Springer Proceedings in Mathematics & Statistics, Vol. 406. Springer, Cham.

FRIGERIO I. DE AMICIS M. 2016. Mapping social vulnerability to natural hazards in Italy: A suitable tool for risk mitigation strategies, *Environmental Science & Policy*, Vol. 63, pp. 187–196.

FRIGERIO I., CARNELLI F., CABINIO M., DE AMICIS M. 2018. Spatiotemporal Pattern of Social Vulnerability in Italy, *International Journal of Disaster Risk Science*, Vol. 9, 249–262.

ISTAT. 2024. *Nota metodologica. L'Indice composito di Fragilità Comunale (IFC)*. Roma: Istituto Nazionale di Statistica.

MAZZIOTTA M., PARETO A. 2024. *Gli indici compositi*, Giappichelli Editore, Torino.

MENTGES A., HALEKOTTE L., SCHNEIDER M., DEMMER T., LICHTE D. 2023. A resilience glossary shaped by context: Reviewing resilience-related terms for critical infrastructures, *International Journal of Disaster Risk Reduction*, Vol. 96.

OECD 2022. *States of Fragility*, OECD Publishing, Paris.

UNESCO 2024. *Education and marginalization*. Paris: UNESCO.

UVAL (Unità di Valutazione degli Investimenti Pubblici) 2014. Strategia nazionale per le aree interne: definizione, obiettivi, strumenti e governance. *Materiali Uval*, No. 31, Roma.

Simona BALLABIO, Istat, ballabio@istat.it
Alberto VITALINI, Istat, vitalini@istat.it

THE WELL-BEING IN THE ITALIAN PERMANENT POPULATION AND HOUSING CENSUS: FIRST RESULTS[1]

Simona Mastroluca, Valeria Quondamstefano, Maria Carmela Russo, Donatella Zindato

Abstract. Since 2022, the Permanent Population and Housing Census, a combined census that integrates data from registers and sample surveys, collects information useful for measuring equitable and sustainable well-being at the local level.
To this end, a set of questions was added to the census questionnaire covering the domains of Social Relationships (network of relationships with relatives, friends and neighbours), Safety (perception of safety when walking alone in the dark and perception of crime risk in the place of usual residence) and Subjective Well-being (satisfaction with life).
This paper presents the first analyses aimed at highlighting well-being disparities across Italian provinces, while also considering demographic dynamics and socio-economic indicators. The Adjusted Mazziotta-Pareto Index (AMPI) was applied to synthesize individual indicators into a single, cohesive measure.

1. Introduction and objectives

In recent decades, the concept of well-being has undergone profound transformations. While economic wealth and the growth of Gross Domestic Product (GDP) were long regarded as the primary indicators of a society's progress, there is now a growing recognition that such measures, though useful, are insufficient to provide a comprehensive picture of quality of life. Well-being is a multidimensional phenomenon encompassing subjective, relational, psychological, and environmental aspects. In this context, indicators measuring how individuals perceive their lives, the social relationships they maintain, and the degree of safety they feel in their place of usual residence have gained increasing relevance. Among these, five are particularly significant: satisfaction with life, the presence of relatives, neighbours, or friends to rely on, and the perception of safety when walking alone in the dark. These are interconnected dimensions that reflect, from different perspectives, the degree of serenity, trust, and integration experienced in everyday life.

[1] Sections are attributed as follows: sections 1 and 2.4 to Simona Mastroluca, sections 3.1, 3.2 and 3.3 to Valeria Quondamstefano, sections 2.1, 2.2 and 2.3 to Maria Carmela Russo and sections 2.5 and 4 to Donatella Zindato.

Starting from 2022, the Permanent Population and Housing Census (PPHC), which employs a combined approach integrating register data and sample surveys, has collected information relevant to the Social Relationships, Safety and Subjective Well Being domains. This new census strategy enables the annual collection and dissemination of data, making the census more responsive to emerging user needs, including the investigation of territorial disparities in well-being indicators down to the NUTS3[2] level and in the largest cities[3]. Taken together, the five census indicators provide insights into fundamental aspects of people's lives that would otherwise be overlooked in a purely economic analysis. However, the integration of structural data and subjective perceptions remains one of the most complex yet essential challenges in the measurement of well-being today. Accordingly, this study first conducts a descriptive analysis of 2022 PPHC data at the provincial level to highlight disparities in subjective well-being across Italy. It then applies the Adjusted Mazziotta-Pareto Index (AMPI) both to these subjective indicators and, separately, to three structural indicators, also derived from Census data: the employment rate (20-64 years), the dependency rate and the percentage of population aged 25-64 with at least upper secondary education. The objective is to assess whether, in contexts of structural well-being, subjective well-being levels also exceed the national average, or conversely, whether these "immaterial" indicators – concerning social relations, feelings of security, and life satisfaction – are largely independent of the socio-economic conditions of the place of residence.

2. Census subjective indicators: a descriptive analysis

This section presents a descriptive analysis of the five subjective well-being indicators: networks of relationships with relatives, friends, and neighbours; perception of safety when walking alone in the dark and satisfaction with life. Using data from the PPHC 2022, the study highlights significant regional disparities, also focusing on gender differences.

[2] The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system that divides the territory of the European Union into different areas, used for statistical purposes. The NUTS3 level corresponds to provinces/metropolitan cities. In the paper, the term "provinces" refers to both provinces and metropolitan cities. In Italy, 14 of the 107 sub-regions at NUTS 3 level are called "metropolitan cities"; these are: Bari, Bologna, Cagliari, Catania, Florence, Genoa, Messina, Milan, Naples, Palermo, Reggio di Calabria, Rome, Turin, Venice, Verona
[3] Largest cities are municipalities with at least 150.000 usual residents.

2.1 Network of relationships with relatives

The network of personal relationships, through the sharing of both material and non-material resources, provides support in daily life and enhances individual well-being. Within the social and relational context, the family network plays a crucial role in Italy. Indeed, 87.2% of the population aged 14 and over report having non-cohabiting relatives they can rely on in case of need. This indicator presents a high territorial variability, with the highest values recorded in Southern Italy and in the Islands. The provinces with the highest shares are Nuoro (94.7%), Enna (93.4%) and Vibo Valentia (92.7%), whereas the lowest values are recorded in Prato (80.2%), Pesaro and Urbino (81.2%) and Rovigo (82.8%). At national level, the share of women who report receiving support from relatives is higher than men with a gap of 2 p.p. (88.2% of women versus 86.2% of men). The analysis confirms this trend across all provinces. The gender gap oscillates between 3.9 percentage points in Foggia, representing the highest disparity observed, and 0.5 percentage points in Rimini, the lowest value recorded.

2.2 Network of relationships with friends

Within the domain of subjective well-being, another dimension related to the sphere of social support is the network of friendships.

In Italy the percentage of people who have friends they can count on is 74.6%. At the provincial level, this value ranges from 87.8% of Nuoro to 67.8% of Rovigo, showing wider geographical differences compared to the data relating to relatives to rely on. Friendship networks are more present in Calabria, Sardinia and in the provinces along the northern border of the country, such as Sondrio and Bolzano/Bozen.

The breakdown by gender reveals only minor differences among males and females: in 81 out of 107 provinces/metropolitan cities the gap is minimal and in favour of men. The widest disparity is observed in L'Aquila (2.2 p.p.), followed by Pistoia (1.9 p.p.) and Lecce (1.6 p.p.). On the other hand, albeit overall low values, Ragusa, Bolzano, and Bergamo report the highest gender disparities in favour of women with 1.6, 1.4 and 1.3 percentage points respectively. Provinces characterized by a uniform gender distribution, where disparities are almost non-existent, include Sondrio, Enna and Prato.

2.3 Network of relationships with neighbours

The final indicator contributing to the social support network is the presence of neighbours (one or more individuals or households) one can rely on in times of need. According to 2022 PPHC data, 71.5% of the population aged 14 and over report being able to count on neighbours. The provinces of Nuoro (84.8%), Isernia (82.1%) record the highest percentages, followed by four of the five provinces of Calabria: Cosenza (81.8%), Reggio di Calabria (81.4%), Vibo Valentia (80.9%) and Catanzaro (79.4%). In contrast, the territories with the lowest percentage values are the provinces of Trieste (64.1%), Barletta-Andria-Trani (65.2%) and Foggia (65.5%). Women, with 72.4%, are more likely to report having access to a neighbourhood support network, (70.7% for men). At the NUTS3 level, gender differences in neighbourhood support networks are more pronounced than those observed in support networks based on relatives and friends. In particular, Ragusa registers the largest gender gap, with a difference of 5 percentage points in favour of women, followed by Foggia (3.6 p.p.) and Gorizia (3.2 p.p.).

Conversely, a limited number of provinces including Teramo, L'Aquila, Perugia, Isernia and Benevento exhibit gender gaps favouring men, although these differences remain marginal (below one percentage point).

2.4 Perception of safety when walking alone in the dark

One of the dimensions of well-being investigated in the PPHC is the perception of safety when walking alone in the dark in one's place of residence. The indicator considers those who declare themselves safe or fairly safe. Feeling safe in one's place of usual residence is essential for quality of life. However, this perception varies greatly according to geographical context and, especially, gender. The provinces with the highest incidence of residents aged 14 or over who consider themselves safe, despite the darkness and loneliness, are Sondrio (84.6%), Aosta (83.2%) and Nuoro (82.9%). On the contrary, in Prato, Naples and Rome the perception reaches the lowest levels with shares of 51.6% for the first one and 56.3% for the remaining two. Therefore, in all sub-regions considered the perception of safety is above 50% but there is a difference of more than 33 percentage points between the highest and the lowest extreme of the ranking.

Men always perceive themselves as more confident than women. More than three-quarters of men feel safe walking alone when it is dark, while the proportion is just over half among women (56.2%). The largest gender gap lies in the province of Monza and Brianza (26.6 p.p.), followed by Modena (25.0 p.p.) and Varese (24.9

p.p.). The distribution between men and women is more homogeneous in Agrigento and Isernia, where the differences are, in both cases, below 12 percentage points.

2.5 Satisfaction with life

Another dimension of subjective well-being investigated in the PPHC is the satisfaction with life, an indicator measuring the extent to which individuals feel their lives align with their expectations beyond temporary circumstances. The indicator considers those who declare themselves very satisfied (i.e. who rate their satisfaction with life between 8 and 10 on a scale of 1 to 10).

The national average life satisfaction rate across all provinces is 53.05%, with substantial intra-regional variability. In all provinces the percentage of people very satisfied is over 40%, though there is a difference of almost 23 percentage points between Bolzano and Napoli, at respectively the highest and the lowest extreme of the ranking. However, no clear territorial pattern or dichotomy North-South can be identified, as the provinces with the lowest shares of very satisfied people can be found all across Italy. In fact, Napoli ranks lowest, with only 41.5% of the population reporting high life satisfaction, followed by another southern province i.e. Taranto (45.2%) and then by the Tuscan province of Prato (45.6%) and by the two northern provinces of Rovigo (45.8%) and Ferrara (47.3%).

Similarly, at the other end of the spectrum, both northern and southern regions are represented, with Bolzano, Enna, Trento, Vibo Valentia and Sondrio displaying the highest life satisfaction shares (with respectively 64.4%, 61.6%, 61.2%, 60.0% and 59.8%). This suggests that regional averages may conceal important local dynamics and that life satisfaction reflects diverse local economic, social, or environmental conditions.

Even in a context of substantial gender balance (the gender parity index, comparing the value of the life satisfaction indicator for the female population to that of the male population falls between 0.95 and 1.05 for the large majority of the provinces), men consistently report being more satisfied with life compared to women. Only exceptions are of the provinces of Bolzano/Bozen, Siracusa, Crotone, Ragusa and Lecce, showing a slight female advantage (less than one percentage point). While on average men are nearly 2 percentage points more likely than women to report being very satisfied, the largest gap (4 or more percentage points) is observed in Tuscany and Emilia-Romagna provinces (Siena, Pistoia, Pisa, Piacenza, Reggio Emilia). The gender gap is usually smaller in provinces that display higher life satisfaction.

3. Methodology and Main results

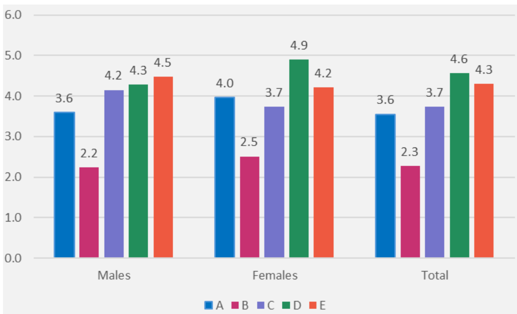
3.1 The Adjusted Mazziotta-Pareto Index

The methodology of composite indices in statistics provides a means of summarizing information from multiple variables into a single measure. This approach is particularly valuable in situations where it is necessary to condense a large amount of data into a more manageable and interpretable form (OECD, 2008). Among the different methods available for synthesising individual indicators into a single, consistent measure, this study employs the Adjusted Mazziotta-Pareto Index (AMPI), a partially non-compensatory composite index designed to standardize indicators at a reference point in time. This standardization ensures that indicators are independent of their original units of measurement (De Muro et al., 2011). By assigning equal weights to all indicators, the AMPI enables absolute temporal comparisons (Mazziotta and Pareto, 2016). The AMPI applies a re-scaling mechanism to individual indicators, positioning them within an open range of 70 to 130. This re-scaling relies on two ‘goalposts’ - minimum and maximum values - that define the potential range of each variable across time and units. These goalposts ensure comparability of indicators over time by providing a consistent measurement framework.

3.2 Application to subjective well-being indicators

An influence analysis of the subjective indicators was conducted to determine the average number of positions by which each territorial unit’s ranking shifts when one indicator is excluded (Figure 1).

Figure 1 – *Influence Analysis of subjective indicators - Year 2022.*



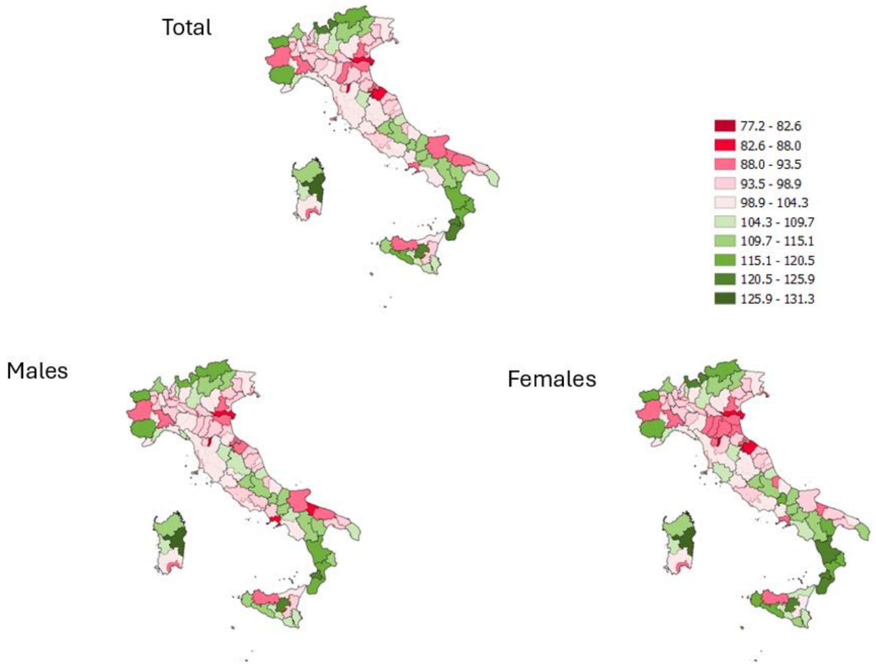
A	Network of relationships with relatives
B	Network of relationships with friends
C	Network of relationships with neighbours
D	Perception of safety when walking alone in the dark
E	Satisfaction with life

Source: *Elaborations on Permanent Population and Housing Census (PPHC) data, Istat.*

The indicator that produces the largest average ranking shifts for women and for the total population was “Perception of safety when walking alone in the dark” (4.2 and 4.3 positions, respectively), whereas for men it was “Satisfaction with life” (4.5 positions).

The AMPI of subjective indicators depicts a heterogeneous picture of Italy, with territorial patterns that are, in some cases, unexpected (Figure 2).

Figure 2 – *Maps of AMPI Subjective Indicators - Year 2022.*



Source: Elaborations on PPHC data, Istat.

Regions traditionally regarded as models of welfare policies and efficient public services, such as Emilia-Romagna, display lower levels of subjective well-being than structurally more disadvantaged areas, such as, for instance, Calabria and Sicilia.

Gender differences in perceived well-being also display distinct regional patterns. In all provinces of Emilia-Romagna, Umbria, and Veneto, women report lower levels of subjective well-being than men. Conversely, in Campania, Calabria, Molise, Apulia, Sicily, and Trentino-Alto Adige, women report higher levels than men, reversing the trend observed in most of the Centre-North.

The gap between the province with the highest AMPI score (Nuoro) and that with the lowest (Prato) amounts to as many as 53.3 points, underscoring marked territorial disparities (Table 1). Remarkably, all five provinces of Calabria rank among the top 12 in the national distribution. Conversely, one of the lowest AMPI score is recorded in Naples followed only by Prato, Rovigo, Pesaro-Urbino, and Barletta-Andria-Trani.

Another metropolitan city in the South, Reggio di Calabria, stands out with one of the highest AMPI scores, exceeded only by Nuoro, Enna, Vibo Valentia, and Sondrio. The data also reveal notable gender disparities in perceived support networks and safety: in Ragusa, Naples and Foggia, women report feeling significantly more supported than men. The opposite occurs in L'Aquila, Pistoia, and Perugia, where the gender gap in favour of men reaches 3.9, 3.5 and 3.2 points, respectively.

Table 1 – *Provinces/Metropolitan Cities with the Highest and Lowest AMPI values – Year 2022.*

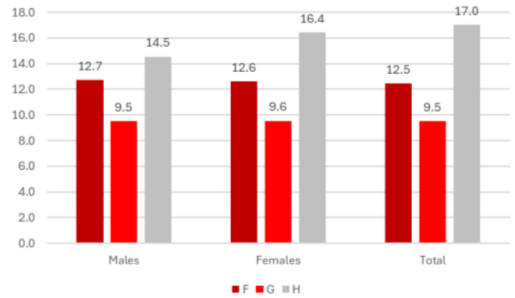
Areas	Regions	Provinces	AMPI
Islands	Sardinia	Nuoro	131.25
Islands	Sicily	Enna	122.39
South	Calabria	Vibo Valentia	121.64
North-West	Lombardy	Sondrio	121.59
South	Calabria	Reggio di Calabria	121.00
…	…	…	…
…	…	…	…
South	Campania	Naples	89.10
South	Apulia	Barletta-Andria-Trani	88.17
Centre	Marche	Pesaro and Urbino	94.17
North-East	Veneto	Rovigo	89.94
Centre	Tuscany	Prato	78.04

Source: Elaborations on PPHC data, Istat

3.3 Comparison with structural well-being indicators

To further investigate the relationship between subjective well-being and objective conditions, the AMPI was also applied to three structural indicators derived from census data: employment rate (20-64 years); dependency rate; percentage of the population aged 26-64 with at least upper secondary education. As with the subjective indicators, the first step of the analysis concerned the influence analysis. In this case, for males, females and the total population the most influential structural indicator is the "Employment rate", which shifts ranking by 14.5, 16.4, and 17.0 positions, respectively (Figure 3).

Figure 3 – *Influence Analysis of structural indicators - Year 2022.*

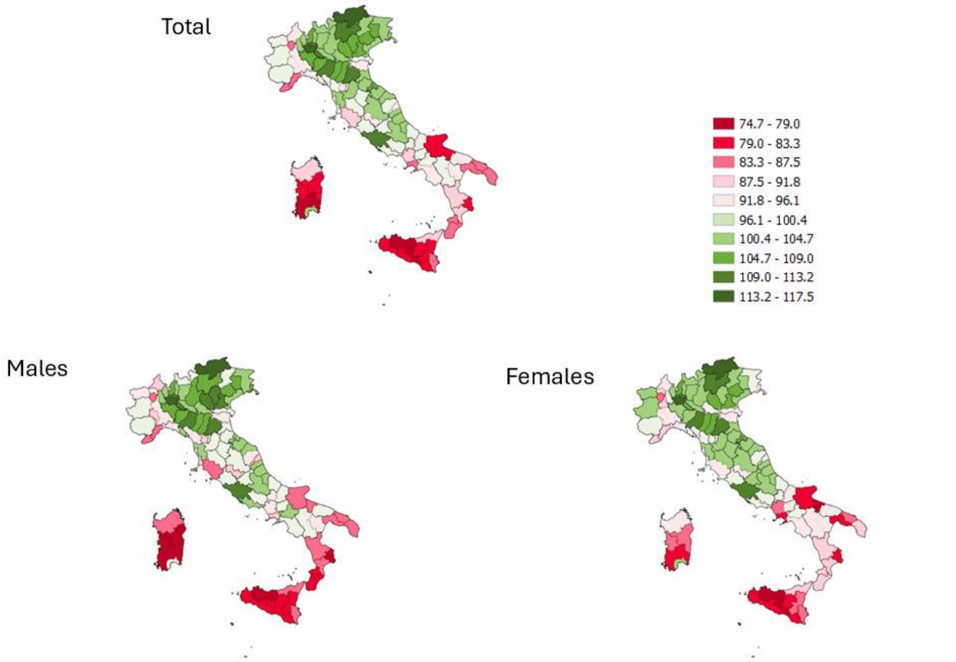


F	Dependency ratio
G	Percentage of population aged 25-64 with a higher qualification than a diploma of lower secondary education
H	Employment rate

Source: *Elaborations on PPHC data, Istat.*

The AMPI analysis of structural indicators reveals a territorial pattern in sharp contrast to that of subjective well-being, with the South and Islands emerging as the most disadvantaged areas. In particular, the Islands and Calabria show the lowest levels of structural well-being, especially among the male population, whereas in other southern regions women tend to experience worse conditions (Figure 4).

Figure 4 – *Maps of AMPI Structural Indicators - Year 2022.*



Source: *Elaborations on PPHC data, Istat.*

The highest levels of structural well-being are observed in the North-East, in Lombardy, and in Rome. In the North-East, only 3 out of 22 provinces (Trieste, Rovigo, and Ferrara) record AMPI scores below 100 (Table 2).

Table 2 – *Provinces with Highest and Lowest AMPI values – Year 2022.*

Areas	Regions	Provinces	AMPI
North-West	Lombardy	Milano	115.63
North-East	Trentino-Alto Adige	Bolzano/Bozen	115.27
North-East	Emilia-Romagna	Bologna	112.00
North-East	Emilia-Romagna	Parma	111.17
Centre	Lazio	Rome	111.11
…	…	…	…
…	…	…	…
South	Campania	Crotone	79.21
Islands	Sicily	Agrigento	79.08
Islands	Sicily	Palermo	77.83
Islands	Sardinia	South Sardinia	77.26
Islands	Sicily	Caltanissetta	77.08

Source*: Elaborations on PPHC data, Istat.*

In North-West Italy, only Lombardy and the province of Novara score above the national average. In the Centre, 13 out of 22 provinces fall below the Italian average, with a notable 7.4-point gap between Rome (111.1) and the next highest-ranking province in the Centre, Pesaro-Urbino (103.7).

In the South and Islands, only Pescara, L'Aquila, Cagliari, and Teramo score above the national average (AMPI > 100). Pescara ranks as the top province in the South, though it is preceded by 25 sub-regions from the Centre-North. The range between the highest (Milano) and lowest (Caltanissetta) AMPI scores across all Italian provinces is 38.5 points.

Interestingly, some provinces in the Centre-North also exhibit critical conditions. Biella ranks 91st out of 107 provinces, making it the lowest-scoring province in this macro-region. Furthermore, Turin and Genoa are the only metropolitan cities in the Centre-North with AMPI scores below 100.

The gender gap in objective well-being is also pronounced in several provinces. In Gorizia, Caserta, and Barletta-Andria-Trani, men report significantly higher levels of well-being than women, with gender gaps of 11.5, 7.7, and 7.1 points, respectively.

Summing up, comparison of AMPI structural values with AMPI subjective values reveals a counterintuitive pattern: populations with higher levels of socio-economic well-being do not consistently report corresponding levels of subjective well-being. This misalignment suggests the need for a more nuanced territorial analysis, and possibly taking into account many more factors such as petty crime (e.g., home burglaries and theft), the ratio of foreigners on the resident population and environmental vulnerabilities (e.g. earthquakes, floods), expectations and cultural norms for what constitutes a “satisfying life”, to gain a more comprehensive understanding of local conditions.

4. Summary remarks and outlook for the future

The inclusion of subjective well-being measures in the Permanent Population and Housing Census represents a major step forward in Italy’s ability to monitor equitable and sustainable well-being across territories on an annual basis. The initial analyses reveal a complex and heterogeneous landscape, with patterns that challenge conventional expectations based solely on economic or structural indicators, suggesting that perceived well-being is influenced by local contexts and does not necessarily align with structural socio-economic conditions (Becchetti L., 2021). As for the subjective well-being, notwithstanding the lack of a clear territorial pattern, it can be clearly observed that regions traditionally seen as models in terms of welfare policies and efficiency of public services show lower levels of satisfaction than areas (such as the South and the Islands) which appear to be more disadvantaged compared to the Centre-North when considering structural indicators. Moreover, the gender analysis highlights regionally differentiated patterns, even though in the Centre-North women generally report a lower perception of subjective well-being compared to men, while the opposite is true in the South and the Islands.

The application of the Adjusted Mazziotta-Pareto Index (AMPI) to subjective well-being dimensions has revealed noteworthy territorial disparities, with provinces in Southern Italy and the Islands often reporting higher subjective well-being than areas socio-economically more disadvantaged. Conversely, the analysis of structural well-being indicators confirms a clear North-South divide, with Northern and Central regions consistently outperforming Southern regions and the Islands, underscoring the persistence of territorial inequalities in economic and educational dimensions.

The coexistence of low structural well-being and higher levels of perceived well-being in several southern provinces underscores the need to complement traditional socio-economic indicators with subjective measures in order to provide a more comprehensive and accurate representation of quality of life at the local level. Furthermore, future studies should also investigate the relations between reported well-being and variables (such as age and citizenship) which previous analyses have shown to have an impact on subjective well-being, as well as the municipality size. Indeed, evidence from other analyses on large cities reveals non-negligible differences in subjective well-being between the capitals and the other municipalities of the respective metropolitan cities (Istat, 2025). This suggests that exposure to the highly competitive and stressful environments typical of major urban centres may affect perceptions of subjective well-being.

Finally, as the set of question on subjective well-being is now part of the PPHC, it will be possible to track the evolution of these subjective indicators over time, providing valuable evidence for policymakers interested in promoting a more equitable and sustainable well-being across the country.

References

BECCHETTI L. 2021. *Misurare il benessere in modo nuovo* in *Lavoce.info* https://lavoce.info/archives/91238/misurare-il-benessere-in-modo-nuovo/

DE MURO P., MAZZIOTTA M., PARETO A. 2011. Composite Indices of Development and Poverty: An Application to MDGs. *Social Indicators Research*, Vol. 104, pp. 1-18.

ISTAT. 2025. *Rapporto Annuale. La situazione del Paese, pp.143-146.*

MAZZIOTTA M., PARETO A. 2016. On a Generalized Non-compensatory Composite Index for Measuring Socio-economic Phenomena. *Social Indicators Research*, Vol. 127, No. 3, pp.983-1003.

OECD. 2008. *Handbook on Constructing Composite Indicators*. Paris: OECD Publications.

Simona MASTROLUCA, Istat. mastrolu@istat.it
Valeria QUONDAMSTEFANO, Istat, quondamstefano@istat.it
Maria Carmela RUSSO, Istat, mariacarmela.russo@istat.it
Donatella ZINDATO, Istat, zindato@istat.it

SPORT AND WELLBEING: NEW RIGHTS, INEQUALITIES AND PERSPECTIVES

Michele Camisasca, Emanuela Bologna, Ilaria Screpante

Abstract. The inclusion of sport among the principles enshrined in the Italian Constitution has represented a significant recognition by the State of the value of physical activity for the development of the individual and society as a whole. However, sport, as a domain of "concurrent legislation" between the State and the Regions, constitutes a subject divided between two legislative bodies: the first (the State), responsible for defining the fundamental principles, and the second (the Regions), tasked with its primary regulation in accordance with the national principles. This work aims to investigate whether and how legislative intervention has impacted the development and dissemination of sports practice across the entire population, with particular attention to the persistent inequalities still present.

1. Introduction – Sports in Italian Legislation

In 2001, for the first time, sports became one of the areas explicitly included within the Italian Constitution, but only with regard to the inclusion of the “sports system” among the subjects of concurrent legislation between the State and the Regions (Article 117 of the Italian Constitution). Consequently, sports became one of the subjects entrusted to the regulation of two legislative bodies: one (the State) responsible for defining the fundamental principles of sports; the other (the Regions) responsible for the primary regulation of sporting practice, in accordance with the general principles established by national laws. Regarding the connection between physical activity and personal development, no particular attention was given in 2001, although this aspect had already been highlighted at the European level within the Amsterdam Treaty of 1997 (Declaration No. 29 attached to the Treaty), which emphasized the social importance of sport. Later, in 2007, with amendments to the Treaty on the Functioning of the European Union, it was explicitly stated that the Union contributes to promoting the European dimension of sport, taking into account its specific features, its volunteer-based structures, and its social and educational functions.

In Italy, however, these values received recognition only later, as the dimension of sport began to emerge within the national education system. In 2009, guidelines were issued for motor education activities in secondary schools, highlighting how sport in the educational context helps increase civic awareness among youth, improve social cohesion, integration, and socialization. In 2015, with the “Reform of the National Education and Training System and Delegation for the Reorganization of Existing Legislative Provisions” (Law No. 107/2015), the “strengthening of motor disciplines” was identified as a key educational goal. The same year, Decree-Law No. 185/2015 recognized the need to develop sports facilities in urban outskirts as a priority tool to address economic and social imbalances and enhance urban safety, establishing the “Sport and Peripheries Fund”.

Subsequently, the 2018 Budget Law (Law No. 205/2017) established the “Single Fund to Support the Enhancement of Italian Sports Movement,” recognizing sports practice as an “inescapable form of personal development for minors,” and supporting the initiation of disabled individuals into sports activities through the use of assistive devices. The 2019 Budget Law (Law No. 145/2018) introduced the “Sport Bonus” measure, allowing taxpayers to benefit from a 65% tax credit on voluntary donations aimed at maintaining and restoring public sports facilities and building new public sports structures. This measure was extended throughout 2020 with the enactment of the 2020 Budget Law. In 2021, pending a comprehensive reform of the school system, the 2022 Budget Law introduced physical education as a subject in primary schools starting from the 2022/2023 school year, explicitly recognizing the role of sports “as an expression of a personal right and a tool for cognitive learning,” and requiring qualified teachers and registration in the related “Physical and Sports Sciences in Primary School” teaching category for fourth and fifth grades. Additionally, the 2022 Budget Law acknowledged a tax credit for documented expenses related to “adapted physical activity,” allowing individuals with chronic illnesses or physical disabilities requiring specialized assistance to recover part of their expenditure

We arrive at the year 2023 with the recognition of sport among the rights acknowledged in the Italian Constitution: “The Republic recognizes the educational, social, and well-being-promoting value of sporting activity in all its forms” (Article 33, paragraph 7, Constitution). Firstly, the educational value emerges, linked to the development and formation of the individual. This is complemented by the social value: indeed, sport often serves as a factor of social cohesion and an instrument of inclusion for individuals or groups facing various forms of disadvantage or marginalization, such as those related to socio-economic, ethnic-cultural, or physical-cognitive conditions. Finally, sport has an undeniable correlation with health, especially when understood in its modern conception as the comprehensive psycho-physical well-being of the person, rather than merely the absence of disease. The phrase recognizing the value of sporting activity “in all its forms” appears to be aimed at explicitly encompassing sport in its broadest sense.

From the analysis of the legislative provisions outlined above, it clearly emerges that, more recently, the socio-educational value of sport has been recognized among the constitutional principles. However, it is also evident that, for many years, the regulation of the subject of "Sport" has been governed through specific regulations and interventions aimed at supporting citizens' motor activity, particularly among the youth. In addition to the aforementioned legislative measures enacted by the State, regional laws have also been adopted to regulate various aspects of sport and physical activity, as well as their dissemination within the relevant territories. Therefore, to understand the role attributed to sport within Italy, it is essential to first consider the regional legislation.

The purpose of this work is therefore to highlight how physical activity has been interpreted by regional institutions; to examine the potential role that regional legislation may have played in the dissemination of sports practice; and to assess the impact of local institutional interventions in promoting the recognition of physical activity and sports as a means of individual and territorial development.

In the following paragraphs, we will describe the main aspects of regional legislation concerning sports and will emphasize, through an analysis of official statistical data on sports and physical activity, how, despite significant institutional recognition of the "right to sport" for all citizens, substantial differences and inequalities persist in the sports participation of the Italian population. This is a complex phenomenon involving gender, economic, social, and geographical factors..

2. Analysis of regional legislation on sports

For the analysis of regional legislation on sports, an initial in-depth study was conducted on the genesis and evolution of regional legislation in Italy regarding sports; a second in-depth study focused on the contents of the current regional laws issued by the 19 Italian Regions and the two Autonomous Provinces (Trento and Bolzano)[1]. Two important aspects have emerged: a) all Regions have issued specific measures regarding sports since the 1990s (approximately), although the first laws

[1] For the analysis of regional legislation and specific content, the laws issued by each region regarding sports were examined. In particular, the current legal texts (and the subsequent amendments and integrations over time) were reviewed: Abruzzo –LR n. 2/2018 s.m.i., LR n. 20/2020; Basilicata – LR n. 26/2004 s.m.i.; P.A Bolzano – LP n. 19/1990 s.m.i.; Calabria – LR n. 28/2010 s.m.i.; Campania – LR n. 18/2013 s.m.i.; Emilia Romagna – LR n. 8/2017, LR n. 2/2024; Friuli Venezia Giulia – LR n. 8/2003 s.m.i.; Lazio – LR n. 15/2002 s.m.i.; Liguria – LR n. 40/2009; Lombardia – LR n. 26/2024 s.m.i.; Marche – LR n. 5/2012 s.m.i.; Molise – LR n. 23/2016 s.m.i.; Piemonte – LR n. 23/2020 s.m.i.; Puglia – LR n. 33/2006 s.m.i., LR n. 14/2017; Sardegna – LR n. 17/1999 s.m.i.; Sicilia – LR n. 29/2014 s.m.i.; Toscana – LR n. 21/2015 s.m.i.; P.A. Trento – LP n. 4/2016 s.m.i.; Umbria – LR n. 19/2009 s.m.i.; Valle d'Aosta – LR n. 3/2004 s.m.i.; Veneto – LR n. 8/2015 s.m.i.

enacted between the late 1980s and the 1990s concern sports facilities and the organization of sporting events; b) in all currently valid regional laws, a clear orientation emerges regarding the social and educational value of physical activity and sports practice, as elements of personal and community growth and development, social inclusion, and reduction of inequalities.

The in-depth analysis of regional legislation regarding sports reveals a fairly homogeneous framework concerning the presence of a territorial discipline, as well as widespread recognition of the central role of physical activity in the social and cultural development of individual territories, albeit with different approaches among the Regions. Some regions (such as Emilia-Romagna, Marche, Basilicata, Liguria, Toscana, Puglia, Autonomous Province of Trento, Umbria, Sardegna, Valle d'Aosta) explicitly acknowledge the social value of sport and/or "Sport of citizenship," often identifying vulnerable groups of citizens to whom special attention should be given (people with disabilities, the elderly, at-risk youth). Legislation in Lazio, Piemonte, and Campania also explicitly recognizes sport as a tool to reduce inequalities and promote social inclusion, directly mentioning disadvantaged subjects, people with disabilities, social/economic hardship, and equal opportunities. Emilia-Romagna appears to have the most detailed and integrated focus on reducing dropout (with a specific law addressing these aspects) and promoting social and economic inclusion through sport. Marche, Liguria, and Toscana emphasize accessibility and usability of facilities for all. Molise and Abruzzo show less emphasis on these specific aspects related to inclusion and inequality reduction. Almost all the examined regions foresee the training of operators and some form of monitoring or data collection. Direct economic support for individual participation, beyond specific cases (such as medical visits or municipal sports vouchers), is not widely evident, as it is typically deferred to sector-specific economic and financial measures. The Autonomous Province of Bolzano and Sicilia seem to focus more on the general promotion of sport, support for facilities and organizations, and, in the case of Sicilia, on the regulation of activities and personnel, with few or no explicit provisions specifically aimed at reducing inequalities or ensuring universal access for disadvantaged groups. In order to summarize the results of the analysis on regional legislation on Sport, some key elements directly connected to the socio-educational value of Sport present in most of the legislative texts under examination have been identified and standardized. Below is a diagram highlighting the areas of discipline uniformly present in regional laws, which highlight the recognition attributed by local institutions to sports-motor practice as a new form of citizenship right.

Table 1 - *Areas uniformly regulated in regional sports laws.*

Common Elements in Regional Laws	Description
Recognition of the socio-educational-health function of Sport	The law explicitly recognizes sport/physical activity as a tool for social inclusion, reduction of inequalities and promotes access for all
Specific target groups of interventions - policies	The law identifies recipients of specific interventions, vulnerable groups or groups at risk of exclusion
Accessibility and Usability of Facilities	The law promotes measures to guarantee access, availability and usability of sports facilities and spaces for all citizens
Financial Support for Participation	The law provides for forms of financial support for individuals or associations, in order to facilitate participation
Promotion in the Educational/School Sector	The law actively promotes sports and physical activity in schools or in collaboration with educational institutions
Training of Operators	The law promotes initiatives for the training and qualification of sports operators
Monitoring and Research	The law provides for the establishment of bodies to monitor the sporting phenomenon (observers, data)
Integration with Other Policies	The law promotes connections and synergies between sports and social policies, health, education, work

Sport, therefore, from what emerges from the analysis of regional legislation, represents in a priority way an activity supported and guaranteed by the institutions for the well-being of people, economic development, social cohesion and the growth of a more integrated society. In light of this sensitivity of the regional bodies towards a vision of sport as a driver of individual and territorial growth, the data of the official statistics on the participation of the population in sporting and physical activities were observed, to understand whether the recognition that the institutions attribute to sport corresponds to a real diffusion of these activities throughout the population.

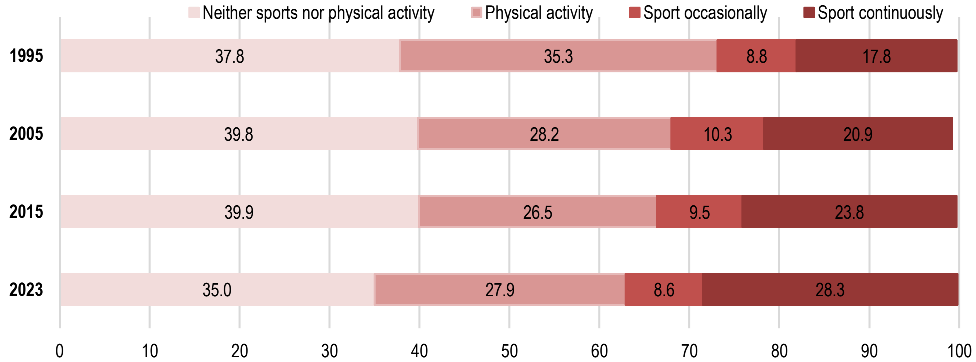
3. Trends in Sport Participation in Italy: a statistical overview

Physical-sport practice represents a fundamental factor in promoting the health and well-being of the population; therefore, the accurate and systematic monitoring of statistical data related to this practice plays a crucial role in deepening the understanding of participation dynamics and supporting the implementation of effective sport policies. In 2023, 36.9% of the population aged 3 and over engaged in sports activities, with 28.3% participating regularly and 8.6% occasionally.

Additionally, 27.9% undertook at least some form of physical activity (including movements and exercises not strictly classified as sports). From 1995 to 2023, there has been a positive trend in sports participation, particularly in regular activity, which increased from 17.8% to 28.3%. This growth persisted steadily throughout the pandemic, despite restrictions on both indoor and outdoor sports (Figure 1).

Concurrently, the proportion of individuals who neither practice sports nor engage in physical activity declined from 37.8% in 1995 to 35% in 2023. This suggests a gradual improvement in physical engagement across the population, with positive implications for public health, given the crucial role of physical activity in preventing chronic diseases and enhancing overall wellbeing.

Figure 1 – *People aged 3 and over who, during their leisure time, either practise sports continuously or occasionally, engage in some physical activity, or do neither. Years: 1995*, 2005, 2015, and 2023. Percentage values.*



*Source Istat, Aspects of daily life Survey * The 1995 data refers to the survey "Leisure Time and Culture".*

Sport participation is more common among young people, especially those aged 6 to 14, where nearly 7 out of 10 engage in regular sporting activities. However, from the age of 15, there is a noticeable decline in sports participation, which is particularly pronounced among girls. This downward trend becomes even more marked after the age of 25, partly due to increased work and family commitments limiting available time for sports.

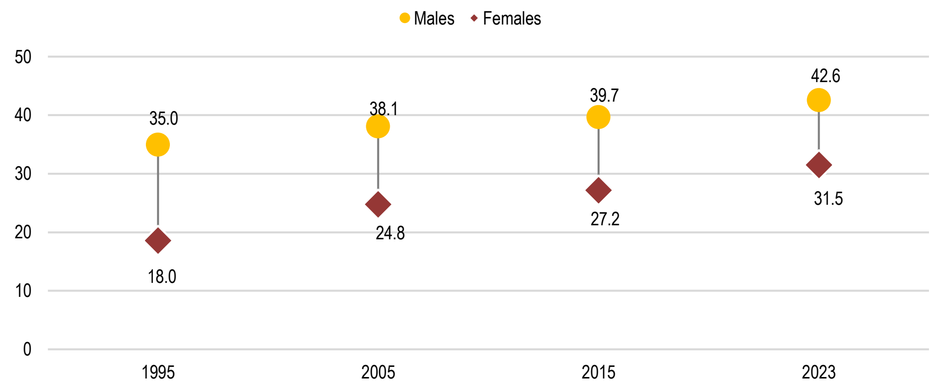
Sedentary behaviour tends to increase with age. Approximately 2 out of 10 adolescents and young adults up to 24 years old (people aged from 15 to 24 years) lead predominantly sedentary lifestyles, while this figure rises significantly among the elderly (people aged 65 and over), affecting nearly 7 out of 10 individuals aged 75 and over[2]. This highlights the need for targeted interventions to promote

[2] https://esploradati.istat.it/databrowser/#/en/dw/categories/IT1,Z0850DAI,1.0/SPORT_FRIENDS

continuous engagement in sport and physical activity throughout life, in order to counteract the negative effects of sedentariness and support a healthy, active lifestyle[3].

Over time, from 1995 to 2023, the gender gap in sports participation has gradually narrowed, decreasing from 17 to 11.1 percentage points (Figure 2). This reduction is especially evident among young people, indicating a closing gap in sports participation levels between males and females in these age groups[4].

Figure 2 – *Percentage values. People aged 3 and over who, during their leisure time, practise sports (continuously or occasionally) by gender. Years 1995, 2005, 2015, 2023. Percentage values.*



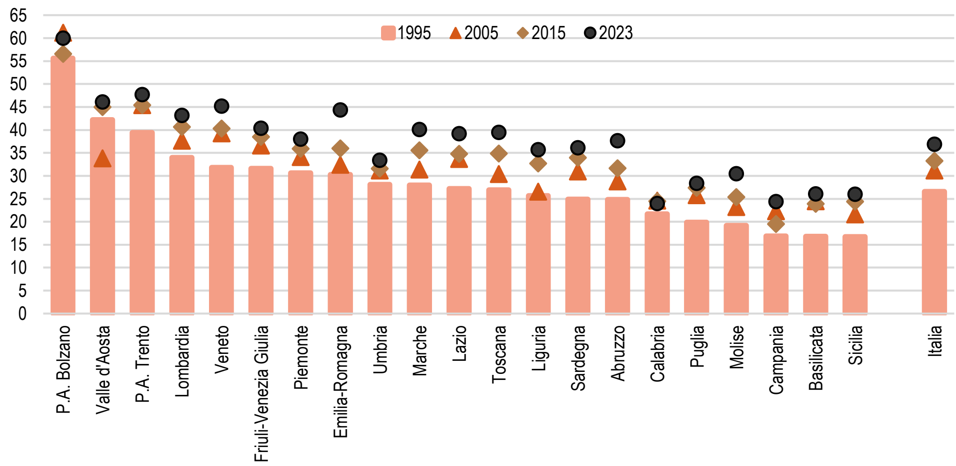
Source Istat, Aspects of daily life Survey.

However, gender differences remain significant in certain age groups, still exceeding 10 percentage points. It is important to note that among children aged 3 to 5 years, no substantial gender differences in sports participation are observed, suggesting near parity in early childhood. With increasing age, gender disparities tend to widen, highlighting the need for targeted measures to promote greater inclusion and female participation in sport, especially in age groups where the gap is most pronounced. From 1995 to 2023, sports participation increased across all Italian regions (Figure 3), although significant territorial disparities persist, particularly the longstanding North-South divide. In Northern Italy, approximately 43% of the population engage in sports, compared to 39% in the Centre and around

[3] Istat (2024), Fumo, alcol, eccesso di peso e sedentarietà – Anno 2023 . Statistica report 17 dicembre 2024.
[4] https://esploradati.istat.it/databrowser/#/en/dw/categories/IT1,Z0850DAI,1.0/SPORT_FRIENDS

27.5% in the South and Islands. This divide is accompanied by higher levels of sedentary behaviour in southern regions, where nearly 49% of the population lead sedentary lifestyles. Some southern regions, such as Basilicata (54.2%), Campania (53.1%) and Sicily (52.7%), report particularly high rates of sedentariness[56].

Figure 3 – *Percentage values. People aged 3 and over who, during their leisure time, practise sports (continuously or occasionally) by regions. Years 1995, 2005, 2015, 2023. Percentage values.*



Source Istat, Aspects of daily life Survey.

The persistent disadvantage of Southern Italy is often attributed to a lack of adequate and easily accessible sports facilities. Supporting this, in 2024 approximately 40% of residents in southern regions consider the available sports infrastructure unsatisfactory—a proportion roughly halved in northern regions[7]. These data highlight the need for targeted interventions to improve the availability and quality of sports facilities in the South, aiming to reduce regional inequalities in sports participation and promote a more active lifestyle nationwide.

Moreover, in 2023, sport participation was highest in central municipalities of metropolitan areas (41.8%) and in nearby zones (35.6%), while smaller municipalities with up to 2,000 inhabitants showed lower levels (28.9%). Between 1995 and 2023, participation increased significantly in large municipalities (+18.7 percentage points), whereas the growth in small municipalities was minimal (+1.5

[5] Idem.
[6] People aged 3 and over who do not engage in either sports or physical activity.
[7] Istat (2025) La pratica sportiva in Italia. Statistica Today, 30 giugno 2025

points), highlighting ongoing territorial disparities. Only 15.3% of individuals with at most a lower secondary education engage in sports, compared to 53.8% of university graduates. Sociocultural differences in sports participation are evident across both genders and all age groups, emphasizing education level as a key differentiating factor. People with higher educational attainment tend to participate in sports more frequently, regardless of age, suggesting education plays a crucial role in encouraging active lifestyles. Although sports participation declines with age across the population, inequalities linked to education persist, underscoring the need for specific strategies to improve inclusion and access to sport among socioeconomically disadvantaged groups[8]. Confirming these observations, the European Health Survey[9] highlights well-established socio-economic inequalities that strongly influence healthy lifestyles. In particular, economic resources play a decisive role in facilitating access to adequate and regular physical activity. In many European countries, including Italy, income disparities translate into significant differences in physical activity levels. These inequalities tend to increase with age and are particularly pronounced among women, illustrating how economic and gender factors interact in shaping opportunities to maintain an active and healthy lifestyle. Such findings reinforce the need for targeted policies that address economic barriers and promote social inclusion, aiming to reduce inequalities and ensure equitable access to physical activity across all population segments[10].

4. Conclusions

This paper has highlighted the evolution of the institutional recognition of sport in Italy, culminating in the inclusion of the right to sport in the Constitution (Art. 33, paragraph 7). This milestone marks a fundamental step in the acknowledgment of sport not only as a means of physical well-being, but also as an educational, social, and inclusive tool. However, the analysis has shown that the formal recognition of the value of sport is not yet fully matched by an equitable and widespread participation in sports activities across the population.

From a legislative perspective, both national and regional laws have increasingly recognized the role of sport in individual development and territorial growth. Regional governments, in particular, have played a key role in promoting initiatives ranging from support for sports facilities to the inclusion of vulnerable groups, demonstrating a growing awareness of sport as a right of citizenship. Nonetheless,

[8] https://esploradati.istat.it/databrowser/#/en/dw/categories/IT1,Z0850DAI,1.0/SPORT_FRIENDS
[9] https://www.istat.it/microdati/indagine-europea-sulla-salute-ehis-file-per-la-ricerca/
[10] Istat (2021) Prevenzione e fattori di rischio per la salute in Italia e in Europa. Statistica report 16 Dicembre 2021.

differences between regions—in terms of policy approaches and available resources—have contributed to maintaining significant territorial disparities, with the South of Italy remaining at a clear disadvantage.

Statistical data confirm the persistence of deep inequalities in access to and participation in sport. Factors such as gender, age, education level, economic status, and geographic location significantly influence the ability to engage in physical activity. Women, individuals with lower educational attainment, residents of southern and rural areas, and economically disadvantaged groups are less involved in sports practice, despite legislative efforts aimed at ensuring universal access.

This mismatch between principles and practice underscores the need for stronger public policies—not only through infrastructure investments, but also via targeted economic support, cultural promotion, and social inclusion strategies. Furthermore, improving policy monitoring, evaluation mechanisms, and the training of sports professionals is essential if sport is to become a structural tool for well-being and equity.

In conclusion, the constitutional recognition of the right to sport presents a valuable opportunity to build a more equitable, active, and cohesive society. For this potential to be fully realized, it is crucial that the right to sport is effectively guaranteed to all citizens, with particular attention to those currently excluded. Only then can sport truly fulfill its educational, social, and health-promoting functions.

References

ISTAT. 2021. Prevenzione e fattori di rischio per la salute in Italia e in Europa. Statistica report 16 Dicembre 2021.

ISTAT. 2024. Fumo, alcol, eccesso di peso e sedentarietà – Anno 2023 . Statistica report 17 dicembre 2024.

ISTAT. 2025. La pratica sportiva in Italia. Statistica Today, 30 giugno 2025.

ISTAT: European health Survey: https://www.istat.it/microdati/indagine-europea-sulla-salute-ehis-file-per-la-ricerca/

ISTAT: Aspects of daily life Survey: https://www.istat.it/en/microdata/multipurpose-survey-on-households-aspects-of-daily-life/

ISTAT: Citizens and leisure time sample survey: https://www.istat.it/en/microdata/citizens-and-leisure-time-sample-survey-ctl-microdata-for-research-purposes/

ISTATDATA: data on sports: https://esploradati.istat.it/databrowser/#/en/dw/categories/IT1,Z0850DAI,1.0/SPORT_FRIENDS

Decree-law of November 25, 2015, No. 185. Urgent measures for interventions in the territory, converted with amendments by Law No. 9 of January 22, 2016 (GU No. 18 of January 23, 2016).
Law of July 13, 2015, No. 107. Reform of the national education and training system and delegation for the reorganization of existing legislative provisions (GU No. 162 of July 15, 2015).
Law of 27 December 2017, No. 205. State Budget for the 2018 financial year and multi-year budget for the three-year period 2018-2020 (GU no. 302 of December, 2017 - Ordinary Supplement no. 62).
Law of December 30, 2018, No. 145. State Budget for the 2019 financial year and multi-year budget for the three-year period 2019-2021 (GU No. 302 of December 31, 2018 - Ordinary Supplement no. 62).
Treaty of Amsterdam amending the Treaty on European Union, the Treaties establishing the European Communities and certain related acts, signed at Amsterdam, 2 October 1997.

Michele CAMISASCA, Istat, michele.camisasca@istat.it
Emanuela BOLOGNA, Istat, bologna@istat.it
Ilaria SCREPANTE, Istat, screpante@istat.it

SUSTAINABLE AND RESILIENT MOBILITY IN THE 14 METROPOLITAN CITIES CAPITALS[1]

Monica Bailot, Anna Maria Cecchini, Livia Fioroni, Patrizia Grossi, Susi Osti, Valeria Quondamstefano, Valentina Spinella

Abstract. In an environmental sustainability perspective, aimed at promoting a cultural shift inspired by environmental respect and the improvement of urban mobility, the research project pursues the objective of determining some indicators sensitive to territorial differences, robust and suitable for exploring the dynamics and factors of sustainable mobility compatible with local development and territorial policies. To conduct the research and allow multidimensional analysis and evaluation of sustainable mobility differences at the territorial level, it was chosen to construct a composite index that allows for the synthetic measurement of complex and multidimensional phenomena. Sustainable urban mobility is analysed considering three domains: private motorization; public transport; active mobility and sharing.

1. Introduction

Sustainable mobility is a crucial issue for governance, the ecosystem, road safety and people's health itself. Starting from the Green Deal (European Green Pact), the strategic initiatives promoted by the European Commission aim to start the EU on the path to a green transition, with the ultimate goal of achieving climate neutrality by 2050. Intermediate objective: reduce greenhouse gas emissions by 55% compared to 1990 levels by 2030 (NetZero2030).

With the aim of raising awareness among citizens and stakeholders, the paper represents the first results of a multidimensional analysis on the 14 capital municipalities of metropolitan cities, part of a broader research project (Sustainable and Resilient Mobility (MOSER) with the following objectives. In particular, the research project proposes to:

- ✓ identify methods and models for the multidimensional analysis and evaluation of sustainable mobility at a territorial level;
- ✓ design an integrated monitoring system on territorial inequalities;

[1] This article is the result of the collaboration between the authors. In particular: Anna Maria CECCHINI wrote Paragraph 1; Susi OSTI wrote Paragraph 2; Monica BAILOT wrote Paragraph 3; Valeria QUONDAMSTEFANO wrote Paragraph 4; Livia FIORONI wrote the Paragraph 5; Valentina SPINELLA wrote the Paragraph 6. However, the present paper is the result of a joint work.

- ✓ increase knowledge of the phenomena, raise awareness among citizens and institutions of the related environmental impact.

The literature on the topic includes some interesting studies based on official data, such as the MobilitAria 2018 Report (Kyoto Club and CNR-IIA), which analyzes the trend of air quality and sustainable mobility in the 14 major Italian cities over the decade 2006-2016; the 21st Report on the Mobility of Italians (Isfort, 2024), which contains a description of individual indicators to represent the trend of the phenomenon, but lacks synthetic indices comparable over time and space. Hence the need to calculate a general composite index.

To carry out the research and enable a multidimensional analysis of territorial differences in sustainable mobility, a composite index was developed to provide a synthetic measurement of complex and multidimensional phenomena. Sustainable urban mobility is analysed considering three domains: private motorisation, public transport and active mobility and sharing. The *Adjusted Mazziotta-Pareto Index* (AMPI) was calculated for each domain, focusing on 14 Metropolitan city capitals, over the period 2016-2022.

2. Data Base

When constructing a synthetic index, the selection of indicators is the result of a trade-off between redundancy and loss of information. The starting database comprises 35 simple indicators from 2016 to 2022, with territorial detail including provincial capitals, geographical subdivisions, and the national total.

Following an in-depth descriptive analysis—including tables, graphs, and statistical measures to facilitate a comprehensive understanding of sustainable mobility dynamics—the trends of several key indicators were examined. The selection of these indicators was guided by a reasoned and conceptually grounded criterion, ensuring that the final set adequately represents all relevant aspects within the three domains: Private Motorisation, Public Transport, and Active Mobility and Sharing. The selected indicators are national in scope, representative of the phenomenon under investigation, and many are aligned with the objectives of specific legislation and are commonly used in international comparisons.

Among other qualities, all indicators are well documented and regularly updated. They are considered robust, as they are based on national and/or international standards and benefit from broad consensus regarding their validity. This ensures their comparability over time and across different geographic contexts. The chosen indicators show variations in time and space, making them largely sensitive to the dynamics of the multidimensional phenomenon to be analysed, and are available in time and spatial series.

In Italy, the private car dominates urban travel. Between 2016 and 2022, in all cities, the motorization rate remained very high, with more than six cars per ten inhabitants. At the end of this period, the number of cars per inhabitant is still growing, driven by the evolution of teleworking/working for home (WFH), although it has not reached 2019 levels. However, the pressure on the environment caused by vehicle traffic depends not only on the number of vehicles but also on their composition. An adequate number of low-emission vehicles can reduce this pressure. In contrast to the motorization rate, the index of pollutant potential associated with vehicles on the road has shown a decreasing trend in recent years: between 2016 and 2022, it fells from 153.0 to 116.2 in Italy.

In the post-pandemic scenario, weak signs of ecological transition are emerging from cities. Public transport is struggling to recover after the drastic reduction in passenger numbers caused by the health restrictions introduced to contain COVID-19 infection. It is still unclear how active walking and cycling will evolve in the years to come, despite having boomed during the pandemic.

The National Plan for Cycling Mobility has set the goal of achieving a density of 32 km of cycling paths per 100 km² by 2024. In Italy, despite a significant increase in the period 2016-2022 (from 21.9 km/100 km² to 27.9 km/100 km²), the gap to the target is still evident.

3. Metadata

The information sources of the selected individual indicators are the Istat survey ‘Urban Environmental data', *Automobile Club Italiano* (ACI) archives and the *Pubblico Registro Automobilistico* (PRA) archives. The following indicators are calculated on data from ACI and the PRA archives. Inhabitant data are derived from the Permanent Population and Housing Census (PPHC); therefore, all indicators per inhabitant are recalculated in time series based on the revision of the intercensary interval of the resident population.

Domain “Private Motorization”:

(A) *Motorization rate for car and motorcycles (per 1,000 inhabitants).* Ratio of the number of passenger cars/motorcycles and the resident population in the reference year multiplied by 1,000 (-);

(B) *Percentage of cars with low-emission.* Ratio of the number of cars with electric traction, hybrids (dual engine, electric and combustion), gas (methane, LPG or hydrogen) or bi-fuel (dual fuel, petrol and gas) and the total number of cars multiplied by 100 (+);

(C) *Percentage of cars with high-emission (Euro 4 or lower).* Ratio of the number of passenger cars with an emission class less than or equal to 4 and the total number of cars multiplied by 100 (-);

(D) *Index of pollutant potential of cars.* Ratio of the sum of the number of high (from Euro 0 to Euro 3) and medium polluting (powered by petrol or diesel from Euro 4 to Euro 6) passenger cars and the sum of medium polluting and low polluting (electric cars and other low emission cars from Euro 4 to Euro 6, hybrid, powered by natural gas or LPG or bi-fuel) ones multiplied by 100 (-).

The source of the following indicators is the ‘Urban Environmental data' survey. Istat launched it in 1998 in the 20 regional capital municipalities and 2 provincial capitals, Bolzano and Catania. Since 2002, this survey has involved all provincial capital municipalities.

Domain “Public Transportation”:

(A) *Demand for local public transportation (annual passengers per inhabitant).* The indicator considers all the following modes of LPT: Bus, Tram, Trolleybus, Underground, Waterborne transport, Funicular, Cable car and other hectometric systems. Suburban or metropolitan rail services are excluded and the indicator corresponds to the average number of LPT passengers per inhabitant (+);

(B) *Availability of buses for local public transport (vehicles per 100 thousand inhabitants).* Ratio of the number of vehicles available for daily public transport operations and the resident population in the reference year multiplied by 100,000 (+);

(C) *Total seat-kilometers offered by local public transport (values per inhabitant).* Ratio of the number of seat-km (summation, for each vehicle used, of the product of available seats and kilometers travelled) and the population resident in the reference year (+).

Domain “Active mobility and sharing”:

(A) *Density of bicycle paths (km per 100 square km of land area).* Ratio of the length of cycle paths, expressed in km, to the reference land area according to the Istat geographic information system (+);

(B) *Vehicle availability of car sharing services (vehicles per 10 thousand inhabitants).* Ratio of the number of public cars available for reservation (station-based or free flow) and the resident population in the reference year multiplied by 10 thousand (+);

(C) *Availability of vehicles used for bike sharing, scooter sharing and electric micromobility services (vehicles per 10 thousand inhabitants).* Ratio of the number of public vehicles for micromobility on reservation and the resident population in the reference year multiplied by 10 thousand (+).

4. Methodology

To evaluate the differences in sustainable mobility at territorial level, taking into account the spread of people's modes of travel, with particular reference to private, collective and smart mobility, a composite indicator was selected.

The construction of a composite indicator is a complex process that involves aggregating individual indicators into a single index, grounded in an underlying conceptual framework that reflects the multidimensional nature of the phenomenon being measured. The main challenges in this approach include the choice of the theoretical framework, the selection of the most representative indicators, and their treatment to compare and aggregate them. The first challenge is the choice of the theoretical framework, which is crucial for guiding the construction of the composite indicator (OECD, 2004). This framework must capture all relevant dimensions of the studied phenomenon, as a lack of clear theory can result in an indicator that does not accurately reflect the multi-dimensional concept. Another critical step is the selection of representative indicators.

The chosen indicators must be relevant, measurable, and reliable to ensure the validity and relevance of the composite indicator. An inadequate selection of indicators can compromise the reliability and validity of the composite index. Once the indicators have been selected, they must be processed and normalized to ensure comparability across units and dimensions. Normalization methods such as standardization or min-max transformation are commonly used, but the right method can vary depending on the nature of the indicators and the theoretical model adopted. The next step is the aggregation of these indicators. This process can be complex, requiring decisions on how to weight the different indicators and which aggregation method to use, such as arithmetic means, weighted means, geometric means, or other advanced statistical methods. After constructing the composite indicator, it is essential to validate its accuracy and reliability through robustness tests, sensitivity analyses, and comparisons with existing measures.

Validation ensures that the composite indicator provides a truthful and useful representation of the multi-dimensional concept.

Finally, the composite indicator must be correctly interpreted and clearly communicated. The results should be presented in a way that is accessible to stakeholders, enabling them to inform policy or managerial decisions effectively.

In summary, constructing a composite indicator requires a series of well-considered steps, each presenting specific challenges. A rigorous methodology and a clear theoretical understanding are fundamental to creating a composite indicator that is useful, accurate, and representative of the multi-dimensional phenomenon under examination.

In this paper, to synthesize the individual indicators into a single measure, the *Adjusted Mazziotta-Pareto Index* (AMPI) is used (Mazziotta-Pareto, 2016), a partially non-compensatory composite index based on a standardization of individual indicators which makes the indicators independent of the unit of measurement (De Muro *et al*., 2011). This summary measure is designed to rescale individual indicators in the range (70; 130) according to two "goalposts," i.e., a minimum and a maximum value representing the possible range of each variable for all periods and all units. This index makes it possible to measure, in a synthetic way, complex and multidimensional phenomena in space and time ensuring robustness.

The AMPI is calculated for the 14 metropolitan city capitals (Turin, Milan, Venice, Genoa, Bologna, Florence, Rome, Naples, Bari, Reggio di Calabria, Palermo, Messina, Catania, and Cagliari) between 2016-2022 years. The comparison at territorial level is facilitated because 100 represents the reference value of Italy in 2016; values higher than 100 indicate an advantageous situation (high sustainable mobility), while lower values indicate a disadvantageous situation (low sustainable mobility).

5. Analyses of results

The following section reports the results for each domain and for the overall AMPI that considers the three domain together.

5.1. Private Motorization

The domain "Private Motorization" take into account four individual indicators: motorization rate for car and motorcycles, percentage of cars with low-emission, percentage of cars with high-emission (Euro 4 or lower) and the index of pollutant potential of cars. Figure 1 shows the values of AMPI and ranking for metropolitan city capital and geographic areas in the period 2016-2022. Values greater than 100 of AMPI indicate a low use of private motorization (sustainable mobility higher than Italy in 2016), lower values indicate a high use.

The AMPI value for Italy showed a steady increase over the period considered, growing from 100 in 2016 to 112.26 in 2022. This indicates that policies aimed at promoting green motorization have had a positive impact, leading to improved sustainable mobility over time. Regarding to metropolitan cities capitals, over the period considered, the top six - Bologna, Venice, Florence, Milan, Turin and Genoa - always maintain the same position. Until 2019, the seventh position was held by Bari (the only southern city with values consistently above 100) that passed to Rome from 2020. Palermo and Cagliari always maintain the ninth and tenth positions, respectively. The 11th position was held by Messina until 2019, after which it was taken by Naples. Reggio di Calabria occupied twelfth place for the first three years, then was overtaken by Naples in 2019 and later by Messina, ultimately falling to thirteenth place. Catania remained in last position throughout the entire period.

Figure 1 – *Domain "Private Motorization": AMPI and ranking of metropolitan cities capitals. Years 2016-2022.*

	2016		2017		2018		2019		2020		2021		2022	
Territory	AMPI	Ranking	AMPI	Ranking	AMPI	Ranking	AMPI	Ranking	AMPI	Ranking	AMPI	Ranking	AMPI	Ranking
Turin	103.46	5	104.66	6	107.32	5	110.11	5	113.00	5	116.43	5	118.33	5
Milan	104.85	4	107.29	4	109.63	4	112.36	4	114.67	4	117.72	4	120.54	3
Venice	109.96	2	112.54	2	114.86	2	117.60	2	119.75	2	122.60	2	124.88	2
Genoa	102.57	6	104.69	5	106.62	6	108.44	6	110.90	6	114.17	6	116.78	6
Bologna	115.14	1	117.34	1	119.59	1	121.94	1	124.27	1	126.88	1	128.73	1
Florence	107.52	3	109.66	3	111.23	3	112.79	3	114.96	3	117.97	3	119.95	4
Rome	99.35	8	102.10	8	104.85	8	107.16	8	109.40	7	112.23	7	114.34	7
Naples	85.83	13	89.18	13	92.01	13	94.71	12	96.52	11	99.02	11	101.10	11
Bari	101.30	7	103.56	7	105.50	7	107.28	7	109.25	8	112.03	8	111.33	8
Reggio Calabria	88.79	12	90.60	12	92.22	12	93.68	13	94.70	13	96.59	13	98.08	13
Palermo	93.50	9	95.53	9	97.28	9	98.83	9	100.03	9	101.94	9	103.71	9
Messina	89.84	11	91.63	11	93.20	11	94.88	11	95.82	12	97.51	12	99.21	12
Catania	78.54	14	80.67	14	82.15	14	83.07	14	84.50	14	86.58	14	87.99	14
Cagliari	90.69	10	92.49	10	94.19	10	95.85	10	97.32	10	99.73	10	101.49	10
ITALY	100.00		102.15		104.08		106.02		107.84		110.32		112.26	

Source: based on Istat data

5.2. Public Transportation

The domain "Public Transportation" consider three individual indicators: demand for local public transportation, availability of buses for local public transport and total seat-kilometers offered by local public transport. Figure 2 highlights, the level of public transportation for each metropolitan city in the period 2016-2022. Values greater than 100 indicate above-average public transport use; lower values display a low use. Between 2016 and 2022, Venice presents the highest level of AMPI, the lowest value (119.55) recorded in 2020 due to the Covid-19 pandemic, while the highest value, it was registered in 2019 (126.66).

Messina ranked last in 2016 and 2019 with an AMPI value of 85.16 and 85.99, but showed a slight improvement over the period considered, reaching to 88.89 in 2022 and surpassing Palermo, Reggio di Calabria and Naples. Palermo ranked last in 2017, 2018, 2020, 2021 and 2022, with an AMPI value of 87.36 in 2017 that further declined to 85.47 in 2021, even if registers a slight improvement to 86.42 in 2022. Venice and Milan break away from the other cities, indicating a more intensive use of public transportation. A clear division emerges between metropolitan city capitals in the North and Center and those in the South. The only exception is Cagliari, which alternates between third and fourth place in the ranking, with AMPI values ranging from a minimum of 108.20 to a maximum of 111.38.

Figure 2 – *Domain "Public Transportation": AMPI and ranking of metropolitan cities capitals. Years 2016-2022.*

	2016		2017		2018		2019		2020		2021		2022	
Territory	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**
Turin	108.21	6	108.57	5	106.77	6	108.97	5	103.33	5	100.74	7	103.39	7
Milan	120.38	2	121.79	2	121.85	2	123.33	2	112.70	2	119.73	2	120.28	2
Venice	**126.48**	**1**	**126.62**	**1**	**126.37**	**1**	**126.66**	**1**	**119.55**	**1**	**123.37**	**1**	**124.50**	**1**
Genoa	105.14	8	105.47	7	104.74	7	105.29	7	99.67	8	100.49	8	102.01	8
Bologna	105.56	7	105.10	8	104.70	8	104.60	8	101.23	7	102.84	6	103.91	6
Florence	108.57	5	109.18	4	111.16	3	112.32	3	107.50	4	110.02	4	110.81	3
Rome	110.78	3	108.33	6	107.20	5	107.63	6	102.76	6	105.36	5	105.72	5
Naples	88.59	11	91.57	11	93.06	11	88.97	11	86.66	12	87.68	12	86.94	13
Bari	95.31	9	93.89	9	94.55	9	94.78	9	93.21	9	93.61	9	93.47	9
Reggio Calabria	88.49	12	88.64	12	88.06	12	87.92	12	88.23	11	88.42	11	88.59	12
Palermo	86.68	13	**87.36**	**14**	**87.16**	**14**	86.85	13	**85.73**	**14**	**85.47**	**14**	**86.42**	**14**
Messina	**85.16**	**14**	87.57	13	87.50	13	**85.99**	**14**	86.58	13	87.26	13	88.89	11
Catania	92.11	10	93.49	10	93.95	10	94.06	10	92.07	10	92.76	10	90.88	10
Cagliari	109.09	4	109.51	3	109.61	4	111.38	4	108.20	3	110.48	3	110.26	4
ITALY	**100.00**		**99.99**		**99.87**		**100.03**		**96.74**		**98.28**		**98.83**	

Source: based on Istat data

5.3. Active mobility and sharing

Regarding the domain "Active mobility and sharing" (Figure 3) that considers three individual indicators (density of bicycle paths, availability of vehicles for car sharing and bike sharing services, scooter sharing and electric micromobility services), an AMPI value higher than 100 indicates above-average use (relative Italy in 2016) of active mobility and sharing services (car, bike, scooter sharing and electric micro mobility), while lower values indicate the opposite.

Figure 3 – *Domain "Active mobility and sharing": AMPI and ranking of metropolitan cities capitals. Years 2016-2022.*

	2016		2017		2018		2019		2020		2021		2022	
Territory	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**
Turin	118.82	2	128.59	3	126.38	3	124.57	3	122.86	4	125.06	3	132.74	2
Milan	**128.30**	**1**	**140.58**	**1**	**144.96**	**1**	**141.12**	**1**	**144.38**	**1**	**143.77**	**1**	**149.43**	**1**
Venice	98.32	7	98.38	9	99.15	8	99.37	8	104.76	5	104.96	6	105.17	6
Genoa	94.39	11	95.07	11	95.03	12	95.34	11	96.02	10	97.78	9	98.45	9
Bologna	104.45	4	104.54	4	116.54	4	120.03	4	124.28	3	127.13	2	128.44	3
Florence	115.77	3	129.99	2	130.42	2	132.43	2	133.84	2	119.30	4	126.52	4
Rome	101.07	5	102.07	5	103.21	5	104.24	5	104.06	7	107.48	5	108.40	5
Naples	95.39	10	94.89	12	95.31	11	95.22.	12	95.58	11	96.74	10	96.75	12
Bari	96.83	9	96.69	10	95.97	10	96.33	10	96.51	9	96.51	11	97.28	10
Reggio Calabria	**93.45**	**14**	**93.66**	**14**	**94.21**	**14**	95.02	13	95.32	13	95.59	12	94.57	13
Palermo	98.41	6	99.20	7	99.48	7	99.71	7	100.01	8	99.70	8	100.22	8
Messina	93.78	12	93.78	13	94.56	13	**94.56**	**14**	**94.56**	**14**	**94.73**	**14**	**93.78**	**14**
Catania	93.72	13	98.71	8	97.56	9	96.61	9	95.36	12	95.47	13	97.16	11
Cagliari	97.57	8	99.57	6	101.13	6	102.76	6	104.74	6	103.80	7	104.79	7
ITALY	**100.00**		**102.00**		**103.22**		**102.85**		**102.87**		**103.33**		**104.74**	

Source: based on Istat data

From 2016 to 2022, Milan has the highest level of AMPI increasing over time, from 128.30 in 2016 to 149.43 in 2022. Over the seven years of analysis, three metropolitan cities alternate the second place in the ranking: Turin (2016 and 2022), Florence (from 2017 to 2020) and Bologna (2021). Between 2016 and 2018, the city that makes the least use of active mobility and sharing services is Reggio Calabria, while since 2019 Messina occupies the last place. Despite being below the threshold value, both show a slight improvement until 2021, while in 2022 the AMPI values decreas. Rome, the most populated city, maintains the fifth position in the years 2016, 2017, 2018, 2019, 2021 and 2022, dropping to seventh in 2020.

Respectively the AMPI value increased from 101.07 in 2016 to 108.40 in 2022, indicating a more intensive use of active and sharing mobility.

5.4. Overall AMPI

The overall AMPI considers the three domain together, providing a synthetic measurement of sustainable mobility in the metropolitan cities capitals. The results show that in Italy, the AMPI values constantly increase between 2016 and 2022, indicating an increasing sustainable mobility over time. The AMPI value grow up from 100 in 2016 to 104.99 in 2022. Milan presents the highest level of AMPI over the period considered, the lowest value (117.04) is recorded in 2016 while the highest value is registered in 2022 (128.64). Catania is always in the last position, with an AMPI value of 87.60 in 2016 and 91.85 in 2022. Among the southern and island cities, the only metropolitan city that presents values above 100 are Cagliari, starting from 2018 and Bari in 2021 and 2022. The overall AMPI confirms a clear distinction between the metropolitan cities in the North-Central and in the South island.

Figure 4 – *Overall AMPI and ranking of metropolitan cities capitals. Years 2016-2022.*

	2016		2017		2018		2019		2020		2021		2022	
Territory	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**	**AMPI**	**Ranking**
Turin	109.79	4	112.97	3	112.76	4	114.11	4	112.50	5	113.19	5	116.94	5
Milan	**117.04**	**1**	**121.71**	**1**	**123.44**	**1**	**124.49**	**1**	**122.22**	**1**	**125.97**	**1**	**128.54**	**1**
Venice	110.39	3	111.33	4	112.36	5	113.42	5	114.26	4	116.36	3	117.47	4
Genoa	100.49	7	101.52	7	101.88	7	102.72	8	101.80	8	103.65	8	105.15	8
Bologna	108.17	5	108.67	5	113.25	3	115.00	3	115.58	3	117.86	2	119.24	2
Florence	110.50	2	115.47	2	116.91	2	118.44	2	117.73	2	115.82	4	118.75	3
Rome	103.49	6	104.08	6	105.06	6	106.32	6	105.33	6	108.28	6	109.36	6
Naples	89.76	12	91.82	11	93.44	11	92.88	11	92.71	11	94.22	11	94.56	11
Bari	97.75	9	97.88	9	98.43	9	99.15	9	99.18	9	100.07	9	100.11	9
Reggio Calabria	90.18	11	90.92	12	91.42	13	92.10	12	92.64	12	93.39	12	93.58	13
Palermo	92.61	10	93.77	10	94.34	10	94.77	10	94.78	10	95.15	10	96.20	10
Messina	89.46	13	90.92	13	91.65	12	91.53	13	92.14	13	92.96	13	93.77	12
Catania	**87.60**	**14**	**90.33**	**14**	**90.74**	**14**	**90.87**	**14**	**90.42**	**14**	**91.45**	**14**	**91.85**	**14**
Cagliari	98.54	8	100.04	8	101.25	8	102.94	7	103.22	7	104.48	7	105.39	7
ITALY	**100.00**		**101.37**		**102.36**		**102.91**		**102.28**		**103.74**		**104.99**	

Source: based on Istat data

6. Conclusions

Realizing that urgent measures and new models of governance are needed, with our MOSER "Sustainable And Resilient MObility" project we intend to address issues related to the current spread of people's modes of travel, with particular reference to private and collective mobility and delving into the issues of road safety and smart mobility. To provide new insights that are also useful to public decision makers.

Through the process of constructing an environmental sustainability index, we have identified basic, individual indicators for sustainable and resilient mobility, which operationalize these concepts. This intricate task involved thorough research and analysis to ensure that the indicators accurately reflect the various dimensions of sustainable mobility. Thanks to the availability of official statistical data, we identified three domains within which the analyses are robust and consistent: mobility infrastructure, public transport availability, and environmental impact. At the current state of the analysis, the work done shows that, in recent years, there has been increased interest in sustainable mobility issues, ranging from green and sharing policies to a growing awareness of the need for enhancing public transport.

This heightened interest is partly driven by global climate change concerns and the need to reduce urban congestion and pollution. However, a significant challenge we face is the limited availability of local-level data and the lack of long-term time series, which are crucial for tracking progress and trends over time.

The results of this analysis between 2016 and 2022 show a varied picture but with clear trends for Italy's major metropolitan cities. Overall, Italy has seen a steady improvement in sustainable mobility, as evidenced by the increase in AMPI values in all three domains analyzed: Private Motorization, Public Transport, and Active and Shared Mobility. This suggests that policies and investments to promote greener alternatives to private transportation are having a positive impact nationwide.

Specifically, for Private Motorization, there is clear success in initiatives to encourage the use of low-emission vehicles, with Italy moving progressively toward greater sustainability. Northern cities such as Bologna, Venice, Florence, Milan, Turin and Genoa are confirmed as leaders in this area. As for Public Transport, Venice and Milan stand out for intensive use, albeit with fluctuations due to the Covid-19 pandemic, while the gap between North/Central and South remains significant, with the exception of Cagliari. Finally, the domain of Active and Shared Mobility sees Milan emerge as an excellence, demonstrating strong adoption of services such as car, bike and scooter sharing, and electric micromobility. In summary, while the country as a whole is progressing, a clear geographic dichotomy remains, with cities in the North and Center leading the transition to sustainable mobility. The cities of the South and Islands, while showing some signs of improvement (such as Cagliari and, to a lesser extent, Bari), are generally in more backward positions, with Catania remaining at the tail end.

These disparities highlight the need for targeted policies and specific investments to reduce the gap and promote more equitable and sustainable mobility throughout the country. The comparative model thus obtained allows the identification of areas where good practices need to be extended and can be replicated on a larger scale.

The pandemic may serve as a pivotal moment, potentially marking a dividing line between an Italy that is "not very green" and an Italy that is eager to adopt more sustainable travel practices. During the pandemic, there was a noticeable shift towards new modes of travel, such as scooters, bicycles, and car sharing, alongside efforts to enhance public transport. These changes were driven by the necessity to maintain social distancing and reduce reliance on crowded public transport, thereby accelerating the adoption of greener travel alternatives.

The project will continue with the analysis of two domains, Land and Environment and Road Incidentally. The Territory and Environment domain will explore how land use and environmental factors influence mobility patterns, while the Road Incidentally domain will examine safety aspects (in relation to deaths and injuries in accidents). A synthesis of all domains will then be conducted to provide a comprehensive assessment of mobility in Italy.

This holistic approach will allow us to identify areas where sustainable mobility can be further promoted and highlight best practices that can be adopted nationwide. Analytical focuses are being developed where data availability allows, ensuring that each domain is thoroughly examined. Additionally, we are implementing a set of simple indicators for each domain to facilitate comparison and monitor progress over time.

The ultimate goal of the MOSER project is to promote a process of cultural change that highlights the environmental benefits achievable through more sustainable mobility. By demonstrating the positive impacts on air quality, public health, and urban livability, the project seeks to drive investment in local development and spatial planning policies that support sustainable practices.

Promoting a transition to greener mobility options not only helps in combating climate change but also enhances the quality of life for residents by creating more livable, efficient, and connected communities. Through education, policy advocacy, and collaboration with local governments and stakeholders, we strive to foster a collective commitment to sustainable mobility and help pave the way for a greener future. In line with what has been stated so far, significant investments in infrastructure and monitoring systems are desirable.

References

ACI, 2024. Note Open Parco Veicoli. https://opv.aci.it/WEBDMCircolante/noteOPV.html

DE MURO P., MAZZIOTTA M., PARETO A. 2011. Composite Indices of Development and Poverty: An Application to MDGs. *Social Indicators Research*, Vol. 104: pp.1-18.

DANIELIS R., MONTE A., ROTARIS L. 2017. *Indicatori di sostenibilità della mobilità urbana. Un'applicazione all'Italia e al Friuli Venezia Giulia*.

ISFORT, 2024. *Rapporto sulla mobilità degli italiani.*

ISPRA, 2024. Mobilità sostenibile: strumenti e indicatori https://www.isprambiente.gov.it/it/servizi/mobilita-sostenibile/strumenti-e-indicatori

ISTAT, 2022. Ambiente urbano. *Nota metodologica* https://www.istat.it/wp-content/uploads/2023/12/Nota-metodologica.pdf

MAZZIOTTA M., PARETO A. 2016. On a Generalized Non-compensatory Composite Index for Measuring Socio-economic Phenomena. *Social Indicators Research,* Vol. 127, No. 3, pp.983-1003.

PIERALICE E., TREPIEDI L. 2015. *Mobilità e Dimensione Urbana. Una Visione Critica: Le Città nella Città*

OECD, 2008. The OECD-JRC, Handbook on Constructing Composite Indicators. *Report*

KYOTO BLUB e CNR, 2018. *Rapporto Mobilitaria* https://iia.cnr.it/mobilitaria-2018-2/

BAILOT M., CECCHINI A.M., FIORONI L., GROSSI P., OSTI S., QUONDAMSTEFANO V., SPINELLA V., 2024. Indicatori della Mobilità Sostenibile in Italia: Città Metropolitane a Confronto, *Istat Poster in XV Conferenza nazionale di statistica* https://www.istat.it/storage/15-Conferenza-nazionale-statistica/poster/06_04_Spinella_POSTER.pdf

Monica BAILOT, Istat, bailot@istat.it
Anna Maria CECCHINI, Istat, amcecchini@istat.it
Livia FIORONI, Istat, fioroni@istat.it
Patrizia GROSSI, Istat, grossi@istat.it
Susi OSTI**,** Istat, osti@istat.it
Valeria QUONDAMSTEFANO, Istat, quondamstefano@istat.it
Valentina SPINELLA, Istat, vspinella@istat.it

INTEGRATED ANALYSIS OF SOCIOECONOMIC INEQUALITIES IN CONTAMINATED SITES: A CASE STUDY [1]

Daniela Vacca, Francesca Abate, Adriano Cabras, Enrico Olla, Giovanna Pala, Stefano Tersigni, Luigi Minerba

Abstract. The study presents a prototype framework for integrated analysis aimed at assessing socioeconomic and demographic profiles in Sites of National Interest (SIN), applied to the Sulcis-Iglesiente-Guspinese SIN site (in south-western Sardinia). The developed model combines demographic, economic, educational attainment, employment data and is structured on two territorial levels: sub-municipal, through the aggregation of enumeration areas, and supra-municipal, through clusters of municipalities, to represent different scales of pollutant exposure. The findings support the creation of spatial profiles to assist territorial planning and policymaking aimed at reducing social and health inequalities.

1. Introduction and objectives of the study

Across the world, the remediation and management of contaminated sites represent a pressing challengefor environmental policy, public health, and territorial governance. These areas often present multiple issues where ecological degradation intersects with economic stagnation and social vulnerability, thereby amplifying existing inequalities (Pasetto *et al*., 2019). Contaminated sites therefore represent contexts of particular concern for public health and social cohesion, since exposure to pollutants frequently overlaps with socio-economic disadvantage. International research on environmental justice has shown that the most vulnerable communities are often those most exposed to environmental pressures, leading to significant health inequalities (Bullard, 2000; Brulle and Pellow, 2006). In Italy, the SENTIERI epidemiological project has documented excess mortality and morbidity among populations residing in Sites of National Interest (SIN), i.e. contaminated areas designated for environmental remediation (see Section 2), with particularly severe impacts on socially disadvantaged groups (Pirastu *et al*., 2013; SENTIERI VI, 2023).

[1] This paper has been jointly written by all the authors but §1 and §3 can be attributed to Daniela Vacca, §2 to Stefano Tersigni, §4 and §5.1 to Giovanna Pala, §5 to Adriano Cabras, §5.2 and §5.3 to Enrico Olla, §6 to Luigi Minerba, §7 to all authors. We sincerely thank Dr. Francesca Abate for her valuable assistance with the textual, critical, and methodological revision of the manuscript.

Within this perspective, this study aims to develop and test a prototype analytical framework applied to contaminated territories, combining demographic, social, economic, health, and environmental data[2]. The framework is designed to provide evidence supporting institutional governance and to generate indicators useful for assessing population exposure and informing territorial planning. From a spatial standpoint, an innovative operational approach is proposed for SINs, structured as a two-level geographical framework within each site: the sub-municipal level allows detailed evaluation of local dynamics and population exposure near pollution sources, while the supra-municipal level groups municipalities by spatial proximity and demographic relevance, forming statistically robust territorial units that support health and environmental governance (Vacca *et al.*, 2025).

The analysis focuses on the Sulcis-Iglesiente-Guspinese SIN, in south-western Sardinia, a historically mining- and industry- affected area that remains characterised by strong social and environmental vulnerabilities. The results reported provide the first findings of the study, aimed at building knowledge frameworks and producing multi-thematic information profiles.

The following sections present the context of SIN, environmental justice and inequalities, geographical levels of analysis, selected indicators, case study results, and future research prospects.

2. Sites of National Interest (SIN) and their management

In Italy, SIN are severely contaminated areas whose environmental damage has a direct impact on human health and ecosystems, requiring complex and particularly costly remediation processes[3]. Their management is entrusted to the State, given the significance of the environmental issues involved and the need for centralised coordination of interventions. Pollution in SINs generally results from prolonged industrial activity, uncontrolled waste disposal and environmental accidents, often dating back to periods when environmental regulations were absent or poorly enforced. The health and environmental consequences of pollution in SINs are numerous and serious: exposure to toxic substances, through inhalation, ingestion

[2] This contribution is part of the institutional collaboration between Istat – Central Directorate for Environmental and Territorial Statistics (DCAT) – and the Department of Medical Sciences and Public Health of the University of Cagliari. The collaboration aims to carry out study on specific territorial areas, focusing on the interactions between environmental, health and socio-economic conditions.

[3] The main contamination sources are chemical/petrochemical plants releasing hydrocarbons, metals and persistent compounds, and steel/metallurgical sites with lead, chromium, nickel, arsenic and asbestos; illegal landfills add solvents, pesticides, pharmaceuticals and emerging pollutants; mining activities, which result in heavy metals and acids due to acid drainage; abandoned military and industrial sites, where explosives, fuels, asbestos and other hazardous agents are found.

(including through the food chain) and skin contact, is associated with the onset of chronic diseases, cancers, neurological and reproductive disorders. At the same time, significant damage to natural ecosystems is observed, including biodiversity loss, water contamination, and soil fertility reduction. The economic impact is also significant: property devaluation, restrictions on land use, loss of tourist appeal and high remediation costs are reported. The complexity of SINs also depends on their territorial extent (sometimes exceeding tens or hundreds of hectares), the variety and persistence of contaminants, the depth of pollution (which can reach aquifers) and the simultaneous presence of different contaminated environmental matrices (soil, subsoil, water and sediments).

The regulatory reference for the remediation of contaminated sites is Legislative Decree No. 152 of 3 April 2006, known as the "Consolidated Environmental Act", which in Article 252 assigns to the Ministry of the Environment and Energy Security (MASE), in agreement with the Ministry of Health, responsibility for the management of SINs, identified on the basis of environmental and health criteria.

3. Environmental justice and inequalities in contaminated contexts

The exposure of resident populations to environmental pollution in contaminated areas is unevenly distributed, often along a socioeconomic gradient, resulting in higher exposure among the most vulnerable communities. This phenomenon, defined as environmental injustice, emerges from the interaction of environmental, economic and institutional factors that lead to a higher concentration of contaminated sites in socially and economically disadvantaged areas, accompanied by a reduced capacity of exposed groups (Bullard, 2000; Brulle and Pellow, 2006).

In SINs, environmental injustice is expressed in the association between greater exposure to pollutants and socio-economic and health disadvantages, particularly for low-income and less educated populations living near contaminated sites (Martuzzi *et al.*, 2010; Pasetto *et al.*, 2019). Additional vulnerability factors, such as limited mobility and restricted access to health services, increase these risks. Economically disadvantaged communities are less likely to move away from polluted areas and often live in locations with poor health and social services, hindering the early diagnosis and management of exposure-related diseases. This situation is observed in some Italian SINs, where populations living near contaminated sites exhibit higher incidence and mortality for respiratory, cardiovascular, and oncological diseases (Pirastu *et al.*, 2013). The SENTIERI VI project (2023) has systematically documented these inequalities, highlighting higher incidence and mortality for the same diseases in SINs, particularly among socially disadvantaged populations. Environmental justice also involves limited participation of local communities in decision-making processes related to remediation and environmental risk management. Decisions are often made without adequate involvement of the most exposed populations, thus reducing the possibility of influencing mitigation strategies and improving living conditions.

In the Sulcis-Iglesiente-Guspinese SIN, the communities most exposed to environmental contamination also have high levels of social deprivation and poorer health outcomes (SENTIERI VI, 2023). This SIN is among the largest and most contaminated in Italy, affected by metallurgical, mining, and other industrial activities. From an administrative-territorial perspective, the SIN includes multiple Health Districts (HDs) across three Local Health Authorities (LHAs): specifically, the Carbonia, Iglesias, and Isole Minori HDs constitute the Sulcis-Iglesiente LHA; the Guspini HD belongs to the Sanluri LHA; and the Cagliari Area Ovest HD is part of the Cagliari LHA[4].

4. Geographical level of analysis

The geographical framework adopted for the socio-economic analysis of SINs is based on a two-tiered model, designed to identify and classify the areas of influence of contaminated sites. It consists of a sub-municipal level, which allows the identification of portions of municipal territory affected by contamination and provides a fine-grained analysis of population exposure near pollution sources, and supra-municipal level, which aggregates multiple municipalities into statistically robust and coherent units, capturing broader territorial patterns and supporting institutional planning (Vacca *et al.*, 2025).

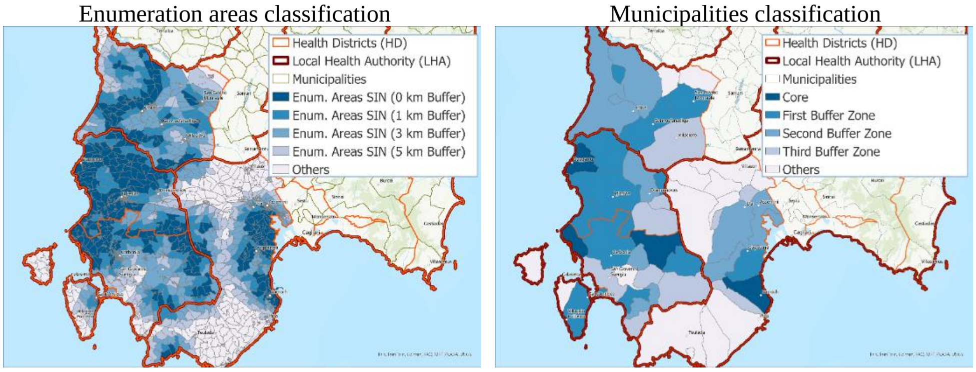
At the sub-municipal level, the territorial reference units are the 2021 Istat enumeration areas, providing the finest resolution available for the spatial interpretation of the resident population (Figure 1). To assess potential environmental impact, buffers of 0, 1, 3 and 5 km were defined around SIN perimeters. The underlying assumption is that the level of potential risk decreases with increasing distance from the contamination source. In this way, it is possible to estimate the “incidence of the population exposed to environmental risk”, defined as the proportion of residents living within the buffer zones around SIN perimeter, an important indicator to assess the magnitude of the problem.

The construction of the supra-municipal level requires an additional step, based on a combined criterion of spatial proximity and demographic relevance. A municipality is included only if a certain proportion of its population resides within

[4] The Sulcis-Iglesiente LHA takes part in the One Health Citizen Science project, funded by the Ministry of Health, which assesses environmental quality, exposure, and health-social impacts in Italian SINs, promoting participation and environmental justice.

the defined buffers. Municipalities are then classified into five categories: Core (within the SIN, with at least 10% of the population), First buffer zone (from 0 to 1 km, ≥20% of the population), Second buffer zone (from 1 to 3 km, ≥30% of the population), Third buffer zone (from 3 to 5 km, ≥40% of the population) and Others municipalities (beyond 5 km). These classifications ensure consistency across governance levels (from enumeration areas to municipalities, HDs LHAs)- supporting planning and decision-making at every scale.

Figure 1 – *Sub-municipal and supra-municipal territorial level.*



Source: our elaboration on MASE and ISTAT data

5. Sources and indicators

The selection of indicators was guided by the need of building a coherent information framework, compatible with the availability of data. At the sub-municipal level, the use of enumeration areas from the 2021 Istat census allows for high spatial resolution, although the available information remains limited in scope. A more detailed and structured set of indicators, while still expandable, makes it possible to outline a profile of the analysed contexts. Among these, a key measure is the incidence of the exposed population, defined as the percentage of residents living within the SIN buffers and therefore potentially subject to environmental exposure**.** This measure provides a foundational context for interpreting the demographic and socio-economic profiles presented in the following sections, which focus on the Sulcis-Iglesiente LHA within the Sulcis-Iglesiente-Guspinese SIN.

The sub-municipal profile reveals a complex demographic structure, with notable differences across territories. In Carbonia HD, the ageing index is very high across all buffers, with a value of 306.7 in buffers 0 and 5, and reaching 362.9 in buffer 1, while the total dependency index consistently exceeds 64% (Table 1).

Table 1 – *Sub-municipal demographic profile by buffer –Sulcis-Iglesiente LHA, year 2021.*

Territories/buffer	Population	Pop_density	Families	Ageing index	Aged_depend _index	Tot_depend _index	% pop 0-14	% pop 75 +
Carbonia HD	**54,858**	**82.1**	**25,240**	**331.8**	**50.1**	**65.2**	**9.1**	**13.5**
Buffer 0	6,992	30.8	3,162	306.7	48.5	64.3	9.6	13.3
Buffer 1	15,996	116.6	7,461	362.9	51.9	66.2	8.6	13.5
Buffer 3	18,095	82.7	8,484	335.0	50.1	65.0	9.1	13.4
Buffer 5	10,353	121.0	4,556	306.7	49.8	66.1	9.8	14.0
Others	3,422		1,577	313.4	46.1	60.8	9.1	13.2
Iglesias HD	**44,687**	**73.9**	**20,568**	**294.9**	**43.2**	**57.9**	**9.3**	**12.3**
Buffer 0	3,666	9.4	1,718	301.5	43.9	58.5	9.2	11.3
Buffer 1	22,905	172.6	10,702	304.9	44.7	59.4	9.2	12.4
Buffer 3	12,695	229.1	5,851	282.4	40.6	54.9	9.3	12.2
Buffer 5	5,418	200.0	2,295	278.9	42.8	58.2	9.7	12.4
Others	3		2		-	-	-	
Isole minori HD	**19,541**	**280.8**	**9,714**	**333.1**	**51.6**	**67.1**	**9.3**	**15.4**
Buffer 0	961	274.1	440	242.2	40.4	57.0	10.6	12.8
Buffer 1	4,296	473.7	2,049	316.0	51.9	68.3	9.8	16.1
Buffer 3	5,287	200.0	2,491	363.2	52.7	67.3	8.7	14.5
Buffer 5	130	4.3	78	400.0	50.0	62.5	7.7	13.8
Others	8,867		4,656	335.4	52.1	67.7	9.3	15.8
LHA	**119,086**	**88.7**	**55,522**	**318.1**	**47.7**	**62.7**	**9.2**	**13.4**

Source: our elaboration on MASE and ISTAT data

In Iglesias HD, too, the values are high, but slightly lower: the ageing index varies from 278.9 (buffer 5) to 304.9 (buffer 1), while the total dependency index stands between 54.9% and 59.4%. The Isole Minori HD, on the other hand, shows greater variability. In buffer 3, the ageing index reaches its maximum (363.2), associated with a total dependency ratio of 67.3%, whereas in buffer 0 it drops to 242.2, indicating a younger demographic composition. Overall, these data suggest that proximity to the SIN does not entail greater demographic vulnerability, as observed differences may reflect other factors, such as settlement patterns, population density and and age structure.

At the supra-municipal level, a wider range of indicators allows for a more detailed territorial profile, although still partial. The selected indicators are distributed along two main dimensions: socio-demographic conditions, which include the age structure of the population, settlement density, natural and migratory balances and the presence of foreign citizens and economic conditions, with particular attention to the labour market and education (Table 2). The resulting profile allows identification of differences among the districts of the Sulcis-Iglesiente LHA in terms of ageing, population density, and family composition, as well as in employment opportunities and educational attainment.

These are elements which, although awaiting integration with the epidemiological profile currently being developed, already provide a valuable initial descriptive tool for understanding social inequalities and guiding interventions. The analysis of district profiles is the subject of the next paragraphs.

Table 2 – *Indicators at the supra-municipal territorial level.*

Themes	Indicators	Year	Abbreviations	Algoritm	Sources
Demography	Territorial surface	2021	Area (km2)	Area (km2)	Istat
	Population density	2024	Pop. dens.	Area (km2)/Pop.	Istat
	Resident population	2024	Population	Pop.	Istat
	Incidence of exposed population by buffer	2021	Exposure pop. Rate	Pop.in buffers/Total pop.*100	Istat
	Percentage change in resident population	2024/04	% Pop. change 24/04	Pop. 2024/Pop 2004*100-100	Istat
	Ageing index	2024	Ageing index	Pop. 65+ /Pop. 0-14*100	Istat
	Percentage of population aged 85 and over	2024	% Pop. 85+	Pop. 85+/Total pop.*100	Istat
	Natural rate per 1000	2024	Natural rate	Birth-death/Pop.*100	Istat
	Migration rate per 1000	2024	Migration rate	(Immigrants - Emigrants) / Pop.*1000	Istat
Economy	Declared income per capita (MEF)	2023	Income p.capita	Icome/Pop.	MEF
	Density of local units per 1000 inh.	2021	LU/1,000 inh.	Local unit/pop.*1000	Istat
	Density of employees in local units total	2021	Emp. in LU / 1,000 inh.	Emp. In local unit/pop.1000	Istat
	Density of local units per 1000 inh., Industry	2021	LU 1,000 – Industry	Industry Local unit/pop.*1000	Istat
	Density of employees in local units per 1000 inh., Industry	2021	Emp. in LU/ 1,000 – Ind.	Industry emp. In local unit/pop.*1000	Istat
Labour and Education	% population aged 25-64 with low education levels	2022	% 25–64 low edu.	Pop. 25–64 with ≤ lower sec./Pop. 25–64*100	Istat
	% population aged 25-64 with upper secondary education diploma	2022	% 25–64 with diploma	Pop. 25-64 with sec. diploma/Pop. 25-64*100	Istat
	% population aged 25-64 with tertiary education diploma	2022	% 25–64 with tertiary edu.	Pop. 25–64 with tertiary/Pop. 25–64*100	Istat
	Employment rate 25-64 years	2022	Employment rate 25–64	Employed 25-64/Pop. 25-64*100	Istat
	Activity rate 25-64 years	2022	Activity rate 25–64	Labor force 25-64/Pop. 25-64*100	Istat
	Unemployment rate 25-64 years	2002	Unemployment rate 25–64	Unemployed 25-64/Labor force 25-64*100	Istat

Source: our elaboration on MASE and ISTAT data

5.1 Carbonia Health District

The Carbonia HD covers an area of 725 km^2 and has a population density of 73.3 inhabitants per km^2. As of 2024, the area has a population of 53,103, 69% of whom are concentrated in the Core and First zones (Table 3). In the Core area, almost 67% of residents live near polluted sites (exposure pop. Rate 2021). Compared to the district average, the Core and Third areas are demographically stable, as evidenced in particular by the lower values of the 2024/2004 population change and the ageing index. From an economic and income point of view, the first two areas have the best profile, although the Core area is characterized by a robust economic structure, accompanied by the highest declared income in the district.

Employment and education data are also better than the district as a whole, although the Core area is weak in terms of tertiary education attainment. The Second buffer zone, on the other hand, has vulnerable socio-demographic and economic conditions, such as high environmental exposure, marked population decline, high ageing, low income, weak employment and critical educational levels, making it one of the most fragile and vulnerable areas of the system.

Table 3 – *Supra-municipal demographic profile by buffer – Carbonia HD.*

Indicators	Core	First buffer zone	Second buffer Zone	Third buffer zone	Others	Carbonia HD	Sardinia
Popolazione e territorio							
Area (km2)	124	238	70	262	31	725	24,099
Pop. dens. 2024	62.5	121.6	46.2	46.8	31.7	73.3	64.8
Population 2024	7,742	28,880	3,235	12,263	983	53,103	1,561,339
Demography, Economy, Labour and Education							
Exposure pop. rate 2021	66.9	55.6	86.5	80.2	100.0	65.6	38.0
% Pop. change 24/04	-11.2	-14.7	-13.1	-9.7	-11.8	-13.0	-4.7
Ageing index	348	414	400	363	422	390	281
% Pop. 85+	4.0	4.3	4.6	4.3	5.0	4.3	4.2
Natural rate	-0.9	-1.1	-1.3	-0.9	-0.7	-1.0	-0.7
Migration rate	-0.2	-0.4	0.2	0.1	0.2	-0.2	0.1
Income p.c. 2023	13,322	12,945	9,602	11,126	10,571	12,335	13,668
LU / 1,000 inh. 2021	52.7	58.7	39.8	49.3	38.2	54.1	76.1
Emp. in LU / 1,000 inh.	381.3	175.3	70.8	130.7	83.0	186.9	223.7
LU / 1,000 – Ind. 2021	5.6	2.9	2.4	4.1	4.9	3.6	5.2
Emp. in LU/ 1,000 – Ind.	198.0	12.8	5.7	15.2	18.4	39.9	26.4
% 25–64 low edu.	50.2	44.4	58.7	53.6	56.5	48.5	42.6
% 25–64 with diploma	39.3	41.3	33.1	36.0	33.3	39.1	38.2
% 25–64 with tertiary edu.	10.5	14.4	8.2	10.4	10.2	12.4	19.1
Employment rate 25–64	58.1	58.2	52.3	55.9	55.6	57.3	62.6
Activity rate 25–64	66.3	67.2	62.0	64.8	66.4	66.2	71.5
Unemployment rate 25–64	12.3	13.4	15.8	13.7	16.4	13.5	12.5

Source: our elaboration on MASE and ISTAT data

5.2 Iglesias Health District

The Iglesias HD covers an area of 605 km^2 and has a population density of 71.5 inhabitants per km^2 (Table 4). As of 2024, the area has 43,282 inhabitants, of whom almost 76% are concentrated in the Core and First zones. In the Core area, approximately 90% of the population lives in the area closest to the contamination sources (exposure pop. rate 2021). Compared to the district average, from a demographic point of view, the Third buffer zone shows relatively stable demographic indicators, with lower indicator values and a migration balance in line with the district average. The Core area presents critical demographic challenges, despite a positive migration balance. From an income point of view, the highest value is found in the First buffer zone with 13,268 euros. The Core and First areas have the most consistent productive structures in terms of the presence of local units and employees per 1,000 inhabitants, while the industrial structure is stronger in the Second and Third buffer zones. The district generally shows low levels of education, particularly in the Third buffer zone, characterised by a high proportion of people aged 20-64 with only middle school and a relatively low percentage of high school and university graduates. Employment performance is also weaker in the Core area, with an activity rate of 63.5%, employment at 55.2% and unemployment at 13.1%

Table 4 – *Supra-municipal demographic profile by buffer – Iglesias HD.*

Indicators	Core	First buffer zone	Second buffer Zone	Third buffer zone	Others	Iglesias HD	Sardinia
Population and territory							
Area (km2)	48	364	81	112		605	24,099
Pop. dens. 2024	21.0	87.1	71.1	43.0		71.5	64.8
Population 2024	1,015	31,737	5,732	4,798		43,282	1,561,339
Demography, Economy, Labour and Education							
Exposure pop. rate 2021	89.3	72.6	84.7	99.9		77.6	38.0
% Pop. change 24/04	-9.9	-12.3	-12.0	-8.6		-11.8	-4.7
Ageing index	355.9	363	332	311		352	281
% Pop. 85+	4.6	4.2	4.0	4.0		4.2	4.2
Natural rate	-1.3	-1.0	-0.7	-0.3		-0.9	-0.7
Migration rate	0.2	-0.1	-0.5	-0.2		-0.2	0.1
Income p.c. 2023	10,922	13,268	12,010	11,855		12,890	13,668
LU / 1,000 inh. 2021	66.5	57.7	52.9	47.7		56.2	76.1
Emp. in LU / 1,000 inh.	190.6	156.1	177.9	107.9		154.4	223.7
LU / 1,000 – Ind. 2021	4.7	4.2	5.6	6.1		4.6	5.2
Emp. in LU/ 1,000 – Ind.	15.7	19.6	32.8	23.9		21.7	26.4
% 25–64 low edu.	51.3	42.7	50.7	56.4		45.5	42.6
% 25–64 with diploma	38.7	41.7	36.9	32.9		40.0	38.2
% 25–64 with tertiary edu.	10.0	15.6	12.4	10.7		14.5	19.1
Employment rate 25–64	52.9	58.5	59.4	58.8		58.6	62.6
Activity rate 25–64	61.7	68.4	68.1	67.7		68.1	71.5
Unemployment rate 25–64	14.3	14.4	12.9	13.2		14.1	12.5

Source: our elaboration on MASE and ISTAT data

5.3 *Isole Minori Health District*

The Isole Minori HD covers an area of 170 km^2 and has a population density of 118.9 inhabitants per km². As of December 31, 2024, the area had 10,451 residents, with almost 55% concentrated in the First buffer zone. This buffer zone is the only part of the health district to have a population, and it is the smallest of the three Health Districts within the Sulcis-Iglesiente LHA. In the First buffer zone, almost 49% of the population lives within 1 km of polluted sites (exposure pop. rate 2021). The demographic profile is relatively critical, marked by a 10.6% population decline over the last two decades, a high proportion of elderly residents (85 years and older) and negative natural and migratory rates. Socio-economic conditions are also fragile, with an average per capita income of €11,711 and a weaker industrial structure compared to the rest of the area and Sardinia. Education and employment indicators highlight a high share of low educational attainment (46.2% with at most middle school diploma), few high school and university graduates, an activity rate of 63.5%, employment at 55.2%, and unemployment at 13.1%.

Table 5 – *Supra-municipal demographic profile by buffer – Isole Minori HD.*

Indicators	Core	First buffer zone	Second buffer Zone	Third buffer zone	Others	Isole HD	Sardinia
Population and territory							
Area (km2)		88			82	170	24,099
Pop. dens. 2024		118.9			105.5	112.4	64.8
Population 2024		10,451			8,666	19,117	1,561,339
Demography, Economy, Labour and Education							
Exposure pop. rate 2021		48.9			100.0	71.9	38.0
% Pop. change 24/04		10.6			6.5	8.8	4.7
Ageing index		382.0			378.4	380.4	281.4
% Pop. 85+		5.0			5.1	5.1	4.2
Natural rate		1.2			1.0	1.1	0.7
Migration rate		0.2			0.1	0.1	0.1
Income p.c. 2023		11,711			13,437	12,491	13,668
LU / 1,000 inh. 2021		56.3			73.9	64.2	76.1
Emp. in LU / 1,000 inh.		151.8			175.6	162.5	223.7
LU / 1,000 – Ind. 2021		4.2			5.7	4.9	5.2
Emp. in LU/ 1,000 – Ind.		13.1			17.6	15.1	26.4
% 25–64 low edu.		46.2			35.9	41.6	42.6
% 25–64 with diploma		38.8			47.2	42.5	38.2
% 25–64 with tertiary edu.		15.0			16.9	15.9	19.1
Employment rate 25–64		55.2			57.4	56.2	62.6
Activity rate 25–64		63.5			66.2	64.7	71.5
Unemployment rate 25–64		13.1			13.4	13.3	12.5

Source: our elaboration on MASE and ISTAT data

6. Research Prospects

The integration of epidemiological data is essential for a comprehensive assessment of environmental and health risks. In particular, the joint analysis of mortality and hospitalisations enables the identification of health outcomes associated with prolonged environmental exposure, providing a robust basis for regional health planning (SENTIERI VI, 2023). In the Sulcis-Iglesiente-Guspinese SIN, the data indicate a critical health situation: overall mortality, measured using Standardized Mortality Ratios (SMR)[5], is slightly higher than expected for both sexes (101 men, 103 women, Table 6). The most significant excess concerns respiratory diseases (SMR 147 men, 123 women), likely reflecting local exposures, including air pollution, fine particulate, and legacy industrial activities. Hospitalization data show stable rates for all natural causes, with a Standardized Hospitalization Ratio (SHR)[6] of 101 in men and 99 in women; however, significant

[5] SMR is calculated as the ratio of observed to expected deaths in population, adjusted for age and sex.
[6] SHR is calculated as the ratio of observed to expected hospitalizations in the population, adjusted for age and sex.

excesses are observed for respiratory diseases (SHR 111 in men and 109 in women) and urinary diseases (SHR 118 in men and 121 in women). Combining these indicators with socio-demographic information yields a multidimensional picture of risk, linking environmental exposures, health inequalities and social vulnerabilities, reinforcing the need for integrated tools for the management of contaminated areas (SENTIERI VI, 2023). The future goal is to update and integrate these epidemiological data with the new supra-municipal geography based on clusters of municipalities, in order to monitor exposure and health outcomes, supporting health and environmental planning, facilitating targeted interventions and coordinated management of areas at greatest risk.

Table 6 – *Mortality and Hospitalization in Sulcis-Iglesiente-Guspinese SIN- SMR and SHR, value for 100,000 inhabitants. Years 2013-2017.*

Mortality	**Male**		**Female**	
	OBS	**SMR**	**OBS**	**SMR**
General mortality	6,344	101 (99-103)	5,956	103 (100-105)
Diseases of the respiratory system	652	147 (138-157)	420	123 (114-134)
Hospitalization	**Male**		**Female**	
	OBS	**SHR**	**OBS**	**SHR**
All natural causes (excluding complications of pregnancy, childbirth, and the puerperium)	40,200	101(100-102)	39,620	99 (99-100)
Diseases of the respiratory system	6,228	111 (108-113)	4,840	109 (107-112)
Diseases of the urinary system	3,035	118 (114-121)	2,272	121 (117-126)

Source: SENTIERI VI Working Group, 2023

7. Conclusions

The adoption of a two-tiered geographical framework for the analysis of SINs allows for integrated and multi-thematic analysis, particularly useful for describing and assessing socio-economic, environmental and health inequalities in contaminated contexts. The application of this empirical approach to the Sulcis-Iglesiente-Guspinese SIN enabled the construction of fine-scale territorial profiles and the estimation of the environmental hazard gradient for the resident population.

It is also necessary to complete the information framework with additional epidemiological data and indicators of socio-economic deprivation in order to build an even more comprehensive and multidimensional model. This integrated approach will promote a better understanding of environmental and health inequalities and represent a transferable model for other SINs at the national level.

Future work will focus on developing integrated territorial information systems, updating epidemiological indicators, and promoting institutional alliances for participatory governance of at-risk territories.

References

BRULLE R. J., PELLOW D. N. 2006. Environmental justice: Human health and environmental inequalities, *Annual Review of Public Health*, No. 27, pp. 103–124.
BULLARD R. D. 2000. *Dumping in Dixie: Race, class, and environmental quality*, Boulder, Westview Press.
MARTUZZI M., MITIS F., FORASTIERE F. 2010. Inequalities, inequities, environmental justice in waste management and health, *European Journal of Public Health*, Vol. 20, No. 1, pp. 21–26. DOI: 10.1093/eurpub/ckp216.
PASETTO R., MATTIOLI B., MARSILI D. 2019. Environmental justice in industrially contaminated sites. A review of scientific evidence in the WHO European Region, *International Journal of Environmental Research and Public Health*, Vol. 16, No. 6, art. 998. DOI: 10.3390/ijerph16060998.
PIRASTU R., PASETTO R., MEZZOI M., SOFIA M., BELLELLI C., IAVICOLI S., COMBA P. 2013. Environmental justice and inequalities in health in Italy: the epidemiological evidence, *Annali ISS*, Vol.49, No. 3, pp. 250–258.
SENTIERI VI WORKING GROUP 2023. Epidemiological Study of Residents in National Priority Contaminated Sites. Sixth Report, *Epidemiologia & Prevenzione*, Vol. 47. No 1–2 Suppl 1, pp. 375–384.
VACCA D., OLLA E., PALA G., TERSIGNI S., MINERBA L. 2025. An experimental geostatistical approach for defining areas of influence and exposed population in contaminated sites. *Presented at the 61st SIEDS conference*, Poverty and Inequalities: People and territories in the age of transitions, Rome, 28–30 May 2025, session Data and methods for monitoring local disparities (S5.7), 30 May.

Daniela VACCA, Istat, daniela.vacca@istat.it
Francesca ABATE, Istat, francesca.abate@istat.it
Adriano CABRAS, Istat, adriano.cabras@istat.it
Enrico OLLA, Istat, enrico.olla@istat.it
Giovanna PALA, Istat, giovanna.pala@istat.it
Stefano TERSIGNI, Istat, stefano.tersigni@istat.it
Luigi MINERBA, Università degli Studi di Cagliari, luigi.minerba@unica.it

HYBRIDISATION BETWEEN LIVING AND WORKING ENVIRONMENT: EVOLUTION OF PREVENTION AND PROTECTION MEASURES IN RELATION TO POPULATION IMPOVERISHMENT AND THE CENTRAL ROLE OF WORKERS [1]

Michele Camisasca, Eva Pietrantonio, Rosario Magro, Alessandro Arborea, Anna Fabiani, Claudio Giordano, Marcella Pietrantonio, Chiara Tresca

Abstract. The Italian National Institute of Statistics (Istat), a public research institute and the main producer of official statistics, has faced the transformations caused by the increasing spread of remote work, accelerated by the pandemic. The aim of this study is to analyse the hybridisation between living and working environment, exploring the implications of work flexibility, the new requirements for adapting prevention and protection measures as well as the socio-economic impacts on workers' well-being. A multidisciplinary approach has been adopted to carry out this analysis, integrating statistical and regulatory aspects with working conditions assessments. To examine strengths, weaknesses, opportunities, and threats associated with the adoption of hybrid work it has been used a SWOT analysis. Sources include anonymised data from health surveillance visits at Istat, Inail surveys on sector specific occupational accidents and diseases, as well as academic studies on new organisational models and on the impact of economic and social inequalities. The study highlights the evolution of the concept of the "workplace" that now can be applied to domestic and shared spaces, requiring the implementation of new prevention and protection measures to ensure workers' safety, health, and well-being. The analysis assesses the risk of disparities in access to technology and optimal working conditions, which could contribute to increased mental workload and work-related stress. The extension of prevention measures to home environments, the promotion of a culture of prevention that incorporate healthy lifestyles, and the enhancement of training on the risks associated with agile work are some of the strategies proposed to mitigate these effects. Finally, the study underlines the need for inclusive policies to reduce gender, territorial, economic, and cultural inequalities, reinforcing the central role of workers in the transition towards more flexible and sustainable work models.

1. Introduction

The National Institute of Statistics is the main scientific, autonomous institution that produces official statistics. Following the inevitable transition to remote working due to the pandemic (Camisasca *et al.*, 2023a), the Institute has faced a central issue in the current scientific and professional debate on different forms of

[1] This article is the result of the common contribution of all and, therefore, the individual contribution is considered equal and equivalent to that of the other co-authors. The authors' points of view expressed in the article do not necessarily reflect the official opinions of the National Institute of Statistics - Istat.

work flexibility. What began as an experiment before the pandemic emergency has become a necessity that has involved millions of workers (Klaser *et al.*, 2023). In accordance with Title III of the CCNL (National Collective Labour Agreement) for the Education and Research sector 2019-2021, the organisational choices of the Institute have been oriented towards the implementation of remote working, leading to a «*workplace*» concept evolution. The increasing use of remote work with new working methods changes the traditional separation between professional environment and private life and redefine boundaries, habits, and interaction. As work changes, it is essential to achieve a new balance between living and working environments, organising both areas while respecting safety and health regulations (Camisasca *et al.*, 2025). This study aims to analyse the interconnection between these aspects and to explore how the increasing hybridisation between work and private life influences individual behaviours and the very concept of workplace. The evolution of working models requires a new approach to the organisation of lifestyles, promoting strategies that can balance productivity and well-being (Dentini, 2025). Economic, social, and technological transformations act as catalysts for this change. Another crucial aspect concerns productivity, recent studies emphasise both the opportunities and the challenges linked to hybrid work. Barrero *et al.* (2023) show how remote work has evolved into a structural feature of the labour market with measurable effects on productivity and organisational performance. Lee (2023) highlights the broader economic and social changes generated by remote work. Although this paper does not directly measure productivity, acknowledging this issue allows for a more comprehensive understanding of hybrid work as a multidimensional phenomenon.

It is crucial to distinguish between the pandemic phase (2020–2021), when remote working was imposed as an emergency prevention measure and therefore not the result of an individual or organisational choice, and the post-pandemic phase (2022–2024), when hybrid work gradually became institutionalised and workers could choose how to perform their work. This difference strongly affects the assessment of its long-term sustainability.

2. Materials and methods

Legislative Decree 81/08 [2] defines workplace as "*any environment intended to accommodate workstations, located within the company or production unit, as well as any other place accessible to the worker in the course of their work*" and emphasizes the health, safety, and suitability of spaces with the aim of preventing risks and ensuring the workers' well-being.

[2] Italian Legislative Decree of 9 April 2008 n. 81. Text coordinated concerning the protection of health and safety in the workplace.

Table 1 - *Organisational features and prevention measures for a traditional office vs. remote working.*

Aspect	Traditional Office	Remote Working	Prevention Measures
Safety	Direct monitoring of working conditions.	Difficulty in implementing safety conditions at home. Particularly in poorer or socially and economically disadvantaged contexts.	Enhance training and workers' awareness.
Ergonomics	Workstations designed according to legal standards.	Makeshift risk and non-ergonomic workstations, particularly in poorer or socially and economically disadvantaged contexts.	Enhance training and workers' awareness
Health and Well-being	Access to services (e.g., canteens, relaxation areas) with moderate stress risk conditions.	Increased risk of stress, isolation, and sedentary behaviour.	Promote healthy lifestyles and prevention.
Training	In-person sessions with direct supervision.	Distance learning, often asynchronous.	Introduce interactive online courses and periodic discussion sessions.
Supervision	Direct hierarchical control.	Greater trust and autonomy required with increased responsibilities.	Adopt non-invasive system of monitoring and feedback sessions.
Work-Life Interaction	Established and fixed schedules and travels.	Flexibility, but risk of overlapping schedules. Burden from family and personal factors. Work-life balance. Reduction in commuting.	Set clear work-life boundaries and defined hours for work and private life (establish availability times to facilitate collaboration with the team).
Techno-stress	Greater opportunity of carrying out other activities.	Need for advanced technologies. Excessive connection times. Including breaks during disconnection moments.	Impose breaks and promote mindful use of technologies.

Source: Istat, Internal documents.

In recent years, the concept of «*workplace*» has evolved, moving from a traditional physical space to a more fluid and virtual dimension. With the rise of remote working, the traditional model of workspace, separated from the personal sphere, has expanded to include new environments, such as home (smart working), coworking spaces, and even outdoor locations. This change has required the application of safety standards to non-traditional contexts as well. It is essential that the worker choses workplaces that meet the minimum safety requirements, including ergonomics, adequate lighting, and appropriate spaces and workstations. Working in non-company environments requires the redefinition of responsibilities, the extension of risk assessment, and the adoption of appropriate behaviours to ensure safety and productivity. While this transformation offers greater flexibility, it can also accentuate disparities between workers, in relation to the access to technologies, adaptability or specific economic and social conditions. Furthermore, new challenges have to be faced, such as the management of work-related stress (Camisasca *et al.*, 2023b) and social isolation, through prevention measures and support strategies. Thus, it is fundamental to identify and reduce new potential risks and hazards to ensure well-being at work. Nowadays, workplace is an integrated ecosystem that combines physical and digital spaces, fostering continuous interaction between technology, human relations, and well-being. The evolution of this concept, guided by Legislative Decree 81/08, requires an innovative and dynamic approach capable of harmonising safety, flexibility, and quality of life in different work contexts. Table 1 shows a comparison between some organisational characteristics and prevention measures that can be applied both to a traditional office and to remote working. Although the analysis is based on Istat data, the trends observed are consistent with INAIL national surveys and with the international scenario. Therefore, while the outcomes cannot be automatically generalised to all work contexts, they provide useful insights for other public institutions and private organisations facing similar organisational and health and safety challenges.

2.1. SWOT analysis about the hybridisation between living and working environment

The SWOT analysis on the hybridisation between living and working environments (Table 2) examines the Strengths, Weaknesses, Opportunities, and Threats, separating the internal and external organisation features and distinguishing the aspects that can facilitate or hinder the goals achievement. Strengths include, e.g., greater flexibility and a better work-life balance, less commuting by minimising daily travel to and from work, greater schedule autonomy and possibility to personalise one's own work environment. Nevertheless, when agile working is excessive some Weaknesses can emerge as the isolation risk, reduced social interaction due to the lack of in-presence events important for strengthening team spirit, difficult risks monitoring, and the possible overlap between personal and professional spheres. Furthermore, the sense of belonging to the organisation and the connection with colleagues can be compromised, reducing team cohesion and work involvement. Talking about the Opportunities, hybrid work is a real asset and allows for a reorganisation of working methods, offering greater autonomy in activities and time management. This is made possible by the adoption of new technologies, the spaces organisation designed to improve worker's well-being, and the promotion of smart working as an organisational model. An additional advantage is the reduction of company costs, thanks to a less need for physical spaces and travels reducing environmental impact. Finally, the analysis highlights the Threats linked to the adoption of hybrid work, such as the risk of economic impoverishment, i.e., the negative impact that excessive remote work could have on the economy or specific categories of workers, the increase in mental burden for workers, as well as inequalities in access to technologies and telecommunication networks.

Table 2 - *SWOT analysis of the hybridisation between living and working environments.*

	Internal origin (features of the organisation)	External origin (features of the environment)
	Strengths	**Opportunities**
Helpful to achieving the objective	▪ Greater flexibility and work-life balance ▪ Reduction in commuting ▪ Greater autonomy in time management ▪ Opportunity to personalise the work environment	▪ Reorganisation of work with different ways of working ▪ Implementation of new technologies ▪ Adaptation of spaces for well-being ▪ Promotion of remote work as the standard ▪ Opportunity to reduce business costs (offices, transportation)
	Weaknesses	**Threats**
Harmful to achieving the objective	▪ Risk of social isolation ▪ Greater difficulty in monitoring risks ▪ Risk of overlap between personal life and work ▪ Workers feel less connection to the organisation and colleagues, with a reduced sense of company belonging	▪ Economic impoverishment ▪ Increased mental workload for workers ▪ Disparities in access to technology and telecommunications networks ▪ Risk of increased inequalities among workers (e.g., some roles requiring physical presence create disparities between those who can benefit from flexibility and those who cannot); economic differences.

Source: Istat, Internal documents.

Not everyone has the resources to bear the additional costs associated with remote work, such as electricity, connection and equipment, nor do they receive allowances to cover them. Furthermore, the flexibility offered by hybrid work could accentuate the disparities between those who can benefit from smart working and those who are bound to physical presence, with economic and professional inequalities. At a macroeconomic level, changes in consumer behaviour, for example, a reduction in purchases in urban centres, could have negative repercussions on specific geographical areas and economic sectors. It is essential to adopt specific strategies to maximise the benefits of hybrid work, while mitigating its risks, such as interventions aimed at fostering a sense of belonging, improving workplaces safety and ergonomics, and ensuring equitable access to technologies for the hybridisation of work sustainable and effective.

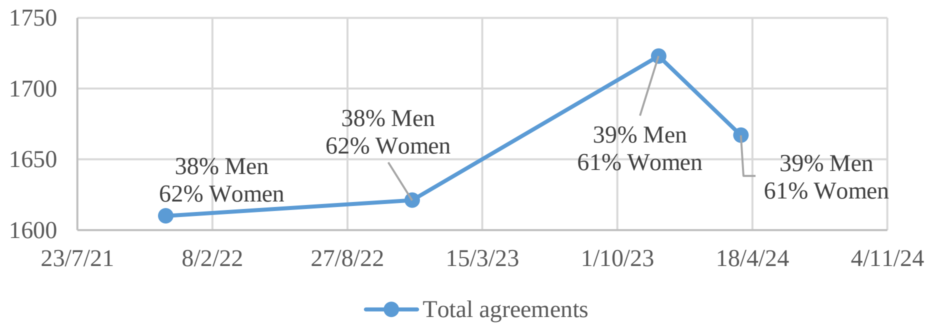
3. Convergence of prevention measures between life and work

According to Law No. 81/2017, Article 18, smart working helps workers to have a better work-life balance and, at the same time, promotes growth. This definition emphasises organisational flexibility, the will of the parties involved, and the use of technological tools that allow remote work (e.g., laptops, tablets, smartphones, etc.). Before the lockdown, Istat tried out agile work that then, during the pandemic, became the primary prevention measure. Agile working has been regulated and structured with the ordinary management - National Collective Labour Agreement 2019-2021 (May-December 2024). As of April 15, 2024, out of approximately 2.000 Istat employees, 1.667 agile working individual agreements have been signed (gender distribution: men=39% and women=61%) [3]. The principle of prevalence regulates work organisation: 51% of working days in-person and 49% in agile work. Furthermore, there are two different profiles of agile working:

1. ordinary agile working: up to a maximum of 20 days on a bi-monthly basis for all employees;
2. enhanced agile working: up to a maximum of 24 days on a bi-monthly basis for employees who document serious, urgent, and otherwise irreconcilable personal or family health situations, including the distance from the workplace.

Figure 1 shows the trend in the total number of individual agreements signed at Istat in the post-pandemic period, from the second half of 2021 to 2024.

Figure 1 – *Number of individual agreements signed in the post-pandemic at Istat.*



Source: Istat, Internal data processing.

A stable growth is observed until January 2024, followed by a decline in the subsequent months. This evolution requires the adoption of appropriate measures to ensure workers' health, safety, and well-being even at home:

[3] LIMITI C., SOLA G., 2024. Agile Working at the National Institute of Statistics. Proceedings CNS15.

- extension of prevention measures to home environment: ensure correct ergonomics, adequate lighting, availability of a first aid kit, and protocols for the proper use of electrical equipment. It is essential that workers take care of their own safety, health, and mental well-being;
- promotion of a healthy lifestyle: encourage regular breaks, a routine of physical exercise, and a balanced time management between private life and work. It is essential to respect the agreed work schedule and adopt habits that promote well-being;
- awareness of common hazards at home and at work and emergency procedures: provide guidelines for the organisation of home spaces and protocols for managing critical situations (fires, short circuits, domestic accidents), with easily accessible emergency numbers.

4. Impact on health, safety and accidents

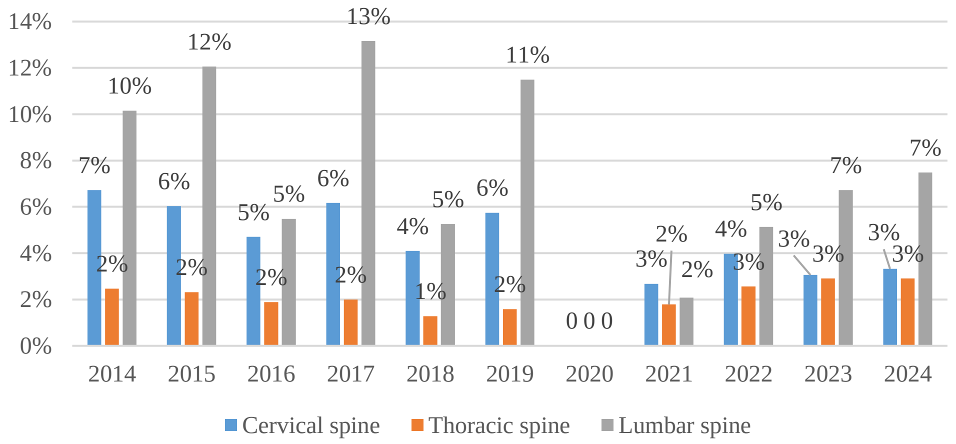
According to INAIL, in the first 11 months of 2024, 543.039 accidents were filed (+0.1% compared to November 2023 and -16,7% compared to the same period in 2022). The increase concerns exclusively commuting accidents. At national level, in the first nine months of 2024, there was a reduction in workplace accidents[4], from 363.064 in 2023 to 361.804 in 2024 (-0,3%), while commuting accidents, i.e., those occurring during the trips between home and work, increased from 67.765 to 71.198 (+5.1%). In the first nine months of 2024, occupational diseases[5] amounted to 65.333, with an increase of 22,0% compared to 2023. The largest increases concern musculoskeletal diseases, followed by those of the nervous system and ear. There is also an increase of tumours and respiratory diseases. Regarding «*domestic accidents*», there is no univocal definition in the National Prevention Plan (PNP) 2020-25[6]. Law No. 493/1999, "*Rules for the protection of health in homes and the establishment of insurance against domestic accidents*" is the regulatory reference. Istat, within the "Multipurpose surveys on Households - Aspects of daily life", defines a domestic accident as *an accidental event that compromises health and occurs within the home or its appurtenances (balconies, garden, garage, stairs, etc.)*.

4.1. Istat data for the period 2014-2024

Figure 2 shows the percentage distribution of musculoskeletal disorders on the total number of individuals monitored in Istat.

[4] https://dati.inail.it/portale/it/tabelle/tabelle-infortuni-sul-lavoro-con-cadenza-mensile.html
[5] https://dati.inail.it/portale/it/tabelle/tabelle-malattie-professionali-con-cadenza-mensile.html
[6] National Prevention Plan 2020-2025. Ministry of Health, Directorate General for Health Prevention. https://www.salute.gov.it/imgs/C_17_notizie_5029_0_file.pdf

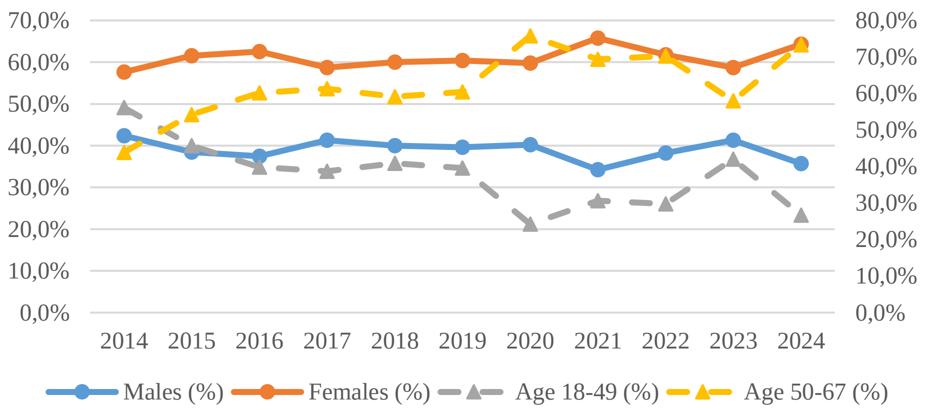
Figure 2 – *Percentage distribution of musculoskeletal disorders on the total number of individuals monitored in Istat.*



Source: Istat, Internal data processing.

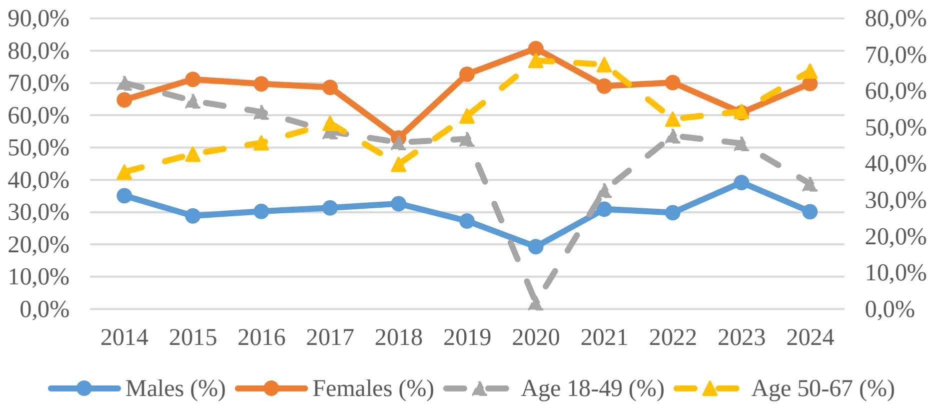
The cervical spine disorders have significantly decreased, from 7% in 2014 to 3% in 2024. Although the absolute number of cases fluctuates, the ratio of reported cases to the monitored population continuously decreases. The lumbar spine disorders are more variable, with peaks of 13% (2017) and drops to 2% (2021). Year 2021 marks a general contraction in all pathologies, with very low percentages. The thoracic spine is the least affected, with values between 1% and 3%. In the 2022-2024 period, the lumbar spine diso1rders values remain between 5% and 7%. Figure 3 shows the percentage of monitored Istat workers, divided by gender and age. The male percentage fluctuates between 30% and 42%, while female percentage represents the majority, with values between 58% and 70%. From 2017 onwards, there is a widening gap between men and women, with more women subjected to monitoring due to the greater female working population. In 2014, the monitored individuals were evenly distributed between the 18-49 age group (56,1%) and the 50-67 age group (43,9%). Since 2016, the percentage of the 50-67 age group has steadily increased, exceeding 60% in 2022 and reaching 73% in 2024. This trend suggests a progressive ageing of the monitored working population. Figure 4 shows the percentage of individuals considered fit for work with limitations, divided by gender and age; as highlighted before, from 2014 to 2024, women have prevailed (values always above 60% and peaks over 70% in the years 2015, 2016, 2019, and 2020). The male percentage has remained lower, between 19% and 39%. In 2020, 80,7% of workers considered fit for work with limitations were women, although the Covid-19 pandemic influenced surveillance activities.

Figure 3 – *Percentage of the number of monitored individuals, divided by gender and age.*



Source: Istat, Internal data processing.

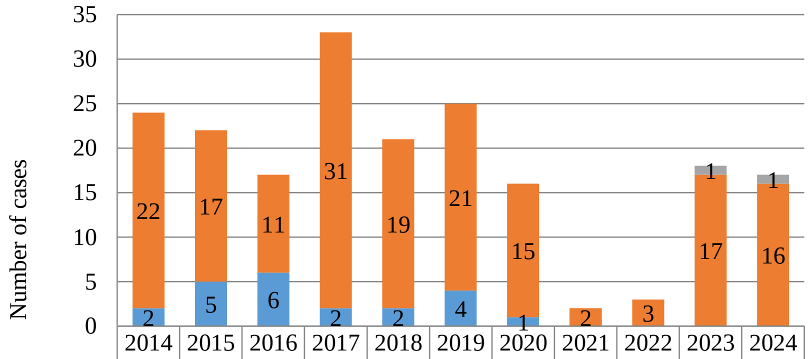
Figure 4 – *Percentage of the number of individuals deemed fit with limitations, divided by gender and age.*



Source: Istat, Internal data processing.

In 2023, the male percentage reached 39,2%, but in 2024 it dropped again to 30,2%. In 2014, most of those considered fit with limitations belonged to the 18-49 age group (62,3%), while 37,7% were in the 50-67 age group. In the following years, the percentage of the 18-49 age group has decreased, while that of the 50-67 age group has increased, reversing the ratio from 2017 onwards. In 2024, 65,5% of workers considered fit with limitations were over 50, confirming the trend of an ageing working population with restrictions. Figure 5 shows accidents data recorded at Istat from 2014 to 2024. Workplace accidents had low values in the period 2014-2019 (with a maximum of 6 cases in 2016). From 2020, these episodes have dropped to zero, following the pandemic, the increased attention to safety, and the spread of agile working. However, with the gradual return to in-person work, commuting accidents are emerging again.

Figure 5 – *Trend of work-related accidents in Istat by categories from 2014 to 2024.*



	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Agile work										1	1
Commuting	22	17	11	31	19	21	15	2	3	17	16
Workplace-related	2	5	6	2	2	4	1	0	0	0	0

Source: Istat, Internal data processing.

Between 2014 and 2019, the trend of commuting accidents showed fluctuations, with a peak in 2017 (31 cases). Starting from 2020, it can be observed a drastic decrease in commuting accidents that hit historical low in 2021 (2 cases), attributable to lockdowns and smart working. There was a rise in the 2023-2024 biennium (17-16 cases), signalling a gradual return to work mobility. Agile working remains the area with the fewest accidents, although individual cases occurred in 2023 and 2024, demonstrating that also remote work can present, even if rarely, safety-related risks. Variations in the percentages should be interpreted with caution, as the analysis is purely descriptive with no control of potential confounding factors (such as age, health status, etc.). Consequently, data cannot be read as evidence of causal relationships but rather as indicative trends for further investigation.

5. Inclusion, well-being and safety policies for sustainable work

The evolution of the world of work requires inclusive policies to reduce gender, territorial, and economic inequalities, strengthening the role of workers in the transition towards more flexible and sustainable models. Disparities in access to healthcare, education, and technology widen the social gap, making investments essential to ensure equity and opportunities. The 2030 Agenda (Goal 10) promotes inclusion as a tool to reduce these inequalities. According to the World Health Organization, *health is a state of complete physical, mental, and social well-being* and it is essential for productivity and quality of life. Healthy lifestyles promotion, smoking reduction, and corporate welfare are fundamental to preventing chronic diseases and improving working conditions. With the increasing median age of the workforce, prevention becomes a priority to ensure active and healthy ageing. Training is crucial for health and safety and should be introduced from an early stage in schools. Recent regulations, such as the Directive of 14 January 2025[7] and the legislation on agile working for SMEs, highlight the need for collaboration between workers and employers. Training is not just an obligation but a strategic opportunity to enhance skills and improve services. In Istat, the strengthening of safety policies has been achieved also through an integrated management of health and safety objectives in the PIAO (Integrated Plan of Activity and Organisation) and in the Risk Management, reinforcing the organisational structure and making prevention policies more efficient. The adoption of more focused strategies makes it possible to protect workers and to optimised risk management.

6. Final considerations

Health and safety in living and working environments are influenced by socio-economic, territorial, technological, and regulatory factors. The development of prevention and protection measures is essential in a context where the boundaries between private life and work are becoming increasingly thinner. The study highlights the following key aspects: the decrease in workplace injuries is offset by a rise in commuting accidents, drawing attention to the need for safe mobility policies; prevention is crucial, from home safety to workplace training; socio-economic and territorial inequalities limit access to training, technology, and healthcare, making inclusive policies essential; digital innovation and AI enhance prevention and personalised care. Moreover, the productivity dimension should not be neglected, remote work is not only a matter of individual choice or organisational flexibility but also a structural factor influencing economic performance and competitiveness (Barrero *et al.*, 2023; Lee, 2023). Finally, new regulations strengthen training and safety, actively involving both workers and employers. Regulatory developments must ensure a dynamic and inclusive approach to safety.

[7] Directive of the Minister for Public Administration of 14.01.2025. Valuing People and Creating Public Value through Training. Principles, Objectives and Tools. https://www.funzionepubblica.gov.it/sites/funzionepubblica.gov.it/files/Direttiva_MinistroPA_14.01.2025_formazione.pdf

References

BARRERO J.M., BLOOM N., DAVIS S.J. 2023. The Evolution of Work from Home. *Journal of Economic Perspectives*, Vol. 37, No. 4, pp. 23–50.

CAMISASCA M., PIETRANTONIO E., MAGRO R., ARBOREA A., FABIANI A., GIORDANO C., PIETRANTONIO M., SPARANO A. 2023. A pandemic health risk management model for the protection of workers: the Istat experience. *The Italian Journal of Economic, Demographic and Statistical Studies*. *RIEDS*. Vol. LXXVII, No.1, pp. 53-64.

CAMISASCA M., PIETRANTONIO E., MAGRO R., ARBOREA A., FABIANI A., GIORDANO C., PIETRANTONIO M., GUERRIERO G., SPARANO A. 2023. Psychosocial risks: work-related stress risk assessment in Istat as a goal for the workers' well-being. *The Italian Journal of Economic, Demographic and Statistical Studies. RIEDS.* Vol. LXXVII, No.4, pp. 181-191.

CAMISASCA M., PIETRANTONIO E., MAGRO R., ARBOREA A., FABIANI A., GIORDANO C., PIETRANTONIO M. 2025. Health and safety management system in the National Institute of Statistics (Istat). *The Italian Journal of Economic, Demographic and Statistical Studies*. *RIEDS*. Vol.LXXIX, No.3, pp. 83-94.

DENTINI A. 2025. Travel analytics for business intelligence: the case of Italian Institute of Statistics, *The Italian Journal of Economic, Demographic and Statistical Studies*. *RIEDS*. Vol. LXXIX, No.2, pp. 211-220.

KLASER K., CUEL R., CASARI P. 2023. The future of hybrid work in Italy: A survey-based Socio-Technical-System analysis. *Journal of Innovation & Knowledge*. Vol. 8, Issue 4, https://doi.org/10.1016/j.jik.2023.100426

LEE K. 2023. Working from home as an economic and social change: A review, *Labour Economics*, Vol. 85.

Michele CAMISASCA, Istat, michele.camisasca@istat.it
Eva PIETRANTONIO, Istat, epietrantonio@istat.it
Rosario MAGRO, Istat, rosario.magro@istat.it
Alessandro ARBOREA, Istat, arborea@istat.it
Anna FABIANI, Istat, anfabian@istat.it
Claudio GIORDANO, Istat, claudio.giordano@istat.it
Marcella PIETRANTONIO, Istat, pietrantonio@istat.it
Chiara TRESCA, External collaborator, Istat, spp@istat.it.

EXPLORING HEALTH DIFFERENCES AMONG WORKERS: INSIGHTS FROM TWO CUORE PROJECT ITALIAN COHORTS

Benedetta Marcozzi, Eleonora Trappolini, Cinzia Lo Noce,
Serena Vannucchi, Luigi Palmieri, Chiara Donfrancesco

Abstract. This study examines sectoral differences in physical and mental health outcomes among Italian workers aged 35–74, using data from two waves of the Italian Health Examination Survey – CUORE Project (2008–2012 and 2018–2019). Four health indicators are analyzed (obesity, depression, work-related stress, and poor self-rated health) across occupational sectors and age groups, separately by gender. Obesity showed age-related patterns varying by sector and gender, with younger men and older women in the commerce-finance-private transport sector at greater risk. Signs of vulnerability also emerged among women in industrial settings. Younger men in the commerce-finance-private transport sector appeared particularly exposed to work-related stress. The results suggest the need to design aging and health policies tailored to specific occupational sectors and gender groups.

1. Introduction and background

Population ageing is profoundly reshaping labour markets across Europe, leading to a growing share of older workers and raising structural concerns about the sustainability of work and labour force participation in later life (Dixon, 2003). In Italy, the demographic shift has contributed to a steady increase in the average age of the workforce (De Rose *et al*., 2019), resulting in more complex career trajectories and increasingly diverse working conditions (Capacci and Rinesi, 2014). In response, various pension reforms have sought to extend working lives (Ministry of Labour and Social Policies, 2025), yet their effects on the health and wellbeing of older workers remain debated.

While some studies suggested that prolonged employment may have health benefit and reduce mortality (Akinwale *et al*., 2011), others argue that extending working life, particularly when driven by economic necessity, can exacerbate health inequalities, especially in contexts with limited welfare redistribution (Baumann *et al.*, 2022). Vulnerable groups, such as women and individuals with pre-existing health conditions, may be disproportionately affected by the pressures of extended employment (Serrano-Alarcón *et al.*, 2023). Moreover, although older workers who remain employed often represent a healthier subset of the population (Sewdas *et al.*, 2020), they still face significant challenges, including both physical and mental strain

(Jones *et al*., 2013). These findings highlight the importance of examining the health profiles of older workers in greater detail.

Among the various health outcomes associated with later-life employment, physical and mental health emerge as key dimensions. Cross-national evidence shows that work stress is linked to depressive symptoms in older employees (Lunau *et al*., 2013), adverse working conditions, such as physically demanding, poorly remunerated, or ungratifying jobs, are consistently associated with poorer mental health, particularly among older workers (Baxter *et al*., 2021; Rugulies *et al*., 2021). Similarly, physical health differences, such as those related to obesity, have been found across occupational categories, with higher prevalence in sectors like healthcare support, administration, and public services, likely due to sedentary tasks combined with work-related stress (Luckhaupt *et al*., 2014).

These health differences appear to vary systematically across the labour market. Employment in Europe and Italy has gradually shifted from agriculture and industry toward the service sector, especially in collective and market services (Eurofound, 2024). In Italy, this has coincided with declining self-employment, rising fixed-term contracts, and an increase in low-skilled service occupations (ISTAT *et al*., 2019). Simultaneously, transformations in technology and work organisation have redefined occupational risks. Automation has altered tasks in industrial sectors, creating new sources of stress and insecurity, particularly for low-skilled workers (Abeliansky *et al*., 2024; Patel *et al*., 2018), while service jobs often combine sedentary routines with high psychosocial demands (Dėdelė *et al*., 2019).

Taken together, these dynamics suggest that occupational sector plays a crucial role in shaping health inequalities among workers. However, little is known about how sectoral differences influence physical and mental health outcomes across different age groups and genders. Despite growing attention to active ageing and extended working lives, as far as we know, no previous research has simultaneously examined multiple health outcomes across occupational sectors with disaggregated analysis by age group and gender. Building on this background, the present study investigates differences in physical and mental health across occupational sectors by age group and gender, among Italian workers aged 35 to 74. Specifically, it explores whether these differences widen with age and differ by gender, shedding light on the role of working conditions in shaping health in later life.

2. Data and methods

2.1 Data sources and study population

This study draws on data from two waves of the Italian Health Examination Survey (HES) – CUORE Project, conducted by the Italian National Institute of

Health as part of the CUORE Project (Istituto Superiore di Sanità, 2025): the 2008–2012 and 2018–2019 surveys. The first wave was carried out in collaboration with the National Association of Hospital Cardiologists (ANMCO) and the Heart Care Foundation (HCF).

The HESs are promoted and partially funded by the Italian Ministry of Health through the National Centre for Disease Prevention and Control. They are part of the National Statistical Programme and follow standardized protocols aligned with the European Health Examination Survey (EHES) framework (European Health Examination Survey, 2025). Laboratory analyses were conducted centrally in national reference centres, and all field staff received standardised training from ISS personnel in accordance with international guidelines (Donfrancesco *et al*., 2021).

The surveys collected detailed information on socio-demographic characteristics, cardiovascular risk factors, health behaviours, and clinical parameters from representative samples of Italian adults aged 35 to 74, randomly selected from municipal population registries. Standardized procedures were used to measure blood pressure, anthropometric measurements, lipid profile, blood glucose, and lifestyles behaviours such as physical activity, smoking, and nutrition.

While the 2008–2012 wave included participants from all 20 Italian regions, the 2018–2019 wave was limited to 10 regions distributed from North, Centre and South of Italy (Piedmont, Lombardy, Liguria, Emilia-Romagna, Tuscany, Lazio, Abruzzo, Basilicata, Calabria, and Sicily). To ensure comparability over time, analyses are restricted to individuals residing in these 10 regions in both waves.

For the purposes of this study, we selected individuals who were currently employed, defined as those who responded "Yes, full-time or part-time" to the question: "Are you currently employed?" The final subsamples included 1,831 individuals (44.3% women) from the 2008-2012 wave and 1,255 individuals (43.3% women) from the 2018-2019 wave.

2.2 Health outcomes

Workers' health status was assessed using four health-related outcomes, each defined by specific criteria.

Obesity was defined as a body mass index (BMI) ≥30 kg/m².

Work-related stress was classified for individuals who reported having jobs involving constant responsibilities and tight deadlines leading to persistent tension, or who indicated low job satisfaction resulting in family-related anxiety.

Depression was identified based on the question: *''Do you know if you suffer from or have ever suffered from depression?''* Respondents answering ''Yes" or ''Yes, but not officially diagnosed" were classified as having depression.

Poor self-reported health (Poor SRH) was defined as a self-assessed health score below 7 on a scale from 1 (poor health) to 10 (excellent health).

These four indicators were selected to capture complementary dimensions of workers' health. Obesity is a prevalent chronic condition among Italian adults (Donfrancesco *et al.*, 2022). Work-related stress and depression address mental health risks commonly linked to adverse psychosocial working conditions. Poor SRH, a widely used global measure, provides a summary assessment of perceived health and wellbeing. Together, these indicators offer a concise but multidimensional profile of health vulnerability across occupational sectors.

2.3 Main explicative and control variables

The main explicative variables are occupational sector and age group. The former is coded according to the International Labour Organization (ILO) classification system, and for analytical purposes was grouped into four broad sectors, as follows: agriculture (S1), industry (S2), services (S3), and commerce-finance-private transport (S4).

The S3 sector includes public or non-profit activities that provide general, technical, or social services, such as healthcare, education, and public administration. In contrast, the S4 sector comprises market-oriented, for-profit activities related to trade, finance, and private logistics.

Age was categorized into three groups: 35–44 years, 45–54 years, and 55 years and over.

In addition, the models were controlled for a set of possible confounding variables, such as socio-demographic variables, the survey wave year (2008–2012 or 2018–2019), the macro-area of residence (North, Centre, or South and Islands), the educational level (which was dichotomised in 'lower education' -elementary or middle school diploma-, and 'higher education' -high school diploma or university degree-), and the marital status (dichotomised in married or cohabiting vs. other - single, divorced or widowed).

2.4 Methods

A descriptive analysis, stratified by survey wave year and sex, was conducted to examine the population characteristics and assess variations across occupational sectors and survey waves (Table 1).

To examine whether health outcomes vary across occupational sector differently by age group, we estimate four logistic regression models, one for each health outcome, including an interaction term between occupational sector and age group (35-44, 45-54, 55+).

Results are presented as predicted probabilities with 83.5% confidence intervals[1] (CI) to facilitate interpretation and ensure comparability across models (Figures 1 and 2).

Due to the small sample size of individuals employed in the agricultural sector (S1), predicted probabilities for this group are not reported in the figures.

All analyses were conducted separately for men and women, given the well-documented gendered patterns in health outcomes and labour market experiences (Di Gessa *et al.,* 2020).

3. Results

The analysis of the 2008-2012 and 2018-2019 survey waves reveals notable changes in the occupational structure of employed individuals aged 35 to 74 years. Employment in the industrial sector (S2) declined in both men and women, with a sharper reduction among men. The service sector (S3), already the most represented in the first wave, further increased its share of employment, especially among women. The commerce-finance-private transport sector (S4) remained relatively stable, with a slightly higher concentration among men in both waves. Employment in agriculture (S1), already marginal, furtherly decreased in both sexes.

Gender differences in health outcomes remain consistent over time. Women reported higher rates of depression and poor SRH in both waves, with a slight increase observed in the second wave. Obesity increased modestly among women, while it slightly declined among men. Work-related stress was more frequently reported by men in both periods; however, its prevalence decreased in the second wave, resulting in a narrowing gender gap.

Figures 1 and Figure 2 show the adjusted predicted probabilities of health outcomes by occupational sector and age group, separately for men and women.

For both sexes, the probability of obesity generally increases with age. However, this trend is less consistent among women in S2, where younger workers (35-44) show the highest predicted value (26%), and among men in S4, where the probability declines from 23% in the youngest age group to 15% in the oldest. Across sectors, predicted probabilities are relatively stable, with slightly higher values observed in S3 and S4 for men, and in S2 for women.

[1] Confidence intervals are centred on the predictions and have lengths equal to 2 × 1.39 × standard errors. This approach follows Goldstein and Healy (1995) and is used to maintain an average level of 5% for Type I errors in pairwise comparisons between groups of means.

Table 1 – *Baseline Characteristics (%) and Total Sample Size - CUORE Project HESs, 2008-2012 and 2018-2019 (Age 35-74 years).*

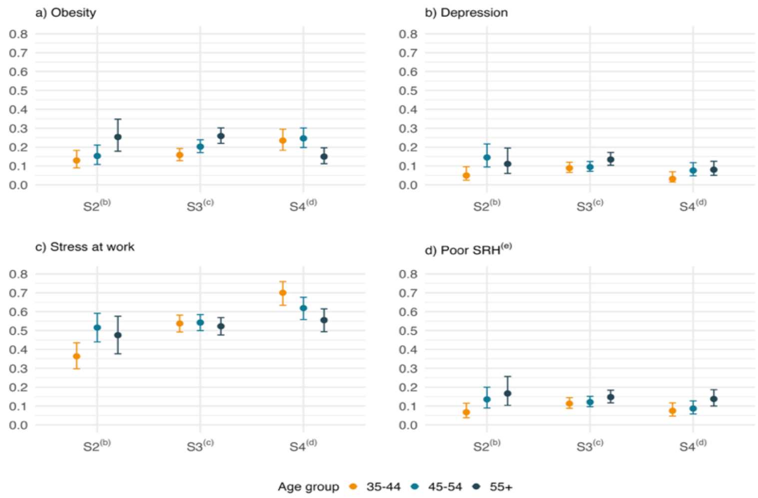
Variables	2008-2012		2018-2019	
	M	W	M	W
Age class				
35-44	34.6	37.5	31.3	31.9
45-54	39.5	40.9	31.9	34.2
55+	25.9	21.6	36.8	34.0
Occupational sector				
S1 [(a)]	2.1	3.3	0.9	1.2
S2 [(b)]	17.5	6.3	10.6	3.4
S3 [(c)]	53.3	77.4	65.2	81.5
S4 [(d)]	27.2	13.0	23.0	13.9
Macro-area of residence				
North	55.1	55.2	41.2	42.1
Centre	30.4	33.1	21.1	19.5
South and Islands	14.5	11.7	37.7	38.4
Educational level				
Lower	34.7	26.3	21.5	18.5
Higher	64.7	73.6	78.5	81.5
Marital status				
Married or cohabiting	80.3	68.8	78.6	66.0
Other	19.7	31.2	21.4	34.0
Health outcomes				
Obesity				
Yes	20.8	15.0	18.2	19.4
No	79.2	85.0	81.8	80.5
Depression				
Yes	6.4	15.5	8.4	17.6
No	92.0	83.4	89.8	81.5
Work-related stress				
Yes	62.8	48.5	50.1	44.9
No	37.2	51.5	49.3	54.8
Poor SRH[(e)]				
Yes	10.9	14.3	11.6	17.6
No	89.1	85.7	88.2	82.4
Total observations	1020	811	687	568

Notes: [(a)]Agriculture; [(b)]Industry; [(c)]Services; [(d)] Commerce-finance-private transport; [(e)]Poor Self-Rated Health. The sample includes only currently employed individuals residing in the following Italian regions: Piedmont, Lombardy, Liguria, Emilia-Romagna, Tuscany, Lazio, Abruzzo, Basilicata, Calabria, and Sicily.
Source: Authors' elaboration on 2008-2012 and 2018-2019 Health Examination Surveys – CUORE Project.

For depression and poor SRH, both outcomes show a similar age gradient, increasing steadily with age for both men and women. Workers aged 55 and older have the highest predicted probabilities. This pattern holds across sectors, although levels are consistently higher among women. Indeed, among the latter, those employed in S2 show the highest predicted probabilities for both depression (from

21% in the youngest age group to 27% in the oldest) and poor SRH (18%-23%), compared to those in S3 and S4, where values remain around or below 20%. Among men, sectoral differences are less marked. However, men in S2 show slightly higher likelihood of poor SRH in the oldest age group (17%) compared to those in S3 (15%) and S4 (14%). Similarly, the highest predicted probabilities for depression among men are observed in S2 (14% for the 45-54 group), though the differences across sectors are more modest than those observed among women.

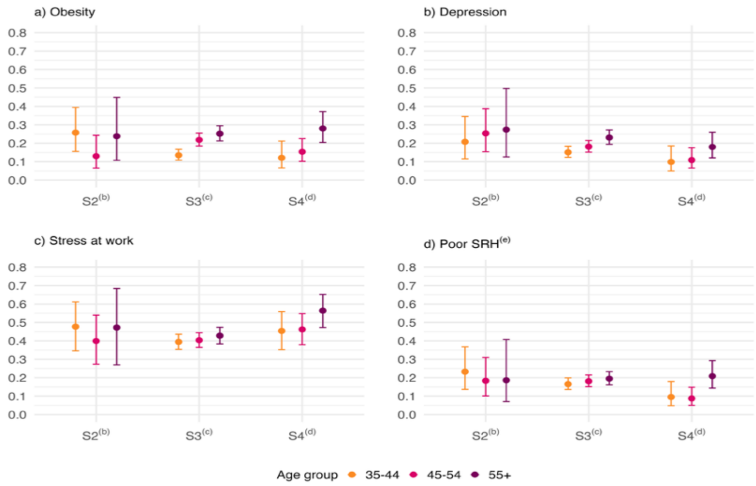
Figure 1 – *Adjusted Predicted Probabilities of Health Outcomes by Sector and Age Group – Men - CUORE Project HESs, 2008-2012 and 2018-2019 (Age 35-74 years).*



Notes: [b]Industry; [c]Services; [d] Commerce-finance-private transport; [e]Poor self-rated health.
Results from logistic regression models adjusted for survey wave, macro-area of residence, educational level, and marital status. In addition, models include an interaction term between occupational sector and age group (35-44, 45-54, 55+). Predicted probabilities are plotted with 83.5% CI.
Source: Authors' elaboration on 2008-2012 and 2018-2019 Health Examination Surveys – CUORE Project.

Work-related stress shows a distinct pattern compared to the other outcomes. Overall, stress is most prevalent in tertiary sectors, particularly S4. In S2, younger women (35–44) report higher levels of stress than their male counterparts (48% vs. 36%). Conversely, in the tertiary sectors, men report higher stress levels, especially in S4, where the predicted probability reaches 70% among the youngest men, compared to 45% among women in the same age group. In S3, stress is relatively stable across age groups for both sexes, though slightly higher for men, remaining around 54% for younger and middle-aged men, and slightly decreasing to 52% for those aged 55+ years. Among women, stress levels in S3 increase marginally with age, from 39% to 43%. S4 exhibits the sharpest gender and age gradients: among men, stress declines notably with age, from 70% (35–44 years) to 62% (45–54) and 56% (55+), whereas among women, it increases with age, rising from 45% in the youngest group to 56% in the oldest.

Figure 2 – *Adjusted Predicted Probabilities of Health Outcomes by Sector and Age Group – Women - CUORE Project HESs, 2008-2012 and 2018-2019 (Age 35-74 years).*



Notes: [(b)]Industry; [(c)]Services; [(d)] Commerce-finance-private transport; [(e)]Poor self-rated health.
Results from logistic regression models adjusted for survey wave, macro-area of residence, educational level, and marital status. In addition, models include an interaction term between occupational sector and age group (35-44, 45-54, 55+). Predicted probabilities are plotted with 83.5% CI.
Source: Authors' elaboration on 2008-2012 and 2018-2019 Health Examination Surveys – CUORE Project.

4. Discussion and conclusions

This study explored how physical and mental health outcomes vary by occupational sector and age group among Italian workers, using data from two nationally representative Italian Health Examination Surveys – CUORE Project. The findings suggest sector-specific health patterns across age groups, with differences between men and women.

Among men, the predicted probability of obesity increased modestly with age in the industrial (S2) and service sectors (S3), pointing to a possible cumulative risk among older workers in these contexts. In contrast, in the commerce, finance, and private transport sector (S4), younger men showed higher probabilities of obesity, which may reflect greater health awareness or selective retention of healthier workers at older ages. Among women, obesity risk increased more consistently with age in S3 and S4, suggesting a gradual accumulation of adverse health conditions across the working life course in these sectors.

Depression and poor SRH exhibited more pronounced age and sectoral gradients, especially among women. Women workers in sector S2 showed the highest predicted probabilities across all age groups. This finding highlights potential vulnerabilities among women in industrial settings, possibly due to cumulative exposure to strenuous or unsupportive work environments over time.

Work-related stress displayed a distinct pattern, with striking age and gender differences. Among men, work-related stress tended to increase with age in S2 and remained stable in S3. In S4, however, a clear inverse relationship with age emerged: younger men reported higher predicted work-related stress than older men. One possible explanation is that, in this sector, older workers may benefit from greater opportunities for career advancement or hierarchical mobility, which could help reducing stress over time. This result is in line with previous studies that found higher levels of stress and burnout among younger workers compared to older ones (Rožman *et al*., 2019). These findings align with our results for men in sector S4, where stress declines with age. However, this trend did not apply to women, for whom stress increased with age in the same sector. This suggests that age alone does not explain differences in stress exposure; sectoral conditions and broader gender-related dynamics in the workplace may also play a role. In line with this, Marinaccio *et al*. (2013) found that women tend to report more negative perceptions of work-related stress risk factors than men, particularly in full-time employment. This may be explained by differences in gender roles at work and at home. Women often take on more caregiving and family responsibilities than men, which can make it harder for them to balance work and personal life.

This study has some limitations that are mainly data-driven. First, it relies on cross-sectional data, which prevents causal interpretation, limits the measurement of health variations and the comparability of cohorts over time. Second, health conditions and occupational characteristics are measured at the same point in time, raising concerns about reverse causality—for instance, it is not possible to determine whether poor working conditions lead to worse health or whether individuals in poorer health are more likely to end up in certain types of jobs. Therefore, our findings should be interpreted as descriptive associations rather than evidence of causal relationships. Third, small sample sizes in certain sectors, particularly agriculture and women in industry, limited the analysis. Agricultural workers were excluded from regression models and estimates for industrial workers (especially women) should be interpreted with caution. Fourth, only obesity was measured objectively; the other health outcomes (depression, stress, and self-rated health) were self-reported and may be influenced by reporting bias or individual perception.

Despite these limitations, the study offers valuable insights. It draws on two nationally standardized health examination data, collected using harmonized procedures consistent with European protocols. The use of both objective and subjective health measures, combined with stratified analysis by age, sex, and occupational sector, provides a nuanced understanding of health inequalities among older workers. These findings are particularly relevant in the context of an ageing workforce and underline the need for workplace health policies that consider gendered and sector-specific risk profiles.

References

ABELIANSKY A.L., BEULMANN M., PRETTNER K. 2024. Are they coming for us? Industrial robots and the mental health of workers, *Research Policy*, Vol. 53, No. 3, 104956.

AKINWALE B., LYNCH K., WIGGINS R., HARDING S., BARTLEY M., BLANE D. 2011. Work, permanent sickness and mortality risk: a prospective cohort study of England and Wales, 1971–2006. *Journal of Epidemiology and Community Health*, Vol. 65, No. 9, pp. 786–792.

BAUMANN I., FROIDEVAUX A., CABIB I. 2022. Health among workers retiring after the state pension age: a longitudinal and comparative study. *BMC Geriatrics*, Vol. 22, No. 1, 984.

BAXTER S., BLANK L., CANTRELL A., GOYDER E. 2021. Is working in later life good for your health? A systematic review of health outcomes resulting from extended working lives. *BMC Public Health*, Vol. 21, 1356.

CAPACCI G., RINESI F. 2014. L'invecchiamento demografico in Italia e nell'Europa del futuro. *Annali del Dipartimento di Metodi e Modelli per l'Economia, il Territorio e la Finanza*, No. 2014, pp. 75–94.

DĖDELĖ A., MIŠKINYTĖ A., ANDRUŠAITYTĖ S., BARTKUTĖ Ž. 2019. Perceived stress among different occupational groups and the interaction with sedentary behaviour. *International Journal of Environmental Research and Public Health*, Vol. 16, No. 23, 4595.

DE ROSE A., RACIOPPI F., CHECCUCCI P., AREZZO M.F., POLLI C. 2019. The workforce aging and challenges for policy and for business. The case of Italy. *Review of European Studies*, Vol. 11, No. 4, pp. 60–69.

DI GESSA G., CORNA L., PRICE D., GLASER K. 2020. Lifetime employment histories and their relationship with 10-year health trajectories in later life, *European Journal of Public Health*, Vol. 30, No. 4, pp. 793–799.

DIXON S. 2003. Implications of population ageing for the labour market. *Labour Market Trends*, Vol. 111, No. 2, pp. 67–76.

DONFRANCESCO C., LO NOCE C., DI LONARDO A., VANNUCCHI S., PALMIERI L. 2021. Progetto CUORE: health examination survey e studi longitudinali a supporto della prevenzione cardiovascolare. *Bollettino Epidemiologico Nazionale*, Vol. 2, No. 4, pp. 12–21.

DONFRANCESCO C., PROFUMO E., LO NOCE C., MINUTOLI D., DI LONARDO A., BUTTARI B., VESPASIANO F., VANNUCCHI S., GALLETTI F., ONDER G., COLIVICCHI F., GALEONE D., BELLISARIO P., PALMIERI L. 2022. Trends of overweight, obesity and anthropometric measurements among the adult population in Italy: The CUORE Project health examination surveys 1998, 2008, and 2018, *PLoS One*, Vol. 17, No. 3, e0264778.

EUROFOUND. 2024. *The Changing Structure of Employment in the EU: Annual Review 2023*. Eurofound research paper, Luxembourg: Publications Office of the European Union.

EUROPEAN HEALTH EXAMINATION SURVEY. (2025, June). *National HES status.* European Health Examination Survey. https://www.ehes.info/national/national_hes_status.htm

GOLDSTEIN H., HEALY M.J. 1995. The graphical presentation of a collection of means. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 158, No. 1, pp. 175-177.

ISTAT, MINISTERO DEL LAVORO E DELLE POLITICHE SOCIALI, INPS, INAIL, ANPAL. 2019. *Il mercato del lavoro 2018: verso una lettura integrata*. Rome: ISTAT.

ISTITUTO SUPERIORE DI SANITÀ. (2025, June). *Progetto CUORE. Indagini di popolazione*. Istituto Superiore di Sanità. https://www.cuore.iss.it/indagini/cuore-data-coorti

JONES M.K., LATREILLE P.L., SLOANE P.J., STANEVA A.V. 2013. Work-related health risks in Europe: Are older workers more vulnerable? *Social Science & Medicine*, Vol. 88, pp. 18–29.

LUCKHAUPT S.E., COHEN M.A., LI J., CALVERT G.M. 2014. Prevalence of obesity among US workers and associations with occupational factors, *American Journal of Preventive Medicine*, Vol. 46, No. 3, pp. 237–248.

LUNAU T., WAHRENDORF M., DRAGANO N., SIEGRIST J. 2013. Work stress and depressive symptoms in older employees: impact of national labour and social policies. *BMC Public Health*, Vol. 13, 1086.

MARINACCIO A., FERRANTE P., CORFIATI M., DI TECCO C., RONDINONE B.M., BONAFEDE M., RONCHETTI M., PERSECHINO B., IAVICOLI S. 2013. The relevance of socio-demographic and occupational variables for the assessment of work-related stress risk, *BMC Public Health*, Vol. 13, 1157.

MINISTRY OF LABOUR AND SOCIAL POLICIES. (2025, June). *Evolution of the pension system*. Ministry of Labour and Social Policies.

https://www.lavoro.gov.it/temi-e-priorita/previdenza/focus-on/previdenza-obbligatoria/pagine/evoluzione-del-sistema-previdenziale

PATEL P.C., DEVARAJ S., HICKS M.J., WORNELL E.J. 2018. County-level job automation risk and health: Evidence from the United States, *Social Science & Medicine*, Vol. 202, pp. 54-60.

ROŽMAN M., GRINKEVICH A.M., TOMINC P. 2019. Occupational stress, symptoms of burnout in the workplace and work satisfaction of the age-diverse employees, *Organizacija*, Vol. 52, No. 1, pp.46-52.

RUGULIES R., SØRENSEN K., DI TECCO C., *et al.* 2021.The effect of exposure to long working hours on depression: a systematic review and meta-analysis from the WHO/ILO joint estimates of the work-related burden of disease and injury. *Environment International*, Vol. 155, 106629.

SERRANO-ALARCÓN M., ARDITO C., LEOMBRUNI R., KENTIKELENIS A., D'ERRICO A., ODONE A., COSTA G., STUCKLER D., on behalf of the IWGRH. 2023. Health and labor market effects of an unanticipated rise in retirement age. Evidence from the 2012 Italian pension reform. *Health Economics*, Vol. 32, No. 12, pp. 2745–2767.

SEWDAS R., DE WIND A., STENHOLM S., COENEN P., LOUWERSE I., BOOT C., VAN DER BEEK A. 2020. Association between retirement and mortality: working longer, living longer? A systematic review and meta-analysis. *Journal of Epidemiology and Community Health*, Vol. 74, No. 5, pp. 473–480.

Benedetta MARCOZZI, Italian National Institute of Health/ University of Rome, Sapienza, benedetta.marcozzi@uniroma1.it

Eleonora TRAPPOLINI, University of Rome, Sapienza, eleonora.trappolini@uniroma1.it

Cinzia LO NOCE, Italian National Institute of Health, cinzia.lonoce@iss.it

Serena VANNUCCHI, Italian National Institute of Health, serena.vannucchi@iss.it

Luigi PALMIERI, Italian National Institute of Health, Rome, luigi.palmieri@iss.it

Chiara DONFRANCESCO, Italian National Institute of Health, chiara.donfrancesco@iss.it

STRENGTHENING PARENTING SKILLS THROUGH COMMUNITY-BASED INTERVENTIONS: EVIDENCE "A VILLAGE FOR GROWING UP" PROJECT

Daniela Del Boca, Luca Mo Costabella, Chiara Pronzato

Abstract. In this paper, we evaluate the impact of participation in *A Village for Growing Up* program. These *Villages* are physical spaces where families with preschool-aged children can engage in shared activities and receive parenting support. We evaluate the program's effects using a difference-in-differences approach combined with propensity score matching. The treatment group consists of families who visited the *Village* at least ten times, while the control group was constructed ad hoc with the support of a survey agency. Our findings show that parents involved in the program read and sing more frequently with their children, and report greater self-efficacy and awareness in their parenting practices.

1. Introduction

The project *A Village for Growing Up* aims to strengthen parenting skills and enhance the quality of educational services for young children in disadvantaged communities, with the ultimate goal of reducing educational poverty.

Until recently, child poverty has been understood primarily in economic terms, measured mainly by parental income and wealth. Following Bronfenbrenner's cognitive development theory (1979), however, the concept has been reframed as multidimensional. Educational poverty is now defined as *"the deprivation experienced by children and adolescents of the opportunity to learn, explore, develop, and freely flourish in their capacities, talents, and aspirations."* This perspective recognizes that, beyond material deprivation, disadvantaged children often lack educational, physical, and socio-emotional growth opportunities. A recent report by Save the Children (2022) highlights the extent of this phenomenon in Italy: 67.6% of children under 17 have never attended a theatre performance, 62.8% have never visited an archaeological site, and 49.9% have never been to a museum. Moreover, 22% do not engage in sports or physical activity, and only 13.5% of children under the age of three attend an early childhood education centre.

In this paper, we evaluate the impact of participation in *A Village for Growing Up* on parenting practices, with the aim of mitigating educational poverty among children. The *Villages* themselves are physical spaces where families with preschool-aged children can participate in shared activities and receive parenting support.

The evaluation focuses on two key dimensions: (1) the frequency of supportive parenting practices and (2) parental conditions. Parenting practices were assessed through a set of activities considered particularly important for stimulating child development and strengthening the parent–child relationship during the early years of life. Parental conditions are measured in terms of self-efficacy and perceived stress, using validated instruments. Self-efficacy is understood as an individual's belief in their ability to manage specific tasks, situations, or aspects of their psychological and social functioning. Findings indicate that participating parents are more frequently involved in joint activities with their children—such as reading and singing—and report higher levels of self-efficacy and awareness in their parenting.

2. Literature review

The theoretical framework *A Village for Growing Up* is grounded in Bronfenbrenner's cognitive development theory (1979), which emphasizes the educational value of family-based services for children—conceptualized as a *Village* surrounding families. This "ecological environment" approach envisions an educational community in which multiple actors contribute to supporting families with young children.

It is well established that parents play a critical role in the development of their children's human capital (Del Boca, 2015; Cunha & Heckman, 2007), and that early childhood educational experiences have long-term impacts on both cognitive and non-cognitive outcomes in education, the labour market, and health. According to Heckman (2008), investing in a child's education also brings significant economic benefits to society, including savings in welfare expenditures. Educational investments in children's human capital and well-being have very important impacts on children from disadvantaged families, where parental inputs are typically more limited due to fewer economic and cultural resources (Brilli *et al*., 2016; Del Boca *et al.*, 2022; Carneiro, 2019). Recent parenting policy experiments have shown positive effects on parenting behaviours and child development (Kiernan & Mensah, 2011). Two recent studies have examined the effects of an Italian program—somewhat analogous to *A Village for Growing Up* —designed for parents and their children. The FACE program (*Becoming an educational community*) ran from April 2018 to July 2021, involving twenty national and local partners across four Italian provinces: Naples, Palermo, Reggio Emilia, and Teramo. Families participated in a range of activities including the exploration of digital, mathematical, and expressive languages (such as reading and singing), infant massage, environmental languages (natural and scientific dimensions), and the languages of the body and food. Led by expert educators, these activities aimed to stimulate and improve parents' manual, expressive, sensory, communicative, and relational competencies. To estimate the program's impact, Del Boca, Pronzato, and Schiavon (2021, 2024) compared "treated" families (those enrolled early in the program) with control families (those enrolled at a later stage). The evaluation reveals positive effects on perceived importance of living in an area rich in opportunities, maintaining quality relationships with family and friends, feeling confident about sharing experiences with other parents, and recognizing the usefulness of digital tools (such as tablets and phones, using appropriate educational apps) for learning. The program also helps parents to achieve greater confidence in interacting and engaging with other parents and adults.

3. The program and its evaluation design

A Village for Growing Up focuses on organizing activities that can be easily replicated within the home environment and thereby enhance the quality of parenting practices. Families are reached and engaged through various strategies, including home visits, social media, and access to existing service networks. The *Villages* operate between three and five days per week, for a total of 10–12 hours. Trained educators lead the activities and regularly exchange experiences to adapt and improve programming. Local stakeholders are also mobilized to disseminate information and contribute to ongoing activities, fostering the creation of new networks between families and services that promote shared values, social inclusion, and sustainability to counteract educational poverty.

The first edition of A *Village for Growing Up* was established in 2018 in disadvantaged communities across nine Italian cities, co-funded by the social enterprise *Con i Bambini* and *The Human Safety Net (Gruppo Generali)*, within the framework of the Italian Fund for Combating Child Educational Poverty. In 2021, with additional funding from the same Fund, the project was expanded. Today, many *Villages* are active throughout Italy: Assisi, Bagaladi, Caprarica di Lecce, Castellammare, Cefalù, Castelbuono, Gualdo Tadino, Lecce, Macerata, Milan, Palermo, Palmanova, Rome, San Benedetto del Tronto, and Trieste.

The aim of the impact evaluation is to determine whether, and to what extent, participation in the project leads to improvements in beneficiaries' conditions, specifically in parenting practices, parental self-efficacy, and perceived stress. It is first necessary to clarify what is meant by "treatment." In this case, in agreement with the Center for Child Health, the evaluation focuses on *Village* attendance. The underlying hypothesis is that any potential benefits derived from participation will emerge after a certain period, during which the experiences and learnings from the *Villages* can progressively influence parenting practices and perceptions. The treatment is defined as having attended a *Village* at least 10 times. The evaluation adopts a difference-in-differences strategy (see Angrist & Pischke, 2008): it assumes that, in the absence of the treatment, the pre–post change in outcomes would have been similar for both groups.

The treatment group consisted of parents of children under the age of six who decide to attend a *Village*. The control group was constructed with the support of a survey research institute, tasked with identifying a group of parents who had not attended the *Villages* but were otherwise similar to the treated group. The sample was designed to match the treated group on key characteristics such as gender and age of the child, geographic location, parental employment and education levels.

Parents in the treatment group were interviewed by the third visit to the *Village* (pre) and after the tenth visit (post). The time between the two questionnaires varies considerably from case to case, depending on how fast each family reached their tenth *Village* visit. The range spans from less than one month to over eight months. 25% of families completed the tenth visit within a month and a half, half within three months, and three-quarters within five months. The survey agency interviewed the control families at a time interval that reflects the distribution observed in the treatment group.

In the difference-in-differences estimation framework, it is assumed that—absent participation in the *Villages*—any initial differences between the two groups would have remained constant over time. This assumption becomes more defensible the more similar the two groups are. For this reason, the estimates were strengthened by a preliminary step aimed at increasing group similarity through matching techniques (see Rosenbaum & Rubin, 1983): each treated family was matched with one or more control families with similar characteristics. The initial questionnaire gathered a broad set of information: age and sex of the responding parent; place of residence; child's age and school attendance (nursery or preschool); family composition; parents' employment and education; availability of third-party childcare support; and the presence of additional caregiving responsibilities not related to children. We use kernel matching, which assigns to each treated unit a weighted average of control units, with weights increasing with the similarity to the treated subject.

4. The outcomes of interest

With regard to the outcomes of interest, the evaluation focuses on two main dimensions: the frequency of certain parenting practices, and specific parental conditions. All outcomes were measured through a questionnaire completed by one parent.

4.1. Parenting practises

As for parenting practices, the focus was on a set of activities considered particularly important for stimulating children and fostering the parent–child relationship, especially in the early years of life. The adoption of these practices was assessed through a series of questions about how often they are performed, such as: "*Which of the following activities do you do with your child*?" The listed activities included: playing together, naming and explaining things that are seen or done with children, reading stories together, singing, and listening to music. Each item allowed for three response options: *never or almost never*, *occasionally*, and *every day or almost every day*. An additional question, potentially correlated with reading habits, concerned the number of children's books available in the home for shared reading with the child.

4.2. Parents' self-efficacy

Concerning parental conditions, the evaluation focused on self-efficacy and perceived stress, both assessed using validated instruments. Self-efficacy is understood as the individual's belief in their capability to manage specific tasks, situations, or aspects of their psychological or social functioning, is measured using the Perceived Maternal Parenting Self-Efficacy (PMP-SE) scale (see Barnes & Adamson-Macedo, 2007). The PMP-SE, developed primarily for mothers of premature infants, comprises 20 items asking mothers to self-evaluate their abilities across various domains—such as feeding the child, interpreting signals, soothing, or engaging the child. The test allows for the construction of four subscale scores corresponding to distinct domains: care taking (e.g., feeding, changing, bathing the child); evoking behavior (e.g., calming the child, making them happy, drawing their attention); reading behavior (e.g., recognizing signs of tiredness, discomfort, preferences); situational beliefs (e.g., quality of interaction, affection expression).

4.3. Parents' perceived stress

Parental stress was assessed through the Parenting Stress Index (PSI; see Abidin, 1997), which evaluates the level of stress in the parent–child relationship. While the original version includes 120 items, for the purposes of this study, a specific subset of 12 items from the short form was used, focusing exclusively on the domain of parental distress, which captures the level of stress perceived by parents due to personal factors directly related to their parenting role.

5. The empirical analysis

5.1. Descriptive statistics

The sample includes 208 treated and 173 control participants. It is important to note that respondents in the treated group were almost exclusively mothers (95%), compared to 86% in the control group. To ensure sample consistency, analyses were restricted to the female subpopulation, which comprises 198 treated and 149 control mothers.

Table 1 – *Mother's characteristics.*

	Treated	Controls	Diff.
Age:			
<=30	0.126	0.188	-0.062
31-35	0.394	0.228	0.166**
36-40	0.283	0.376	-0.093*
>= 41	0.197	0.208	-0.011
Area:			
North	0.576	0.658	-0.082
Centre	0.263	0.289	-0.026
South, islands	0.162	0.054	0.108**
Nationality:			
Italian	0.838	0.960	-0.121**
Education:			
Compulsory	0.086	0.074	0.012
Secondary	0.369	0.376	-0.007
Tertiary	0.545	0.550	-0.005
Employed	0.606	0.678	-0.072
Observations	198	149	

*** significant at 5% level, * at 10%*

Tables 1 and 2 summarize the characteristics of the treated and control groups, referring respectively to mothers and households. Mothers in the treatment group tend to be somewhat younger than those in the control group. While nearly four in ten treated mothers are between 31 and 35 years old (39%), this is the case for less than a quarter of mothers in the control group (23%). By contrast, the largest share of control-group mothers falls in the 36–40 age bracket (38%), compared to 28% among treated mothers. Differences also emerge in terms of geographical location: treated mothers are more likely to reside in the South and Islands (16% versus 5% in the control group). Moreover, foreign families are more represented among participants, accounting for 16% of the treatment group compared to only 4% of the control group. No major differences, however, are observed in mothers' educational attainment. In both groups, more than half hold a university degree and over one-third a secondary school diploma (Table 1).

Table 2 – *Household's characteristics.*

	Treated	Controls	Diff.
Children			
1	0.788	0.685	0.103**
2	0.197	0.262	0.065
3+	0.015	0.054	-0.039
Workers:			
0	0.040	0.027	0.014
1	0.359	0.349	0.010
2+	0.601	0.624	-0.023
Child's age:			
<= 6 months	0.207	0.248	-0.041
7-12 months	0.293	0.255	0.038
13-24 months	0.313	0.255	0.058
>= 25 months	0.187	0.242	-0.055
Childcare:			
Early childcare	0.197	0.195	0.002
Pre-school	0.106	0.148	-0.042
Support	0.515	0.617	-0.102*
Other caring	0.015	0.034	-0.018
Make ends meet			
Lots of diff.	0.045	0.154	-0.109**
Some diff.	0.273	0.497	-0.224**
Quite easily	0.455	0.282	0.173**
Very easily	0.227	0.067	0.160**
Observations	198	149	

*** significant at 5% level, * at 10%*

Mothers attending the *Villages* are more likely to have only one child (79% compared to 69% in the control group) and are less likely to receive help from relatives or friends in childcare (52% versus 62% in the control group). The most striking difference, however, concerns economic resources. Families in the control group report far greater financial strain: 65% of them declare having “many” or “some” economic difficulties, compared with only 32% of families in the treatment group. Other characteristics, by contrast, appear well balanced across the two groups: about half of the children are in their first year of life, around 20% attend nursery care, and an average of 12% are enrolled in preschool (Table 2).

Table 3 – *Starting levels.*

	Treated	Diff. treated-controls
Self-efficacy:		
Care taking	0.835	-0.031**
Evoking	0.796	-0.001
Reading	0.811	-0.025**
Situational	0.862	-0.016
Stress	0.459	-0.018
Playing:		
Never	0.015	-0.005
Occasionally	0.035	-0.032
Every day	0.949	0.037
Explain:		
Never	0.030	-0.023
Occasionally	0.056	-0.065**
Every day	0.914	0.089**
Reading:		
Never	0.066	-0.142**
Occasionally	0.313	0.072
Every day	0.621	0.071
Singing:		
Never	0.040	-0.134**
Occasionally	0.131	-0.359**
Every day	0.828	0.493**
Music:		
Never	0.020	-0.101**
Occasionally	0.182	-0.254**
Every day	0.798	0.355**
Books:		
0	0.015	-0.106**
1	0.106	-0.075*
2-5	0.172	-0.043
6+	0.707	0.224**

*** significant at 5% level, * at 10%*

An additional description of baseline conditions is presented in Table 3, which summarizes the initial mean levels of the outcomes, as captured by the first questionnaire. For consistency, self-efficacy and stress variables are scaled to a 0–1 range. Two key baseline differences emerge. First, mothers in the treatment group report lower levels of self-efficacy, particularly in the dimensions of *care taking* and *reading*. At the same time, however, they are more frequently engaged in child-stimulating activities: they are more likely to explain concepts, read to their children, listen to music together, and sing to or with them.

Table 4 – *Impact of the program.*

	Starting level	Estimated effect
Self-efficacy:		
Care taking	0.835	0.072**
Evoking	0.796	0.064*
Reading	0.811	0.065*
Situational	0.862	0.062*
Stress	0.459	0.013
Playing:		
Never	0.015	-0.011
Occasionally	0.035	0.005
Every day	0.949	0.006
Explain:		
Never	0.030	-0.019
Occasionally	0.056	0.046
Every day	0.914	-0.027
Reading:		
Never	0.066	-0.018
Occasionally	0.313	-0.109**
Every day	0.621	0.127**
Singing:		
Never	0.040	-0.084**
Occasionally	0.131	0.122**
Every day	0.828	-0.038
Music:		
Never	0.020	-0.025
Occasionally	0.182	-0.045
Every day	0.798	0.071
Books:		
0	0.015	0.000
1	0.106	-0.004
2-5	0.172	-0.043
6+	0.707	0.047

*** significant at 5% level, * at 10%*

5.2. The impact of the program

Table 4 presents the impact estimates. When applying a difference-in-differences approach, we compare how much the treatment group improves (or worsens) between the pre- and post-intervention periods relative to the control group. For instance, the *care taking* dimension of self-efficacy is 0.835 in the treatment group before the intervention (a slightly lower baseline value than in the control group, see Table 3), but increases by an additional 0.072—relative to the control group—between the pre- and post-treatment observations. We also observe improvements in the other three dimensions of self-efficacy, each amounting to somewhat less than

10% in relative terms: the ability to evoke (e.g., calm the child) increases from 0.796 to 0.860; the ability to read (e.g., recognize signs of tiredness) from 0.811 to 0.876; and the ability to understand the situation (e.g., affection expression) from 0.862 to 0.924. No significant change is observed in parenting stress. Regarding interactive and stimulating parenting practices, improvements are found in singing with the child and in reading with the child: the probability of reading on a daily basis increases from 0.621 to 0.748, while the probability of singing occasionally increases from 0.131 to 0.253.

5.3. Heterogeneous effects

The study also explores potential heterogeneity within the target population.[1] We consider various forms of heterogeneity by analysing subgroup effects based on the mother's educational level, employment status, the age of the youngest child, and the time elapsed between the baseline and end line surveys. The distinction between employed and non-employed mothers does not reveal substantial differences in benefits; in contrast, the increase in parental self-confidence is especially pronounced among more educated mothers. With respect to the child’s age, we find that the program’s benefits are greater for parents of children older than one year. A final distinction concerns the time elapsed between the completion of the initial and of the final questionnaire. We divide respondents into those who completed them within three months and those who took more than three months. This distinction reflects a different “intensity” of treatment—not dictated by the project design, but by how mothers engaged with the project. As such, one would expect the estimated effects to be stronger for the first group, as indeed confirmed by the analysis. In conclusion, the most pronounced results are observed when attendance at the *Villages* is more intensive.

6. Conclusions

The results obtained offer several important insights. Overall, we find positive effects on parenting practices and on parents’ sense of self-efficacy, but no impact on stress levels.

Before turning to these outcomes, however, it is worth reflecting on the families who attend the *Villages*. More than half of the participating parents hold a university degree. By comparison, among Italian women of childbearing age, the share of graduates is about 35% (IT-SILC, 2023). This indicates that the project does not fully reach its intended target, which should include not only middle-income households

[1] Results availbale upon request.

but also families facing greater economic and cultural vulnerability. The literature consistently shows that these families are those who would benefit the most from such initiatives. We also observe that the families attending the *Villages* are generally already more engaged in stimulating activities with their children (e.g., singing, listening to music). At the same time, they report feeling less effective in many dimensions of parenting, which may help explain their motivation to join the program.

Another interesting aspect concerns the age of the children: half are under one year old, while the other half are between ages 1 and 5. One possible explanation is that mothers, supported by maternity or parental leave, have more time to dedicate to their infants during the first year of life. The results, however, suggest that the program generates greater benefits for families with relatively older children. This may indicate the value of scheduling activities closer to dinnertime to better accommodate families with children in this age group.

Finally, the lack of measurable impact on stress deserves consideration. While one of the program's aims is to reduce parental stress, in the short term this potential benefit may be offset by the additional commitment required to attend *Village* activities. Longer-term evaluation would be needed to better understand these dynamics.

Acknowledgements

We are grateful to the social enterprise *Con i Bambini* and to *The Human Safety Net* (*Gruppo Generali*) for funding this research. We are also thankful to Anduena Alushaj and Giorgio Tamberlini (*CSB onlus – Centro per la Salute delle Bambine e dei Bambini*) for their valuable input, and for their support throughout the research.

References

ABIDIN R.R. 1997. Parenting Stress Index: A measure of the parent–child system. In ZALAQUETT C.P., WOOD R.J. (Eds.) *Evaluating stress: A book of resources*. Lanham: Scarecrow Education.

ANGRIST J.D., PISCHKE J.S. 2008. *Mostly harmless econometrics*. Princeton: Princeton University Press.

BARNES C.R., ADAMSON-MACEDO E.N. 2007. Perceived Maternal Parenting Self-Efficacy (PMP S-E) tool: development and validation with mothers of hospitalized preterm neonates, *Journal of Advanced Nursing*, Vol. 60, No. 5, pp. 550–560.

BRILLI Y., DEL BOCA D., PRONZATO C. 2016. Does the availability of child care impact child outcomes?, *Review of the Economics of the Household*, Vol. 14, No. 1, pp. 27–51.

BRONFENBRENNER U. 1979. *The ecology of human development. Experiments by nature and design*. Cambridge: Harvard University Press.

CARNEIRO P. 2019. *Parental beliefs, investments, and child development: Evidence from a large-scale experiment*. IZA Discussion Paper No. 12506.

CUNHA F., HECKMAN J.J. 2007. The technology of skill formation, *American Economic Review*, Vol. 97, No. 2, pp. 31–47.

DEL BOCA D. 2015. Child care choices and child outcomes. *IZA World of Labor*, No. 134.

DEL BOCA D., MARTINO E., PRONZATO C. 2022. Non-cognitive outcomes and early childcare attendance, *Review of Economics of the Household*, Vol. 20, No. 4, pp. 1059–1085.

DEL BOCA D., PRONZATO C., SCHIAVON L. 2021. How do parenting courses affect families' time use? Evidence from an RCT experiment in Italy, *Rassegna Italiana di Valutazione*, No. 79, pp. 39–59.

DEL BOCA D., PRONZATO C., SCHIAVON L. 2024. The Evaluation of Short Repeated Programs: The Case of Parenting Skills Programs, *Evaluation and Program Planning*, Vol. 106, 102476.

HECKMAN J.J. 2008. Schools, skills and synapses, *Economic Inquiry*, Vol. 46, No. 3, pp. 289–324.

KIERNAN K.E., MENSAH F.K. 2011. Parenting, family resources and children's early educational attainment: the mediating role of parenting, *British Educational Research Journal*, Vol. 37, No. 2, pp. 317–336.

ROSENBAUM P.R., RUBIN D.B. 1983. The central role of the propensity score in observational studies for causal effects, *Biometrika*, Vol. 70, No. 1, pp. 41–55.

SAVE THE CHILDREN. 2022. *La povertà educativa. Necessario un cambiamento di passo nelle politiche di contrasto*. Roma.

Daniela DEL BOCA, Collegio Carlo Alberto, daniela.delboca@carloalberto.org
Luca MO COSTABELLA, ASVAPP, lmocostabella@asvapp.org
Chiara PRONZATO, Università di Torino, chiaradaniela.pronzato@unito.it

EMPLOYMENT, QUALITY OF WORK AND MONEY: A REGIONAL PERSPECTIVE ON GENDER DISPARITIES IN ITALY[1]

Daniela Fusco, Cira Acampora, Paola Giordano, Maria A. Liguori, Margherita M. Pagliuca

Abstract. This study aims to investigate the gender disparities in the Domain Work and Money, considering three sub-domains: Employment, Quality of Work and Money, analysing regional trends in Italy before and after the COVID-19 pandemic. In this study, we investigated the gender disparities in the Domain using both Istat and INPS data sources. For the Domain, a set of 13 simple indicators, both objective and subjective, more closely related to work well-being and divided by sub-domain, was selected. The analysis covered 5 years (2018–2022), spanning both the pre- and post-COVID-19 periods. The analyses were carried out with a model based on the use of multivariate analysis techniques and the construction of synthetic indices, one for each sub-domain, capable of provide a comprehensive overview, that would not emerge from individual indicators alone. The synthetic index was named Gender Equality Regional Index (GERI). The results revealed that the regions showed different behaviour over the years and different reactions to the COVID-19 pandemic, with important spatial patterns.

1. Introduction

The Gender gap, in a work and economic context, is manifested through systemic disparities between men and women in terms of labour-force participation, earnings and limited access to career advancement for women. According to the International Labour Organization (ILO), despite the progress made in recent decades, the gender gap is still persistent worldwide. In 2024, only 46.4 percent of women of working age were employed compared to 69.5 percent of men and, in more than 30 years, the gender gap in employment has narrowed by only 4 percentage points (ILO, 2025). In terms of pay, women continue to earn less than men, with the average gender pay gap in all OECD countries standing at 12 percent, with significant variations between member countries. Women, despite having similar or better qualifications than men, often work in low-paid jobs (OECD, 2022).

[1] Authors contributions: Daniela Fusco paragraphs 3.3 and 4; Cira Acampora paragraph 1; Paola Giordano paragraph 2.1, Abstract and References; Maria A. Liguori paragraphs 3, 3.1 and 3.2; Margherita M. Pagliuca paragraph 2.2.

According to the Gender Equality Index, gender inequalities are particularly marked in the labour and money sectors, areas in which Italy continues to record values below the European average (EIGE, 2024).

As many international studies point out, women show a greater propensity to orient themselves towards work positions compatible with family responsibilities, in particular childcare; on this front, there is a lot of evidence to support the so-called "motherhood penalty" (Ortiz-Ospina *et al.*, 2018). Underlying these gender differences are above all cultural and social factors, which attribute to women the main role in the management of domestic and care responsibilities. Recent studies show, in fact, that women are more inclined to prefer part-time jobs or jobs with flexible hours to better adapt to family needs (Jost and Möser, 2023).

The available data therefore highlight the need for a multidimensional approach in measuring and overcoming the gender gap, which integrates objective and subjective indicators for a more complete assessment of the phenomenon. For this reason, in addition to the employment and economic aspects, the component of the quality of work was also considered, more closely linked to the satisfaction of professional needs and individual well-being.

The study focuses on the regional geographical level, with the aim of understanding how gender inequality presents itself in Italy today, to grasp territorial differences and fill the lack of specific analyses on this scale.

Despite growing attention to gender equality, subnational analyses remain limited in Italy. Amici and Stefani were among the first to adapt the Gender Equality Index to the regional level, highlighting significant territorial disparities, with southern regions notably disadvantaged compared to those in the North and Centre (Amici and Stefani, 2013).

Another relevant contribution is the Regional Gender Equality Index (R-GEI), which takes up the conceptual model of EIGE divided into six domains (work, money, knowledge, time, power, health), applying it to the Italian regions. The results confirm the territorial gaps by highlighting the analytical value of local measurements for the definition of effective policies (Di Bella *et al.*, 2021).

The decision to analyse gender disparities at the regional level is based on both empirical and policy considerations. An analysis on a regional scale makes it possible to highlight the profound socio-economic differences between the different areas of the country, which directly affect gender inequalities, and to guide targeted policy interventions. This approach is in line with the perspective of EIGE, which stresses the importance of localised strategies in addressing structural inequalities.

Starting from the domains proposed by the GEI, integrated with some of the indicators of the Sustainable Development Goals (SDGs) , in order to measure the changes in gender inequalities with the arrival of the pandemic and how they differ in the Italian regions (Fusco *et al.*, 2023), this work intends to identify a measurement of the phenomenon, from 2018 to 2022, in the Work and Money Domain, specifically in its 3 sub-domains: Employment, Quality of Work and Money by filling a gap in the literature on the post-pandemic period and offering new evidence on a crucial phase for the labour market and social resilience.

2. Methodology

2.1. Data sources and selected indicators

Developing a synthetic measure is a challenging process that necessitates making a multitude of crucial decisions across conceptual, analytical, and empirical dimensions.

For each sub-domain, a study of sources was carried out in order to identify the most suitable indicators to represent the phenomenon. All the analysed data come from institutional surveys (Table 1).

Therefore, the elementary indicators were chosen based on their significance and their capacity to represent the phenomenon under analysis. Additional selection criteria covered the availability of statistical data at regional level; data timeliness to ensure an adequate time comparison; thematic appropriateness feasibility (the availability of obtaining and processing updated data in a simple way has been taken into account). The analysis covered 5 years (from 2018 to 2022).

This analysis highlights how overall indicators vary between gender and over time, providing a detailed picture of gender differences in various domains.

Table 1 - *Basic indicators, algorithms and data sources for the Work and Money Domain.*

Sub-domain	Basic indicator	Data source
Employment	Young people neither in employment nor in education and training (NEET)	Istat - Labour force survey
	Employment rate of parents in a couple (15-64 years old)	Istat - Labour force survey
	Employment rate (20-64 years old)	Istat - Labour force survey
Money	People at risk of poverty	Istat - Statistics on Income and Living Conditions (EU-SILC)
	Gross hourly wages per hour paid for job positions	Istat - Annual register of wages, hours and labour costs for individuals and enterprises (RACLI)
	Pension expenditure	INPS - Statistics of social security and social assistance
Quality of work	Cultural employment	Istat - Labour force survey
	Share of employed persons with temporary jobs for at least 5 years	Istat - Labour force survey
	Share of over-qualified employed persons	Istat - Labour force survey
	Involuntary part time	Istat - Labour force survey
	Share of employed persons who feel their work unsecure	Istat - Labour force survey
	Job satisfaction	Istat - Labour force survey

Source: authors' elaboration.

2.2. Method

In order to synthesize each sub-domain in a single composite index, we use a formative measurement model that is the indicators are considered as causing the gender gap (rather than being caused by it, such as in the reflective approach), so, the correlations between basic indicators are not very relevant. Gender equality is based on the existence of a gender role, so the aim of the indicators is to establish the relative situation of men and women and the changes that have occurred at different moments in time. Therefore, in line with previous studies (e.g., Bericat, 2012; Klasen, 2006; Cascella *et al.*, 2022) we didn't use absolute levels of indicators, but we have calculated female-to male ratios (R), because they can be interpreted as a measure of the gender gap. The ratios measure the level attained by women in relation to the status attained by men: in this way it is possible to capture the different forms of inequality rather than the single levels. A value R = 1 indicates perfect parity; a value 0 < R <1 indicates inequality favourable to men; and a value R > 1 indicates inequality favourable to women (Permanyer, 2010). Until now, some examples of indices that have used ratios are the Gender Equality Index developed by European Institute for Gender Equality (EIGE) and the Extended Regional Gender Gaps Index (eRGGI) (Cascella *et al*., 2022).

We apply the Adjusted Mazziotta–Pareto Index (AMPI) because it enables both spatial and temporal comparisons across units. Moreover, since our aim is to capture heterogeneous dimensions of gender gap without assuming full compensability, we adopt a formative measurement approach. In this perspective, indicators are considered to shape the latent construct (gender inequality), are non-interchangeable, and do not need to co-vary. The AMPI is a non-compensatory (or partially compensatory) composite index, based on the assumption of non-substitutability of indicators—i.e., a "deficit" in one dimension cannot be entirely offset by a "surplus" in another (Mazziotta and Pareto, 2016). It is based on a non-linear function which, starting from the arithmetic mean, introduces a penalty for the units with unbalanced values of the indicators.

Individual indicators are normalized by a re-scaling in the range (70;130) according to two 'goalposts', i.e. a minimum and a maximum value which represent the possible range of each variable for all time periods and for all units. Such type of normalization allows to perform absolute comparisons over time. Given the matrix $X = \{x_{ijt}\}$, where x_{ijt} is the value of the indicator j for the unit i (the Italian regions) in time t (2018, …, 2022 years), we calculate the matrix R of normalized scores r_{ijt} as follow:

$$r_{ijt} = \begin{cases} \dfrac{\left(x_{ijt} - Min_{x_j}\right)}{\left(Max_{x_j} - Min_{x_j}\right)} * 60 + 70, & if\ the\ indicator's\ polarity\ is\ positive \\ \dfrac{\left(Max_{x_j} - x_{ijt}\right)}{\left(Max_{x_j} - Min_{x_j}\right)} * 60 + 70, & if\ the\ indicator's\ polarity\ is\ negative \end{cases}$$

where Min_{x_j} and Max_{x_j} are the "goalposts" for the indicator j.

Denoting with M_{r_i}, $S_{r_i}$and cv_{ri} respectively, the mean, standard deviation and coefficient of variation (S_{r_i}/M_{r_i}) of the normalized values (r_{ijt}) of the unit i, for each year the generalized form of AMPI is given by:

$$AMPI_i^{+/-} = M_{r_i} \pm S_{r_i} cv_{ri} \tag{1}$$

The AMPI decomposes the score of each unit in two components: the mean level (M_{r_i}) and the penalty ($S_{r_i}cv_i$). The penalty reflects the variability of the indicators in relation to the mean value ('horizontal variability') and it is used to adjust the score of each unit. This penalty is either added to or subtracted from the mean, depending on the direction of the elementary indicators with respect to the phenomenon under analysis. Specifically, if the composite index is "positive"—that is, increasing values correspond to favorable variations of the phenomenon—then $AMPI^-$ is applied. Conversely, if the composite index is "negative"—that is, increasing values correspond to unfavorable variations of the phenomenon—then $AMPI^+$ is applied.

In our application, the penalty is treated differently depending on whether the mean value exceeds or falls below the threshold of 100, which represents the condition of parity. Specifically, if the mean value is greater than 100, the penalty is added; conversely, if the mean is less than 100, the penalty is subtracted. This approach allows us to take into account not only the variability of the indicators but also the direction of the imbalance with respect to the parity condition. In this context, the penalty is not solely intended to capture internal disparities among the indicators; rather, it serves as a mechanism to amplify or mitigate the expression of advantage or disadvantage relative to the normative threshold of 100, which holds substantive significance in the assessment of gender equality.

Furthermore, to better analyse and represent the phenomenon, a synthetic index, named Gender Equality Regional Index (GERI) was constructed for each sub-domain of the Work and Money Domain. This index facilitates spatial and temporal comparisons across Italian regions from 2018 to 2022, aiming to identify changes influenced by the COVID-19 pandemic.

3. Results

The three selected sub-domains (Employment, Quality of Work and Money) were explained through 13 indicators.

The composite index captures the complexity but it reduces the dimensions in space with an evident loss of information. Therefore, to understand the results of the index, reading the indicators is the best choice. The values of indicators and indices for each sub-domain will be presented in the following paragraphs. The GERI was been represented with radar charts comparing data from 2018 to 2022 across all Italian regions. The chart features five coloured lines representing each year and regions are labelled around the perimeter.

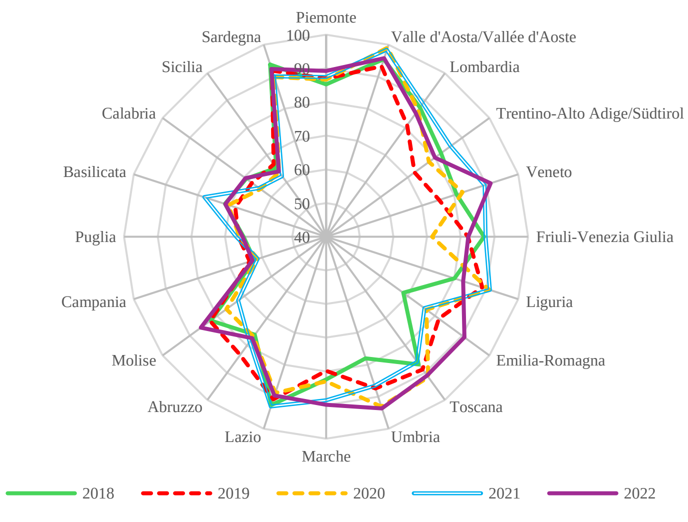
3.1. Sub-domain Employment

On average, throughout the period considered, the share of NEETs made up of women is higher than that of men (22.9 percent and 19.6 percent, respectively), with an increase for both genders during the pandemic peak. Notably, in 2020, the gap widened (3.9 points difference) and then narrowed over time. 2022 saw the lowest share of NEETs (19.4 percent for women and 17.1 percent for men) and the smallest gap (2.3 points).

The employment rate (20-64 years) sees large gender differences, with male values exceeding female values by almost 20 percentage points on average (73.0 percent for men and 54.0 percent for women). The range is constant over the period under consideration, while rates, with the exclusion of the pandemic years, tend to be equally upward for both sexes. If we consider parents in couples, the gap tends to be larger (almost 30 percentage points) and remains constant over time.

The GERI results underscore the marked gender employment gap between the territorial divisions: with the understanding that levels of absolute parity will be recorded only in Valle d'Aosta in 2020-2021, the gap in the southern regions is far greater than in the central and northern regions (Figure 1).

Figure 1 – *Regional synthetic Index of sub-domain Employment. Years 2018-2022.*



Source: authors' elaboration.

At the tail end is still Campania, where throughout the period under consideration index values are always below 64, with a peak during the pandemic of 61.4.

The southern regions with higher index values are Molise and Basilicata, with the latter having the highest index value (78.0) in 2021, unlike the national trend. Values more in line with the Central-North are recorded by Sardegna with a 5-year average of 91.6. Among the northern regions, the lowest values of the index are the preserve of Trentino-Alto Adige, where it goes from 82.2 in 2018 to 79.9 in 2022, with an increase in the index value in 2021 (85.8). Emilia-Romagna is the region with the best increase, rising from 68.3 in 2018 to 90.8 in 2022.

3.2. Sub-domain Quality of Work

On average, throughout the period under consideration (2018-2022), those in scientific-technological professions and with a university education, the so-called cultural employers, are represented in the majority by the female sex (22.6 percent compared to 13.6 percent), with no major changes over the five-year period.

With regard to medium-term workers (employed on fixed-term contracts for at least 5 years), there are no particular differences between the sexes and over the period under consideration (on average 18.1 percent for women and 17.6 percent for men). The same cannot be said for the over-qualified employed: in this case, in fact, women prevail over the so-called stronger sex (on average 27.7 percent and 25 percent, respectively), although the pandemic seems to have had a negative influence (in fact, it goes from 26.4 percent in 2018 to 28.1 percent in 2022).

Regarding involuntary part time, the gender gap is evident: on average, in the 5 years considered, almost one-fifth of females are in this situation, compared to 6.5 percent of males. By 2022, however, the trend seems to be evolving positively, from 19.4 percent in 2018 to 16.5 percent.

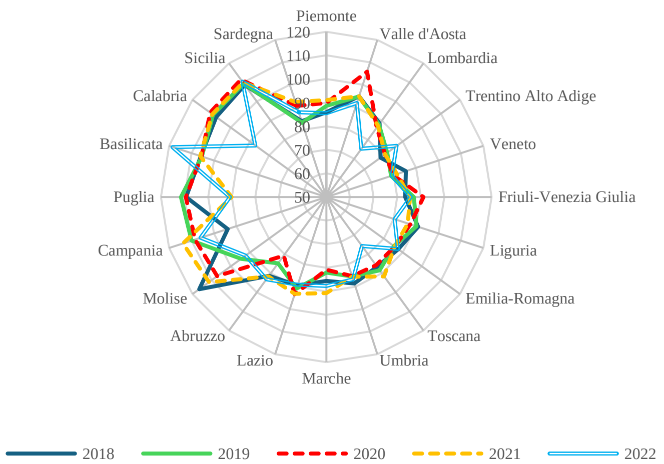
The percentage of people who fear losing their jobs tends to be low for both sexes (on average 6.6 percent of females and 5.8 percent of males), with the highest, albeit slightly significant values recorded in 2020.

Slightly less than half of workers of both sexes consider themselves, on average, to be very satisfied with their jobs with a tendency for satisfaction to improve over the entire five-year period for the following aspects (except for 2022 for females, which record a 0.4 percent decline from 2021): earnings, career opportunities, number of hours worked, job stability, home-to-work distance, and interest in work. The period under consideration shows a steady trend of improving satisfaction with one's job.

The Figure 2 shows the results of the GERI for this sub-domain. In no region does the quality of work reach parity; generally, in the central and northern regions, disparities favour men, while in the south and Sicilia, they favour women. The years 2020 and 2021 are when the average values are closest to parity. One possible explanation for this phenomenon could be the increase in flexible work measures, such as remote work, which were more requested by women. This phenomenon is particularly pronounced in the south (excluding Abruzzo) and in Sicilia, where high index values indicate disparities favouring women.

In 2022, the average values are even lower than those of 2018, likely because the return to in-person work had a greater impact on women's work quality than on men's. While in southern Italy this led to a move closer to parity, in central and northern Italy, the result was a further deterioration in women's situation, especially in Lombardia (75.2) and Toscana (75.8). The region that, over time (excluding 2022), remains closest to parity is Valle d'Aosta.

Figure 2 – *Regional synthetic Index of sub-domain Quality of Work. Years 2018-2022.*



Source: authors' elaboration.

3.3. Sub-domain Money

On average, throughout the period under consideration (2018-2022), women living in households at risk of poverty are about two percentage points higher than men. The onset of the pandemic has led to a worsening for both sexes, with the gap narrowing by 2021. However, for women the situation worsened again in 2022.

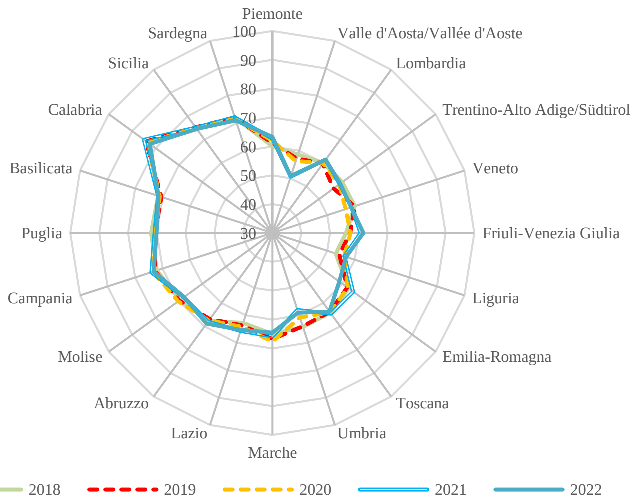
Regarding gross hourly wages, women earned an average of 12.9 euros, also during the period under consideration, compared with an average amount for men of 14.2 euros. Pay tends to increase over time, albeit slightly, with no consequences due to the pandemic. The average monthly pensions amount for females is almost half that of males (644.9 and 1,133.5 euros, respectively). For both sexes, however, the amount tends to increase over time.

Confirming the large monetary disadvantage of women, the GERI values for this subdomain are far from parity and clearly to the disadvantage of “pink quotas”: in fact, they range between 50.6 (Valle d'Aosta 2021) and 85.0 (Calabria 2021). Except in the case of a few regions, the phenomenon has remained rather unchanged over time, with a slight worsening in the last year considered, and, as shown in Figure 3, the values furthest from parity were recorded in the regions of the Central-North.

The regions furthest from parity on average are those in the North. Specifically, Valle d'Aosta and Liguria have the lowest values: the index never reaches 60, and Valle d'Aosta in 2021 and 22 recorded the lowest values ever (50.6 and 50.7).

In general, in the South, the situation is slightly better. Calabria is the region with the overall highest values of the index: between 82.5 and 85.0 but no temporal trend is evident. The other regions with higher GERI values, albeit far from parity, are Sicilia and Campania.

Figure 3 – *Regional synthetic Index of sub-domain Money. Years 2018-2022.*



Source: authors’ elaboration.

4. Final remarks

The gender gap, in a work and economic context, is caused by several factors, including gender discrimination, differences in work experience, and occupational segregation, with some regional variances.

In the labour market, women suffer persistent disadvantages in employment compared to men. The difficulty is more evident in the southern regions and in particular in Campania, Puglia and Sicilia. Career interruptions due to childbirth lead to a reduction in female employment over the life course (Ortiz-Ospina *et al*., 2018; EIGE, 2024). There are large gender differences in the amount of work done by women and men and in the type of job and contract performed. Among the three sub-domains considered, the one relating to the Quality of work presents taller values of GERI, particularly in south (except Abruzzo) and Sicilia. So, in these regions, the disparity is in disfavour of male.

Gender inequalities measured in the domain of Money are the visible outcomes of wide-ranging inequalities in other domains of life. Women's lower access to financial and economic resources reflects their heavier load of unpaid care within the household, so they are more likely to be at risk of poverty (Ortiz-Ospina *et al*., 2018; EIGE, 2024). Besides, women are more likely than men to work in sectors characterised by lower pay. Finally, women pensions amount is almost half that of men. The disparity is more accentuated in the regions of Central-Northern Italy, particularly in Valle d'Aosta and Liguria.

Of course, in Italy there are many important regional differences in the aspects considered in this research, so regional monitoring is necessary. The results for the three sub-domains showed that the construction of an index capable to read the distance from equality, alongside the reading of the individual indicators, could be an important informational input for monitoring the issues over time.

This would be useful for policymakers considering our nation's programmatic choices within the context of the national strategy for gender equality, since the goal must be equality for both sexes.

Acknowledgements

The authors would like to thank Matteo Potenzieri for his support in data quality control.

References

AMICI, M., STEFANI, M. L. 2013. A gender equality index for the Italian regions. *Questioni di Economia e Finanza (Occasional Papers),* No.190.

BERICAT E. 2012. The European Gender Equality Index: Conceptual and Analytical Issues, *Social Indicators Research*, Vol. 108, pp. 1–28.

CASCELLA, C., WILLIAMS J., PAMPAKA M. 2022. An Extended Regional Gender Gaps Index (eRGGI): Comparative Measurement of Gender Equality at Different Levels of Regionality. *Social Indicators Research*, Vol. 159, pp. 757–800.

EUROPEAN INSTITUTE FOR GENDER EQUALITY (EIGE). 2024. Gender Equality Index.

DI BELLA, E., LEPORATTI, L., GANDULLIA, L., MAGGINO, F. 2021. Proposing a regional gender equality index (R-GEI) with an application to Italy. *Regional Studies* Vol. 55, No. 5. pp. 962-973.

FUSCO, D., LIGUORI, M.A., PAGLIUCA, M.M. 2023. Gender gap of Italian region: a synthetic index proposal. *Rivista Italiana di Economia Demografia e Statistica* Vol. LXXVII No.4 Ottobre-Dicembre 2023.

INTERNATIONAL LABOUR ORGANIZATION (ILO). 2025. Women and the economy: 30 years after the Beijing Declaration.

KLASEN S. 2006. UNDP's gender-related measures: Some conceptual problems and possible solutions, *Journal of Human Development and Capabilities*, Vol. 7, No. 2, pp. 243–274.

JOST M., MÖSER S. 2023. Salary, flexibility or career opportunity? A choice experiment on gender specific job preferences. *Front. Sociol*. Vol. 8:1154324.

MAZZIOTTA M., PARETO A. 2016. On a generalized non-compensatory composite index for measuring socio-economic phenomena, *Social Indicators Research*, Vol. 127, pp. 983–1003.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2022. Same skills, different pay: Tackling gender inequalities at firm level.

ORTIZ-OSPINA E., HASELL J., ROSER M. 2018. Economic Inequality by Gender. Our World in Data. (Last change to text: March 2024).

PERMANYER I. 2010. The measurement of multidimensional gender inequality: Continuing the debate. *Social Indicators Research*, Vol. 95, pp. 181–198.

Daniela FUSCO, Istat, dafusco@istat.it
Cira ACAMPORA, Istat, acampora@istat.it
Paola GIORDANO, Istat, pgiordano@istat.it
Maria A. LIGUORI, Istat, liguori@istat.it
Margherita M. PAGLIUCA, Università Parthenope, margherita.pagliuca@uniparthenope.it

THE BRAIN WASTE IN ITALY: QUANTIFICATION AND CHARACTERISTICS OF OVER-QUALIFIED EMPLOYMENT[1]

Simone De Angelis, Valeria Quondamstefano

Abstract. Overqualification describes an employment situation where an employee's education exceeds job requirements. This mismatch between acquired skills and and what is necessary to perform a certain occupation is an important indicator of labour market inefficiencies, with potential repercussions both at the individual level (in terms of job satisfaction, motivation and professional development) and at the collective level (in terms of optimal utilisation of human capital). The analysis is based on the data made available by the 2021 edition of the Permanent Population and Housing Census, which will be compared with those from the 2011 General population Census in order to provide an evolutionary picture of the phenomenon over the decade considered. The study aims to capture the changes that have taken place from a quantitative perspective, i.e. in terms of the incidence and distribution of overqualified employment, also with reference to the socio-demographic characteristics of the people involved. For a more detailed understanding of the issue, specific indicators have been developed to describe its incidence in relation to gender, age and citizenship. The analysis have been carried out with a high geographical detail, down to the NUTS3 level, in order to detect any local inequalities to also support the definition of targeted policies on the territory.

1. Introduction

In recent years, the Italian labour market has shown signs of recovery, as evidenced by both a decline in the unemployment rate and an increase in employment. However, despite this growth, the employment rate for the entire population aged 15 years and over remains below the OECD average, as does that of university graduates, which, although rising, remains below the mean of the most advanced countries.

It is nevertheless undeniable that holding a university degree provides a clear employment advantage over a high school diploma. In 2024 (Istat, 2024), in the 15–89 age group, the employment rate for high school graduates stood at 57.8%,

[1] Sections are attributed as follows: section 1 to Simone De Angelis and sections 2, 3 and 4 to Valeria Quondamstefano

whereas for graduates or postgraduates it was over fifteen percentage points higher, at 73.4%.

Investing in education thus continues to represent a decisive factor for improving job opportunities and reducing the gap with the rest of Europe. However, the Italian labour market also exhibits significant contradictions: in recent years, there has been an increase in the number of highly educated workers employed in positions that do not require such qualifications. This phenomenon reflects a vertical skill mismatch — a misalignment between the qualifications held and those required — which differs from horizontal mismatch, where a worker is employed in a qualified position that does not match their field of study (Brandi *et al.,* 2017).

In the 1970s, some economists hypothesized that the increasing enrollment in tertiary education in developed countries would generate an oversupply of graduates that the labour market would be unable to absorb. In 1976 was coined the term overeducation to describe this situation (Freeman, 1976). However, these predictions were not borne out: already at the end of Fordism, and later in the 1980s with the rise of automation and IT-driven production processes, the demand for skilled labour expanded. The labour market responded to any temporary oversupply of qualified workers by further developing knowledge-based industries.

The prolonged economic crisis that began in 2007–2008, accompanied by a sharp reduction in the availability of venture capital for investments in high-growth sectors, disrupted this self-regulation mechanism. This has pushed many highly educated individuals to accept any type of job to avoid long periods of unemployment. As a result, overqualification is now a widespread and growing phenomenon.

The causes are not to be found solely in an excess supply of graduates relative to demand but also in other structural factors. In seminal work *Human Capital* (Becker, 1964) analysed the relationship between education, training and earnings, arguing that investment in skills improves not only individual economic outcomes but also collective welfare. Human capital, in fact, consists not only of formal education but also of general and job-specific work experience.

In Italy, the weak integration between the education system and businesses is considered one of the main factors behind both the difficult labour market entry for young people and the high incidence of overqualification. To address this issue, both the Jobs Act and the education reform known as "*La Buona Scuola*" (Law 107/2015) introduced measures to strengthen on-the-job training pathways, integrating them into the curricula of schools and universities.

In addition to the quality of education, the economic context must also be considered. A recent OECD report (OECD, 2018) highlights that, despite improvements in employment rates, productivity in Italy remains unsatisfactory. This is attributed to relatively low skill levels, weak demand for advanced skills from firms and limited use of available skills. The report describes Italy as being "trapped in a low-skill equilibrium," a situation in which a low supply of skills is accompanied by low demand from businesses. While some large firms compete successfully on global markets, many others — often small or medium-sized and family-owned — exhibit low levels of managerial, minimal investment in technologies and productivity-enhancing work practices, and limited incentives to invest in human capital.

The OECD identifies a central cause in the Italian business structure: 85% of Italian firms — employing about 70% of the total workforce — are family-owned, mostly small or medium-sized, and operate in medium- or low-technology sectors. These firms generally do not prioritize hiring highly qualified personnel but rather focus on minimizing labour costs to remain competitive.

A more recent framework, the career mobility theory (Sicherman, Galor, 1990), posits that individuals initially accept jobs below their qualification level in order to acquire work experience and job-specific human capital, ultimately facilitating upward career mobility through promotions or transitions to positions requiring higher skills. From this perspective, overqualification would represent a physiological and temporary stage, reflecting a deliberate strategy among younger workers who accept the first available job but continue searching to improve their position, either within the same company or elsewhere. However, as will be shown in the following empirical analysis, current data do not seem to fully support this interpretation.

This study, therefore, aims to explore the phenomenon of overqualification, defined as the condition in which an individual is employed in a job that requires a lower level of education than they possess. This misalignment between acquired skills and job requirements represents a significant indicator of labour market inefficiencies, with potential negative consequences both at the individual level (in terms of satisfaction, motivation, and career development) and at the collective level (in terms of human capital wastage, lower productivity, and inefficient public spending on education).

2. Data source and criteria of selection

To study the phenomenon of overqualification between 2011 and 2021, microdata from the 15th Population and Housing Census (2011) and the Permanent Population and Housing Census (2021) were considered. The Permanent Census collects information on occupations according to the ISCO 08 COM classification (10 categories relating to large occupational groups). Of these 10 categories, the first six requiring the lowest level of education were selected for analysis. For each individual, their level of education and occupation were considered. If the profession requires a lower level of education than that possessed by the individual, then that individual is classified as overqualified. Specifically, regarding education, the following levels were included: Bachelor's degree or first-level academic diploma, Higher Technical Diploma, University Diploma; Master's degree or second-level academic diploma (including the former Academy and Conservatory programmes); and Research Doctorate (PhD) or advanced research academic diploma. As for employment, the categories of occupation examined were: Manual and unskilled labor (Farm hand, Custodian, Construction worker, Domestic assistant, Dishwasher, Usher, Porter, Hospital attendant, Refuse collector, Stablehand); Operation of manufacturing systems, machinery and assembly lines, driving vehicles (Forklift operator, Assembler of electric devices, Truck driver, Taxi driver, Automatic loom operator, Rolling mill operator, Oil mill operator); Skilled labour (Bricklayer, Mechanic, Heating system installer, Shoemaker, Tailor, Carpenter, Blacksmith, Upholsterer); Plant cultivation and/or animal breeding (Farmer, Fruit grower, Stockman, Fish farmer, Reforester, Gardener, Fisherman); Retail sales and services (Shopkeeper, Police officer, Hairdresser, Cook, Waiter, Flight attendant, Baby sitter, Nanny, Salesperson); and administrative support (Secretary, Postal service counter worker, Switchboard operator, Administrative assistant, Service counter staff). An example of overqualification could be a graduate working as a farm hand.

3. Overqualification in Italy: a descriptive analysis

In 2011, the total number of overqualified workers was 993,206, representing 22.3% of the employed with a high level of education (4,446,437 individuals). By 2021, this number had increased to 1,731,747, accounting for 29.9% (5,791,999 individuals). This meant that between 2011 and 2021, there was therefore an absolute increase of 738,541 overqualified workers, corresponding to a relative growth of +74.4%.

The presence of overqualified workers is particularly high in the more industrialized provinces and metropolitan cities with greater employment potential. In 2011, 271,865 overqualified workers (accounting for 27.4% of the total) resided in Rome, Milan, Turin, Naples and Bologna. By 2021, this number had increased to 490,522, representing 28.3% of all overqualified workers in these same metropolitan cities (Table 1). The overqualification rate is defined as the percentage ratio between the number of overqualified individuals (numerator) and the total number of highly educated workers (denominator).

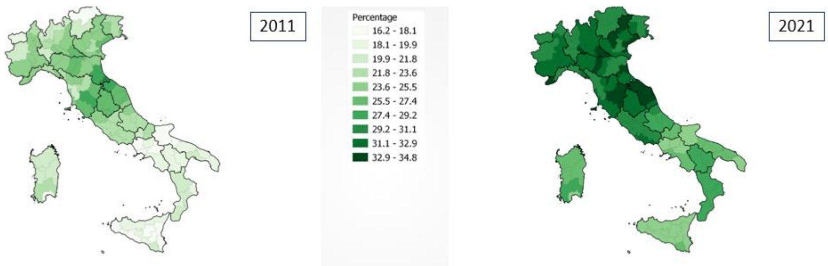
Table 1 – *First 10 Provinces/Metropolitan Cities with Highest level of overqualification – Years 2011-2021.*

Provinces/Metropolitan cities	Absolute Value	Percentage	Provinces/Metropolitan cities	Absolute Value	Percentage
Rome	100,049	10.1	Rome	169,483	9.8
Milan	75,117	7.6	Milan	149,570	8.6
Turin	38,975	3.9	Turin	69,295	4.0
Naples	30,894	3.1	Naples	57,724	3.3
Bologna	26,830	2.7	Bologna	44,450	2.6
Florence	21,809	2.2	Florence	37,395	2.2
Padua	18,941	2.0	Brescia	34,730	2.0
Brescia	18,496	1.9	Padua	32,667	1.9
Genoa	17,275	1.9	Bergamo	31,251	1.8
Venice	16,490	1.7	Monza and Brianza	30,132	1.7

Source: *Elaborations on data from Permanent Population and Houses Census (PPHC), Istat*

Figure 1 shows that in 2011, the provinces with the highest overqualification rates were Rimini (29.6%), Ravenna (28.3%), and Siena (28.0%). In contrast, the lowest rates were observed in Bolzano/Bozen (16.2%), Agrigento (16.5%), and Cagliari (16.7%). By 2021, the areas most affected by overqualification were Mantua (34.8%), Prato (34.7%), and Macerata (33.8%), while the lowest rates were recorded in Enna, Messina, and Agrigento (24.0%).

Figure 1 – *Overqualification rate: Provincial breakdown – Years 2011-2021.*



Source: *Elaborations on 15th Population and Houses Census (PHC) and PPHC data, Istat.*

3.1. Overqualification of certain disadvantaged categories: women, young people and foreigners

In 2011, the overqualified females were 630,567, representing 63.5% of the total amount of overqualified workers (993,206 individuals). By 2021, this number had risen to 1,029,156, accounting for 59.4 per cent of the total overqualified (1,731,747 persons). This meant that between 2011 and 2021, there was therefore an absolute increase of 398,589 overqualified workers, corresponding to a relative growth of +63.2%.

As can be seen from Figure 2, in 2011 the highest levels of females overqualification were recorded in the provinces of Rimini (35.4%), Ravenna (33.7%), and Pesaro-Urbino (33.2%), whereas the lowest levels were found in Agrigento (17.2%), Foggia (18.7%), and Bolzano/Bozen (19.3%). By 2021, the phenomenon was most pronounced in Arezzo (50.5%), Belluno (50.3%), and Siena (49.3%), while it remained comparatively limited in Agrigento (25.1%), Caltanissetta (25.6%), and Siracusa (26.4%).

Figure 2 – *Overqualification rate: female breakdown – Years 2011-2021.*

Source: Elaborations on 15th PHC and PPHC data, Istat.

An analysis of the population by age group reveals that overqualification decreases progressively with age (Table 2).

Table 2 – *Overqualified workers by age group – Years 2011-2021.*

Age group	Overqualification Rate (%) Year 2011	Overqualification Rate (%) Year 2021
15-29	38.1	41.5
30-34	29.4	36.4
35-39	24.7	33.4
40-44	20.8	31.5
45-49	17.3	27.3
50-54	14.0	25.0
55-59	11.2	21.6
60-64	9.8	19.1
65 and over	6.2	13.4
Total	22.3	29.9

Source: Elaborations on 15th PHC and PPHC data, Istat.

In 2011, the share of underemployed university graduates dropped below 25% starting from the 35–39 age group. By 2021, however, this threshold was not reached until the 55–59 age group. This trend suggests that overqualification became a more structural and persistent phenomenon over the decade 2011–2021, with individuals who were underemployed in 2011 likely remaining in the same condition in 2021.

The most critical situation in both years under analysis was observed among young individuals aged 15–29 entering the labour market, who faced the highest incidence of overqualification, highlighting persistent structural barriers to adequate job matching at labour market entry. As shown in Figure 3, in 2011, the overqualification rate among 15–29-year-olds was highest in Grosseto (48.2%), Venice (45.2%), and Gorizia (45.1%), and lowest in Bolzano/Bozen (25.1%), Agrigento (30.5%), and Isernia (30.7%). A decade later, in 2021, the highest rates remained in Grosseto (47.0%), Pesaro-Urbino (46.9%), and Livorno (46.5%), while the lowest rates have been recorded in Isernia (31.8%), Enna (35.5%), and Agrigento (35.7%). These findings suggest that, despite some regional variation, overqualification among young entrants has remained a widespread and enduring phenomenon over the past decade.

Figure 3 – *15-29 years olds Overqualification Rate: Provincial Breakdown -Years 2011-2021.*

Source: Elaborations on 15th PHC and PPHC data, Istat.

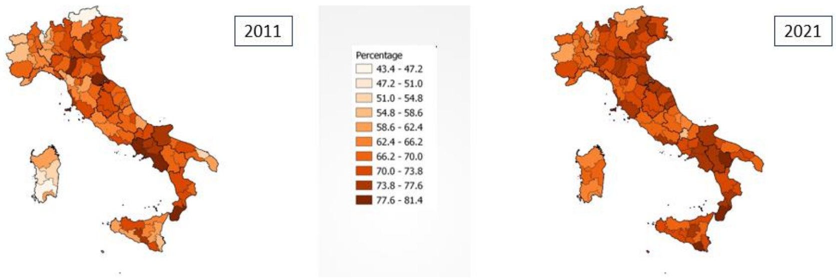
The overqualification also affects foreign workers[2]. In 2011, there were 160,126 overqualified foreigners in Italy, representing 16.1% of the total overqualified workers (993,206). By 2021, this number had increased in absolute terms to 211,669 individuals (an increase of 51,543). However, in relative terms, in ten years their share decreased to 12.2%, marking a decline of 3.9 percentage points.

When looking at provincial data (Figure 4), in 2011, foreign overqualification was most widespread in Caserta (80.9%), Reggio di Calabria and Naples (80.7%), while the lowest rates were observed in Bolzano/Bozen (43.4%), South Sardinia (47.0%), and Oristano (48.5%). By 2021, the highest figures were recorded in

[2] Foreign nationals, who obtained their highest qualification abroad, must select the corresponding qualification in Italy.

Ragusa (81.4%), Reggio di Calabria (78.2%), and Grosseto (77.7%); the lowest remained in Isernia (56.2%), Trieste (59.6%), and Turin (60.7%).

Figure 4 – *Foreign Overqualification Rate: Provincial Breakdown – Years 2011-2021.*



Source: *Elaborations on 15th PHC and PPHC data, Istat.*

4. Conclusions and next steps

Between 2011 and 2021, overqualification — reflecting a significant waste of human capital — has become even more pronounced.

From a territorial perspective, overqualification is more prevalent in provinces and metropolitan cities with higher employment potential, particularly in Northern and Central Italy. Data from the 2021 Census, when compared to 2011, reveal an increased disadvantage for women, young people, and foreign citizens.

It is interesting to observe, on 2021, that the provincial distribution of overqualification among women and young people largely reflects the general pattern observed at the national level. In fact, the provinces recording the highest incidence of overqualification are predominantly located in central and northern Italy—areas traditionally characterized by more dynamic labor markets and higher levels of educational attainment. A few notable exceptions emerge, however, such as Naples and Bolzano, where the rates deviate from this general trend.

In contrast, the phenomenon of overqualification among foreign workers displays a markedly different geographical pattern, being more widespread in the southern provinces and on the islands. This divergence may be linked to structural factors such as limited employment opportunities, labor market segmentation, and the concentration of migrants in low-skilled sectors, which tend to offer fewer opportunities for the full utilization of educational and professional skills.

The greater incidence of overqualification among women is associated with several structural factors, including persistent gender inequality in the labour market, their concentration in low-paid or part-time jobs, limited access to leadership roles, and challenges in balancing work and family responsibilities.

Among young workers, the higher rates of overqualification are mainly linked to limited work experience despite high educational attainment, high competition for qualified positions, mismatches between academic training and labour market needs, difficulties accessing high-skilled entry-level jobs and a greater willingness to accept any employment in order to gain experience.

For foreign workers, overqualification is driven by multiple barriers, such as problems in the recognition of educational qualifications obtained abroad, language obstacles that limit access to skilled positions, regulatory or bureaucratic hurdles in achieving job equivalency, length of stay in Italy, and the need to accept any available job to ensure economic survival or residency stability.

Future steps in the analysis will include a detailed examination of overqualification by economic activity sector, as well as by status in employment, distinguishing between employees and self-employed workers. Additional studies will consider the type of employment, including disparities in working hours and contract types. An in-depth study of large municipalities (more than 150,000 usual residents) will also be conducted to better understand urban dynamics. In addition, the combined effects of various socio-demographic and labour market factors on the probability of overqualification will be identified and quantified through multivariate analysis techniques.

References

BRANDI M.C., CARUSO M.G., DE ANGELIS S., MATROLUCA S. 2017. L'evoluzione del mercato del Lavoro degli immigrati laureate: il fenomeno della "overeducation" dai censimenti del 2001 e 2011, *International Journal of Migration Studies*, Vol. LIV, No. 206, pp. 181-213.

BECKER G. S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, New York: National Bureau of Economic Research.

FREEMAN R., 1976. *The Overeducated American*. New York: Academic Press.

ISTAT, 2024, Istat databrowser. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,Z0500LAB,1.0/LAB_OFFER/LAB_OFF_EMPLOY/DCCV_TAXOCCU1/IT1,150_915_DF_DCCV_TAXOCCU1_2,1.0

OECD, 2018. *Skills Strategy Diagnostic Report: Italy 2017*, Paris: OECD Publishing. https://doi.org/10.1787/9789264298644-en

SICHERMAN N., GALOR O. 1990. A Theory of Career Mobility, *Journal of Political Economy*, Vol. 98, No.1, pp. 169-192. https://doi.org/10.1086/261674

Simone DE ANGELIS, Istat, sideange@istat.it
Valeria QUONDAMSTEFANO, Istat, quondamstefano@istat.it

JOB VACANCY RELATION WITH LABOUR INPUTS AND UNEMPLOYMENT: PRELIMINARY EVIDENCE FROM ECONOMIC SECTORS

Pasquale di Padova, Annalisa Lucarelli, Emilia Matera

Abstract. Job vacancy statistics (JVS) are key leading indicators of the business cycle and are essential for structural labour market analysis, particularly in identifying mismatches between labour demand and supply. Despite their importance, Italy has yet to fully exploit demand-side data or evaluate the predictive capacity of JVS for employment and unemployment trends. This paper investigates whether changes in job vacancy rates can serve as leading indicators of economic activity by analysing their relationship with various labour input indicators, including occupied positions, hours worked, and jobs filled by temporary agency workers. The analysis uses quarterly time series and growth rate comparisons from 2016 to 2024, based on data from three official Istat sources: Employment, Wages, Salaries, and Social Contributions; Job Vacancies and Hours Worked; and Labour Force. A central component is the construction of Italy's Beveridge curve to explore the dynamics between labour demand and supply. Results indicate that, overall, the vacancy rate leads labour input growth by one quarter, though some sectors show contemporaneous relationships. The findings also reveal both cyclical and structural shifts in the vacancy-unemployment relationship, laying the groundwork for future modelling of job matching dynamics.

1. The JVS Survey within the labour market framework

The labor market is shaped by the constant interaction between the supply and demand for labor. Supply is represented by individuals seeking work, while demand is expressed by enterprises needing employees. Employment outcomes result from the alignment of these two forces through job matching, which determines the efficiency of resource allocation. This interaction creates three main areas of analysis. The first is the matched segment, where supply meets demand. This segment, covering the number of employed individuals and hours worked, is a key indicator of market performance. The other two areas highlight unmet needs: unemployment (individuals seeking work but unable to find it) and job vacancies (open positions without suitable candidates). These imbalances signal structural or cyclical issues that hinder efficient resource allocation.

To understand the full complexity of the labor market, it's crucial to adopt a dual perspective, analyzing both individual behaviors and enterprise decisions. This approach requires data from both households and businesses to measure not only employment but also vacancies, hours worked, and job quality. This integrated view helps identify mismatches between skills and demand, as well as broader socioeconomic dynamics.

The provision of timely, accurate, and harmonized labour market statistics is fundamental for economic analysis, policymaking, and business strategy at both national and European levels. A key instrument in this domain is the quarterly Job Vacancy and Hours Worked Survey (JVS), a comprehensive statistical framework designed to measure critical labour market dynamics. The survey provides indispensable data on labour demand, labour input, and emerging economic pressures, adhering to stringent European Union regulations while also generating valuable national-level indicators.

The JVS is a quarterly sample survey designed to be representative of the private non-agricultural industry and service sectors. The target population consists of active enterprises and private institutions resident in the national territory with an employment size ranging from 1 to 499 employees. The survey operates with a substantial sample of approximately 29,000 enterprises. This sample is strategically stratified to ensure robust estimates, comprising around 16,000 enterprises with fewer than 10 employees and approximately 13,000 enterprises in the 10-499 employee range. To maintain the statistical validity of the panel over time while minimizing the response burden on individual units, the sample is rotated by one-third annually. This rotational design ensures that two-thirds of the sample remains consistent between consecutive years, thereby improving the accuracy of year-on-year comparisons.

The survey's economic scope is defined according to the NACE Rev. 2 statistical classification of economic activities. It covers sections from B (Mining and quarrying) to S (Other service activities), comprehensively encompassing the majority of the private industry and service economy. Notably, it excludes section A (Agriculture, forestry and fishing) and section O (Public administration and defence; compulsory social security), which are outside its primary focus.

Data collection is conducted using a modern, efficient methodology. The primary instrument is a quarterly questionnaire administered via Computer-Assisted Web Interviewing (CAWI). Respondents access and complete the questionnaire through the official ISTAT Business Portal, a secure online platform. This approach facilitates timely data submission and has proven effective, yielding a high response rate of approximately 70%.

The survey captures the total number of job positions by measuring the stock of employees at the end of both the reference quarter and the previous quarter. This longitudinal element is supplemented by flow data, which includes workers who both started and ended their employment during the reference quarter.

The definition of a job vacancy is strictly aligned with EU Regulation. A job vacancy is defined as a paid post that is newly created, unoccupied, or about to become vacant for which the employer is taking active steps to find a suitable candidate and intends to fill either immediately or within a specific period of time. The survey further specifies that a vacancy exists at the point where a suitable candidate could have started working immediately. An additional qualitative dimension is captured by identifying vacancies for which enterprises report particular difficulties in finding suitable candidates, offering insight into skills mismatches in the labour market.

The concept of hours worked is disaggregated into three primary components to provide a detailed view of labour input: ordinary hours; overtime hours; hours paid but not worked. Furthermore, the survey monitors key indicators of labour underutilization. It collects data on hours lost due to strikes, for reasons both related and unrelated to the specific employment relationship. It also measures the impact of state-supported schemes through the short-time working allowance, a government instrument providing financial support when businesses are forced to downsize or suspend activities. This includes ordinary, exceptional, and extraordinary interventions. Finally, it quantifies the reduction of working hours resulting from the application of solidarity contracts (as per Law 863/84).

To enhance the robustness of its estimates and provide a complete picture of the economy, the JVS integrates data from other key statistical sources.

For enterprises with 500 or more employees, which are outside the direct scope of the JVS sample, the survey incorporates micro-data on employees, hours worked, and job vacancies from the Monthly Large Enterprises Survey (LES). This integration ensures that the final estimates reflect the entire business economy. Additionally, the JVS utilizes both micro-data and macro-data on job positions from the Quarterly Survey on Employees, Gross Wages and Social Contributions (OROS). This information is crucial for estimation and calibration processes, ensuring the final survey weights accurately reflect the structure of the target population.

The JVS produces a series of key statistical outputs with staggered release-timetables, adhering to the requirements of European regulations. The main outputs include:

- Flash Job Vacancy Rate: preliminary estimates are released within 45 days of the end of the reference quarter, providing an early signal of labour market tightness.
- Final Job Vacancy Rate: the definitive rate is published within 70 days, in compliance with Regulation (EC) No 453/2008.

- Total Hours Worked: comprehensive data on hours worked is disseminated within 90 days, following the framework of Regulation (EU) No 2019/2152 on European business statistics.

The indicators produced are particularly relevant to the labour market. Notably, the Job Vacancy Rate is included among the Principal European Economic Indicators (PEEIs) — a set of key macroeconomic indicators that provide timely and high-quality information for monitoring short-term economic developments of the European Union (EU) and the euro area.

In addition to the above-mentioned indicators, required by Eurostat, the survey generates a set of valuable indicators for national-level dissemination. These include per capita hours worked, overtime hours as a percentage of total hours worked, and the rate of short-time working allowance hours per 1,000 hours worked, offering deeper insights into productivity, work intensity, and the economic cycle.

2. The relation between vacancies and labour inputs

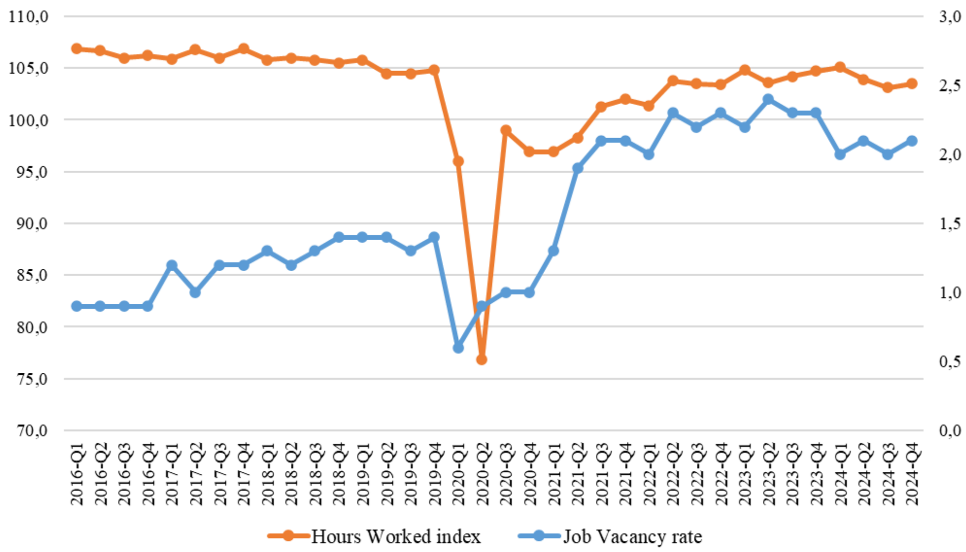
In the short term, labour demand is closely linked to fluctuations in enterprise output. When production increases, firms initially respond by extending the working hours of current employees, both through ordinary and overtime hours. If the growth in output persists, businesses initiate new recruitment activities, which eventually lead to a rise in employment levels. This sequential response introduces a temporal gap between the initial rise in output and the actual hiring of additional staff, during which job vacancies emerge.

Job vacancies thus function as an intermediate stage in the adjustment of labour inputs and serve as early signals of forthcoming employment changes. Because they reflect the hiring intentions of firms, vacancies are particularly valuable as short-term indicators of economic activity. An observed increase in vacancies often anticipates a rise in employment in subsequent periods, making them a key variable for economic forecasting. Job vacancy statistics, therefore, play a critical role in linking short-term economic dynamics to labour market adjustments.

Building on this theoretical relationship between labour demand and labour inputs, this section delves into the empirical connection between job vacancies, hours worked and employment. A key mechanism through which firms respond to short-term fluctuations in labour demand, particularly during periods of economic expansion, is by adjusting regular working hours before altering headcount.

Figure 1 illustrates the strong correlation between changes in per capita hours worked and the job vacancy rate, a widely used proxy for labour demand. This dynamic adjustment was particularly pronounced during the COVID-19 pandemic. In the early stages of the crisis, total hours worked declined sharply, mirroring the abrupt drop in labour demand triggered by widespread lockdowns. As economic conditions began to stabilize, firms quickly increased hours worked to meet rising demand, even before significantly expanding their search for new workforce.

Figure 1 – *Quarterly Job Vacancy rate (right scale) vs per capita Hours Worked index. Seasonally adjusted data 2016-2024; NACE Rev.2 sectors B to S; 1+ employees.*



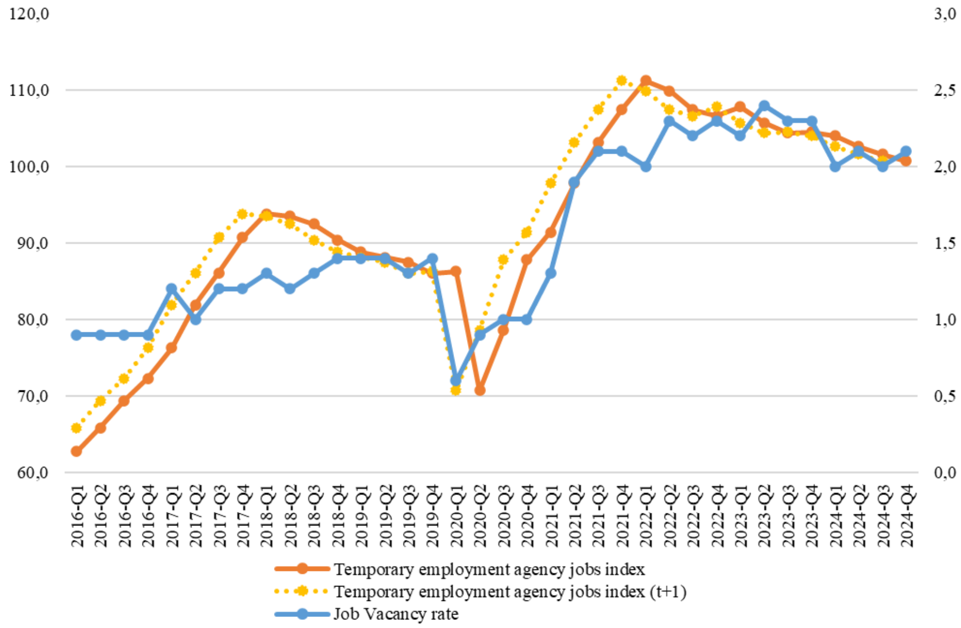
Source: JVS.

From the third quarter of 2021 onward, hours worked once again began to closely align with movements in the vacancy rate, reflecting a normalization of labour market dynamics. This pattern underscores the role of working hours as a flexible and immediate response tool for firms managing emerging needs in labour demand.

Following the previously discussed dynamics of per capita hours worked, which similarly reflect firms' short-term responses to shifting labour demand, this section turns to the next stage in the labour input adjustment process: the use of temporary agency work. This form of employment offers firms a flexible means of scaling labour capacity in response to cyclical fluctuations, particularly when demand increases are perceived as potentially transitory.

Figure 2 presents a comparison between the job vacancy rate and the temporary employment agency jobs index, examined both contemporaneously and with a one-quarter lag applied to the latter. The time series analysis reveals a clear pattern: the vacancy rate consistently functions as a leading indicator for other labour input measures. The delayed response observed in temporary agency work highlights how firms initially rely on job postings and internal labour adjustments, such as regular and overtime hours, before expanding their use of external, flexible labour arrangements.

Figure 2 – *Quarterly Job Vacancy rate (right scale) vs Temporary employment agency jobs index. Seasonally adjusted data 2016-2024; NACE Rev.2 sectors B to S; 1+ employees.*



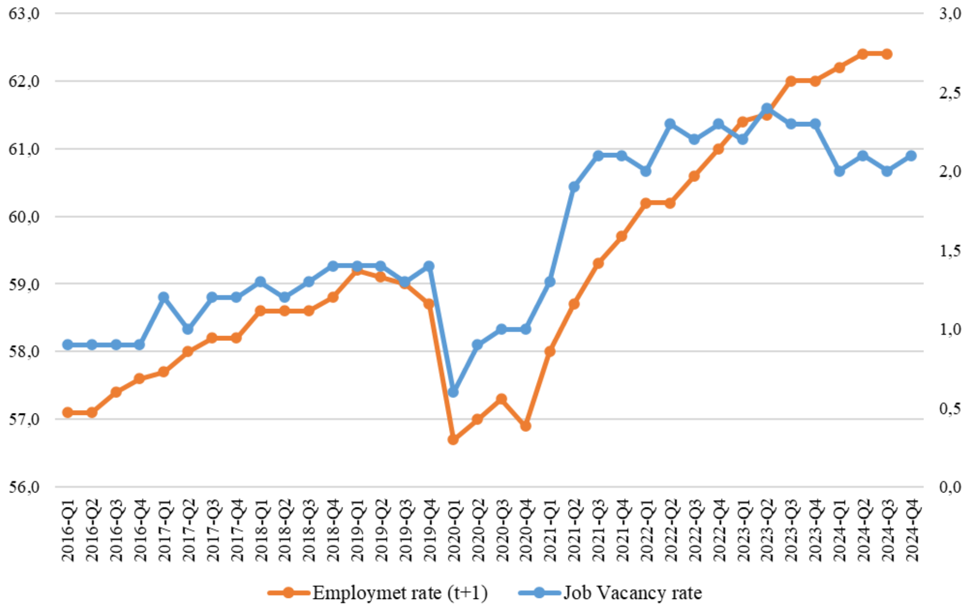
Source: JVS and OROS.

This relationship is especially evident in the periods both preceding the COVID-19 pandemic and following the onset of 2022, during which the lagged temporary agency jobs index aligns closely with movements in the vacancy rate. The coherence underscores the instrumental role of temporary employment in firms' adaptive strategies during economic recovery and expansion phases.

The last relationship explored in this section concerns the link between job vacancies and overall employment. This relationship can be examined through two lenses: the number of individuals employed or the total number of job positions. Both approaches offer valuable insights into how shifts in labour demand translate into realized employment outcomes.

Figure 3 illustrates the employment rate with a one-quarter lag, allowing for a clearer view of the temporal relationship between labour demand and subsequent employment developments. The data reveal a consistent pattern: changes in the job vacancy rate tend to precede corresponding movements in employment, typically with a lag of approximately one quarter. This suggests that job postings are a reliable early signal of hiring trends, reflecting firms' forward-looking behavior in response to evolving economic conditions.

Figure 3 – *Quarterly Job Vacancy rate (right scale) vs Employment rate (15-64 years). Seasonally adjusted data 2016-2024; JV rates for enterprises in NACE Rev.2 sectors B to S and with at least 1 employee.*



Source: JVS and LFS.

Although the figure focuses on the employment rate, similar conclusions hold when analyzing the total number of job positions. Due to space constraints, the corresponding graph for job positions is not included here, but the underlying dynamics remain consistent: both metrics reinforce the role of vacancies as a leading indicator in the employment adjustment process.

To provide a comprehensive overview of the results discussed above, Table 1 reports the cross-correlations between the job vacancy rate and different labor input measures at various time lags. For completeness, the correlation with the unemployment rate is also included and will be examined in greater detail in the following section.

The results indicate that the correlation between job vacancies and hours worked peaks at a one-quarter lag. A similar pattern is observed for the other labor input indicators. Specifically, the use of temporary agency work is strongly associated with job vacancy dynamics, reaching the highest correlation at t+1. The leading role of the job vacancy rate with respect to employment dynamics is likewise evident, as the correlation between the two series reaches its maximum values starting from a one-quarter lag. Finally, the inverse correlation with the unemployment rate highlights a consistent relationship between rising labor demand and falling unemployment. This relationship is statistically strong at a one-quarter lag and strengthens further over the next two quarters.

Table 1 – *Cross-Correlations among Job Vacancy rate with labour inputs and Unemployment rate at different time lags.*

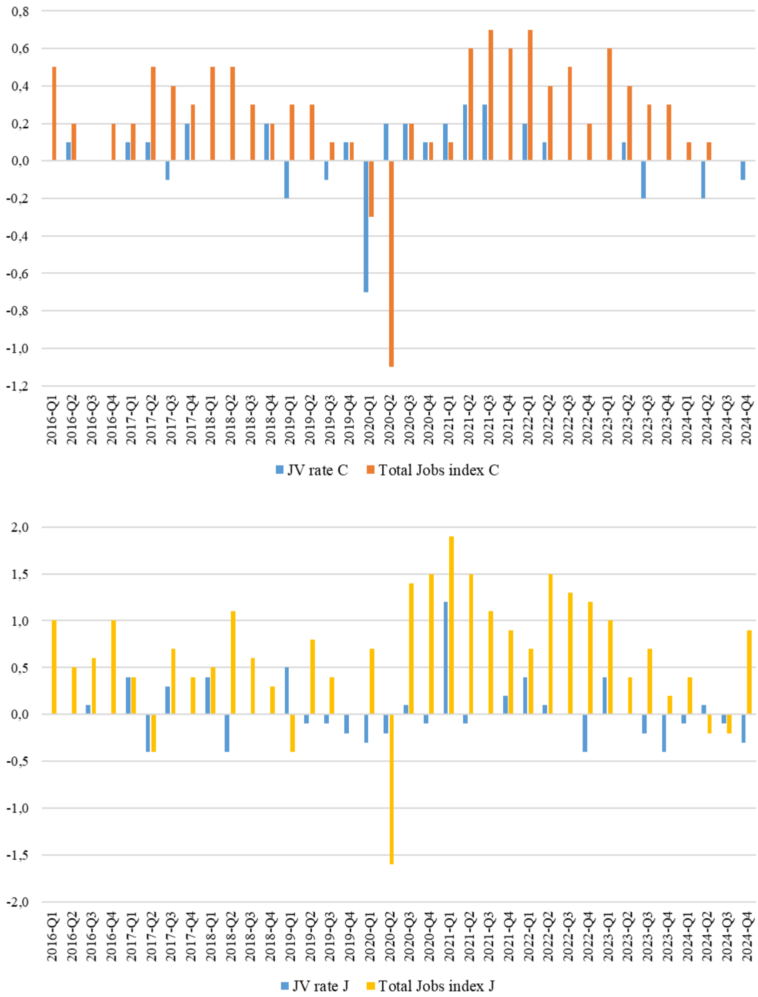
	Hours Worked index	Temporary employment agency jobs index	Employment rate	Unemployment rate
t-3	-0.14	0.77	0.51	-0.68
t-2	-0.03	0.85	0.63	-0.71
t-1	0.04	0.88	0.73	-0.75
t0	0.15	0.89	0.83	-0.78
t+1	0.24	0.90	0.90	-0.82
t+2	0.09	0.81	0.91	-0.88
t+3	0.13	0.70	0.91	-0.87

While the aggregate relationship between labour demand and employment is well established, a sectoral analysis offers valuable insights into how this dynamic may differ across distinct areas of the economy. Examining labour market adjustments at the sectoral level allows for the identification of structural or behavioral factors that may influence the responsiveness of employment to changes in labour demand.

To this end, the analysis focuses on the degree of alignment over the reference period between short-term fluctuations in the job vacancy rate and the corresponding percentage changes in the employment index across NACE Rev. 2 sections. This comparison enables a more nuanced understanding of how labour demand translates into employment outcomes in different economic contexts.

Figure 4 draws attention to two sectors in particular where this relationship appears especially robust: Section C (Manufacturing) and Section J (Information and Communication). In both sectors, more than in others, employment dynamics closely track movements in job vacancies, suggesting a heightened sensitivity of filled positions to shifts in labour demand. One plausible explanation for this strong correlation lies in the high concentration of skilled labour in these sectors. Highly qualified workers typically face fewer barriers to mobility and are often better positioned to respond swiftly to new opportunities.

Figure 4 – *Job Vacancy rate vs Total Jobs index in Manifacturing and Information and Communication sectors. NACE Rev.2: C, J; 1+ employees; absolute differences and percentage changes over the previous quarter.*



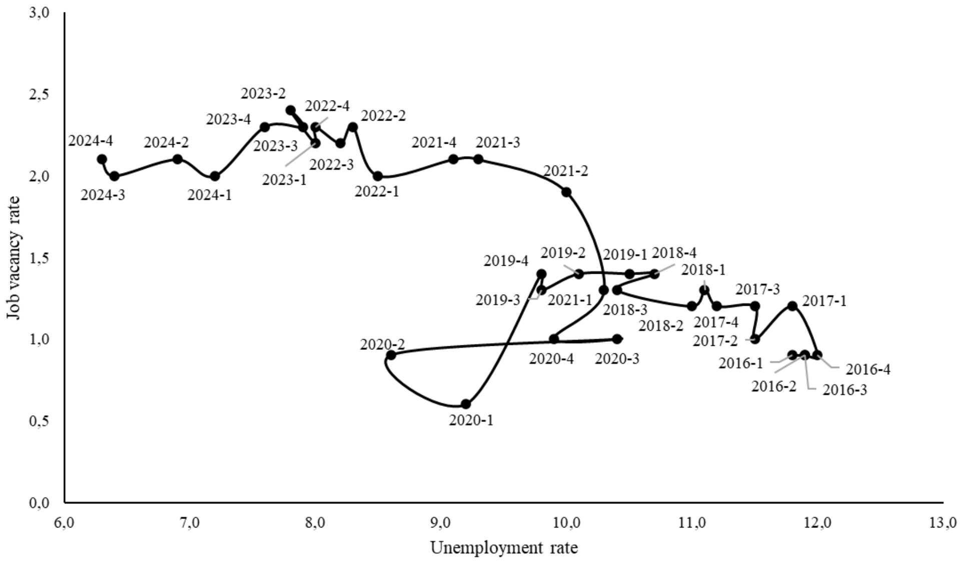
Source: JVS and OROS.

3. Job vacancy and unemployment: the Beveridge curve for Italy

While the relationship between job vacancies and various labour inputs has been analyzed in the previous section, further investigation is needed into the implications of mismatches between labour demand and supply, particularly in relation to unemployment.

A widely used analytical tool for exploring this relationship is the Beveridge Curve, which graphically illustrates the inverse relationship between the unemployment rate and the job vacancy rate. It is a standard instrument in labour economics for assessing the efficiency of labour market matching, also making possible to assess whether the dynamics of vacancies and unemployment reflect cyclical movements.

Figure 5 – *Job vacancy and Unemployment rates (15-64 years): Beveridge Curve. Seasonally adjusted data 2016-2024; JV rates for enterprises in NACE Rev.2 sectors B to S and with at least 1 employee*.



Source: JVS and LFS

As shown in Figure 5, in the Italian labour market over the period 2016–2024, a general trend of improved matching between labour demand and supply can be observed, particularly following the economic recovery after the double-dip recession[1]. Between 2016 and 2019, the unemployment rate steadily declined while labour demand increased, reflecting a phase of cyclical recovery accompanied by improved labour market efficiency. In 2020, the job vacancy rate dropped sharply due to the pandemic, although this was not immediately followed by a rise in unemployment, thanks to policy interventions such as dismissal bans.

With the gradual return to normality, the period from 2020 to 2021 saw a brief

[1] An analysis of the Beveridge curve for years 2008-2017 in Italy can be found in Istat *et al*. (2017).

uptick in unemployment, followed by a sharp and sustained decline beginning in 2021. At the same time, labour demand began to recover from late 2020 onward, with the job vacancy rate rising steadily through the end of 2023. In 2024, labour demand appeared to level off, but this stabilization did not reverse the downward trend in unemployment, which continued to fall, albeit at a slower pace.

This pattern — also reflected at the European level (Eurostat, 2025) — may point to a phase of relative efficiency in labour market matching, even amid slower job creation. Several factors could help explain this development: improved job search behavior among the unemployed, more targeted and effective recruitment strategies by employers, better alignment between available skills and job requirements, or a reduction in frictional unemployment due to enhanced information flow and mobility.

4. Conclusions

This study has examined the Italian labour market through the lens of the Job Vacancy and Hours Worked Survey (JVS), highlighting the critical role of job vacancies as both indicators of firm-level labour demand and predictors of broader employment dynamics. By integrating data from multiple sources and employing a harmonized statistical framework, the JVS provides a comprehensive view of labour input adjustments, including hours worked, temporary agency work, and changes in overall employment.

First, a robust positive correlation was established between the job vacancy rate and hours worked, highlighting that firms initially respond to shifts in output demand by adjusting the hours worked of their existing staff. This represents the most immediate and flexible margin of labour input adjustment. Subsequently, the analysis demonstrated that the job vacancy rate also precedes changes in temporary agency employment, typically with a temporal lag of one quarter. This sequential adjustment process reveals that after exhausting internal flexibility, firms turn to external, non-permanent staffing solutions. Ultimately, the study confirmed that job vacancies are a reliable leading indicator for overall employment, with changes in the vacancy rate consistently anticipating movements in the employment rate by approximately one quarter. This relationship was found to be particularly pronounced in skill-intensive sectors such as Manufacturing (NACE Rev. 2 section C) and Information and Communication (NACE Rev. 2 section J), suggesting that labour mobility and market responsiveness are heightened where specialized qualifications are in demand.

Beyond the direct relationship with labour inputs, the analysis extended to the broader interplay between labour demand and supply through the lens of the

Beveridge Curve. The examination of the vacancy-unemployment relationship in Italy from 2016 to 2024 revealed a notable improvement in the labour market's matching efficiency. Despite the significant economic shock induced by the COVID-19 pandemic, the curve's trajectory indicates an enhanced capacity to align job seekers with available positions, particularly during the post-pandemic recovery phase. This inward shift of the Beveridge Curve suggests that factors such as more effective recruitment strategies, improved skills alignment, or reduced frictional unemployment have contributed to a more efficient structural performance of the Italian labour market.

In summary, the converging evidence from the analysis of hours worked, temporary employment, aggregate employment levels, and the Beveridge Curve collectively affirms the central role of job vacancy statistics in labour market analysis. This encouraging preliminary body of evidence, which assesses and confirms that changes in the number of vacancies are a leading indicator, with a one-quarter lag, of labour input dynamics, paves the way for future research and analysis. In particular, applying multivariate time-series methodologies, such as vector autoregressive and error correction models, offers a promising way to capture the dynamic interdependencies among vacancies, hours worked, and employment. This will therefore be useful for monitoring and predicting labour market developments with greater precision.

References

EUROSTAT 2025. Job vacancy and unemployment rates - Beveridge curve, *Statistics Explained*, May 2025, https://tinyurl.com/buufa3xc.

GARASTO S., DJUMALIEVA J., KANDERS K., WILCOCK R., SLEEMAN C. 2021. Developing experimental estimates of regional skill demand, *ESCoE Discussion Paper 2021-02*, Economic Statistics Centre of Excellence, National Institute of Economic and Social Research.

ISTAT, MLPS, INPS, INAIL, ANPAL 2017. *Il mercato del lavoro: verso una lettura integrata*. Roma: Istituto nazionale di statistica.

LOVAGLIO P.G. 2022. Do job vacancies variations anticipate employment variations by sector? Some preliminary evidence from Italy, *Labour*, Vol. 36, No. 1, pp. 71-93.

LUCARELLI A. 2010. Job vacancies and hirings: first evidence coming from enterprise-level analysis, *SIS*, Vol. 1, pp. 159-164.

Pasquale DI PADOVA, Istat, pasquale.dipadova@istat.it
Annalisa LUCARELLI, Istat, anlucare@istat.it
Emilia MATERA, Istat, ematera@istat.it

INEQUALITIES IN ITALIAN COMPETITIVE RESEARCH FUNDING: EVIDENCE FROM THE PRINWINNERS DATASET[1]

Venera Tomaselli, Andrea Orazio Spinello, Giulio Giacomo Cantone

Abstract. The adoption of competitive grants for allocating public research funds is based on the assumption that competition enhances the efficiency of public expenditure, promoting higher quality in research processes and outcomes. However, the literature on the topic has pointed out some drawbacks of the competitive model. This analysis focuses on the indirect effect of competitive funding in enforcing inequalities among academic institutions and disciplinary macro-sectors in Italy. The study presents the PRINWINNERS dataset, encompassing approximately 6,500 research projects funded to over 18,500 recipients across 4 rounds of the PRIN programme from 2017 to 2022, including the extraordinary NRRP-labelled round in 2022. The general PRIN regulation sets an equitable distribution of funds across Life Sciences, Physical Sciences and Engineering, and Social Sciences and Humanities, but such equity is not always reflected within these areas. Using the proportion of tenured professors as benchmark, some disciplinary macro-sectors have benefited more from competition than others within the same area. A considerable amount of variability is observed in the median funding per recipient across academic institutions. Data analysis also shows that the universities specialised in bio-medical or technological research receive high median funding whereas those specialised in social sciences receive less, with Bocconi University being a noteworthy outlier. However, disciplinary specialisation alone does not predict the allocations well enough, indicating the need to investigate additional factors to better understand the dynamics shaping competitive outcomes.

1. Introduction

Over the last decades, the progressive introduction of competitive mechanisms in research funding has led to the proliferation of a growing number of national R&D programmes supporting a selected number of research projects based on a set of merit criteria such as scientific quality, originality and potential impact. The expansion of this funding model assumes that competition for financial resources fosters a more efficient use of public funds while improving the quality of research through processes of selection designed to identify promising projects and successful research teams (Lepori *et al.*, 2007).

[1] Data curation and resources: Andrea Orazio Spinello; Formal analysis: Giulio Giacomo Cantone; Supervision: Venera Tomaselli. All three authors contributed equally at conceptualisation, methodology and writing.

However, the adoption of competitive models for allocating research funding has raised concerns about the ambiguities and challenges associated with a fair distribution of research resources, as well as the systemic effects that such mechanisms may generate, including the penalisation of institutions with fewer resources when competing with excellent institutions, the adoption of opportunistic strategies to increase the chances of success in funding rounds, the reinforcement of cumulative advantage ("Matthew effect”), and a limited capacity to foster substantial disciplinary innovation (Laudel, 2006; van den Besselaar *et al*., 2017; Wang *et al*., 2018; Reale & Zinilli, 2020).

Some studies have addressed inequalities in the allocation of competitive research funding that can arise at both the disciplinary and institutional levels when some fields of study or institutions receive systematically higher shares of funding. For instance, a recent large-scale bibliometric analysis found that Life and Earth sciences hold a significant advantage in attracting grant support, with a high proportion of articles in these fields acknowledging funding over a ten-year period in the Web of Science database (Tian *et al.*, 2024). At the institutional level, studies have pointed to the over-representation of élite universities in successful funding bids, for example in the case of the 24 Russell Group universities in the United Kingdom (Liyanage *et al.*, 2024). This contribution investigates similar potential inequalities in the Italian context, with a particular focus on the PRIN programme.

2. Background: the PRIN programme in Italy

Historically, Italy has been a country characterized by a research funding distribution model based mainly on core (institutional) funding. Project-based funding in Italy has been marked by a limited number of instruments and a centralised administration at the ministerial level (Spinello *et al*., 2023) and, according to the latest data available before the introduction of NRRP-related initiatives, only 10% of the country's research funding has been allocated through competitive mechanisms (Reale, 2018).

The main national competitive R&D instrument is the PRIN (*Progetti di Rilevante Interesse Nazionale*) established in the late 1990s, and it has always been managed directly by the Ministry of University and Research (MUR). Its historical aim has been to support collaborative projects of high scientific quality across Italian public research institutions, encouraging the creation of research networks. As a *curiosity-driven* programme, it has offered researchers the possibility of freely choosing topics and methods for their proposals, allowing greater freedom of exploration of innovative (inter)disciplinary topics compared to other instruments of research funding. PRIN projects follow the labelling system of the European Research Council (ERC) instead of the Italian disciplinary sectors, and the applicants are not formally bound to propose projects in the disciplinary areas of scientific qualifications[2] they are affiliated with (for

[2] Established by MUR Ministerial Decree 855/30-10-2015 and by related previous decrees.

instance the subject they teach at university). It follows that e.g. a Professor of Literature or Biology can still be recipient for a project labelled with the ERC "The Human Mind and Its Complexity", which roughly corresponds to Psychology.

A historical limitation of the PRIN programme is that funds are allocated without scheduled calls or parameterised budgets. It allowed MUR to arbitrarily skip the funding of PRIN for one or more years: until 2010, PRIN rounds were published annually but then their frequency became biennial or triennial, often with a highly variable financial budget. The 2010 round allocated approximately 170 million euros, then only 32.3 million in 2012, and 91.9 million in 2015.

The 2017 round marked a significant increase in available funds, reaching 391 million euros. The programme introduced two important innovations, recognising potential unfairness in the outcomes of previous rounds: the inclusion of a preferential line reserved for projects led by young researchers under 40 years old, with the aim of encouraging generational turnover; and a clause to allocate a percentage of the budget of the round to projects where all the recipients are affiliated with research institutions in Southern Italy. Since the 2017 round an increasing number of scholars from non-academic institutions have had access to funding. The year 2022 saw a relevant surge in the budget thanks to the implementation of the fourth mission of National Recovery and Resilience Plan (NRRP) "Education and Research": MUR has launched two rounds for proposals of the PRIN programme "2022" (742 million euros) and "2022 NRRP" (420 million euros). The NRRP-labelled round was aimed to align national research with the strategic priorities of European research. In fact, the programme has introduced a specific requirement to include in the proposals emerging strategic themes in line with the objectives of the European Framework Programme for Research and Innovation 2021-2027. Additionally, with the NRRP-labelled round the MUR adopted a guideline requiring that 60% of the budget be allocated to Southern Italy.

3. Analysis of the PRINWINNERS dataset

PRINWINNERS is a dataset, managed by CNR-IRCrES, that provides a collection of three 'regular' funding rounds of the PRIN programme from 2017 to 2022, plus the extraordinary NRRP-labelled round of 2022 (hereafter referred to as NRRP). The unit in the database is the funding recipient, who can be the national-level principal investigator, or the researcher responsible of a local unit of it. Information about the project, e.g. title, ERC label of the project, etc., is linked to specific information about the recipients, e.g. the funding they received, their scientific affiliations, etc. A recipient may appear in multiple rounds. Academic recipients constitute the majority of all recipients, and most of the funding is concentrated in the 2022 round (Table 1)[3].

[3] Data sources for PRINWINNERS are the Directorial Decrees of the MUR for funding admission and reallocations: a corpus of around 100 documents which have been digitally organized for the

Table 1 – *Summary of the PRINWINNERS dataset.*

	Projects	Recipients	Academic Recipients	€
2017	664	2,763	2,536	381,226,706
2020	308	1,259	1,153	178,649,593
2022	3,754	10,110	9,003	741,925,632
NRRP	1,780	4,611	4,105	419,973,843
Total	6,506	18,743	16,797	1,721,775,774

3.1. Inequalities among scientific disciplines

The ERC macro-labels correspond vaguely with traditional academic disciplines and are grouped under three areas: Life Sciences (LS), Physical Sciences and Engineering (PE), and Social Sciences and Humanities (SH). Each macro-label is associated with more specific labels (see Tables 2.a, 2.b, 2.c).

The allocation of funds follows the guidelines highlighted in the PRIN calls to minimise inequalities among these three areas of the ERC classification[4]: when summing all funding in PRINWINNERS, LS and PE receive the same allocation of around 607M € (LS: 2,241 funded projects, PE: 2,292). SH received around 16% less funding (507M € in 5,907 projects). Within these areas, the proportion of funds has been stable over time except for the NRRP round, which was more focused on applied science (also, some labels did not exist before 2022, e.g. Environmental Engineering).

The allocation has not been uniform among the micro-labels. Life Sciences are dominated by applications, Diagnostic tools, Physiology, Biotechnology, and Public Health; PE includes a larger number of labels, with only the dominance of Products and Processes Engineering. Among the SH there is more uniformity with remarkable allocation of funds for the study of Cultures and Cultural Production, which received around 20% of all funding. It is not straightforward to determine if these proportions among ERC labels are justified, given that at macro level these are fixed instead. It is noteworthy that inequalities may be driven by the participation of academic recipients on projects that do not correspond to their discipline.

In Table 3 the following benchmark:

$$\frac{\text{n of tenured professors in the discipline}}{\text{N of tenured professors}}$$

constitution of this dataset. Information about recipients have been retrieved by matching data on academics with the “Cerca Università” (“Search for University”) web portal managed by MUR. The dataset was finalized in the early months of 2024; therefore, it does not account for the reallocations of PRIN 2022 round, issued in September 2024.

[4] By Art. 3 of MUR Directorial Decree 3728/27-12-2017; and by Art 4. of MUR Directorial Decrees 1628/16-10-2020, and 104/2-2-2022, and 1409/14-09-2022. These can be accessed at https://prin.mur.gov.it/Iniziative .

is checked against the proportions of allocation of funds.

Table 2.a – *Relative allocations of funds to projects in Life Sciences.*

ERC Label: Life Sciences (LS)	2017	2020	2022	NRRP
Applied Life Sciences and Biotechnology	.157	.172	.158	.163
Cellular and Developmental Biology	.07	.078	.072	.067
Diagnostic Tools, Therapies and Public Health	.168	.185	.171	.173
Ecology, Evolution and Environmental Biology	.065	.082	.088	.075
Genetics, Genomics, Bioinformatics and Systems Biology	.082	.08	.085	.086
Immunity and Infection	.093	.086	.085	.076
Molecular and Structural Biology and Biochemistry	.065	.073	.083	.102
Neurosciences and Neural Disorders	.126	.107	.116	.114
Physiology, Pathophysiology and Endocrinology	.175	.137	.141	.144

Table 2.b *Relative allocations of funds to projects in Physical Sciences and Engineering.*

ERC Label: Physical Sciences and Engineering (PE)	2017	2020	2022	NRRP
Computer Science and Informatics	.092	.098	.088	.1
Condensed Matter Physics	.08	.081	.078	.072
Earth System Science	.089	.095	.094	.1
Environmental Engineering	0	0	.054	.068
Fundamental Constituents of Matter	.099	.093	.089	.036
Mathematics	.072	.06	.069	.059
Physical and Analytical Chemical Sciences	.07	.067	.075	.081
Products and Processes Engineering	.221	.208	.185	.226
Synthetic Chemistry and Materials	.089	.1	.102	.128
Systems and Communication Engineering	.127	.141	.113	.119
Universe Sciences	.06	.059	.053	.01

Table 2.c – *Relative allocations of funds to projects in Social Sciences and Humanities.*

ERC Label: Social Sciences and Humanities (SH)	2017	2020	2022	NRRP
Cultures and Cultural Production	.221	.209	.201	.175
Human Mobility, Environment, and Space	0	0	.078	.116
Individuals, Markets and Organisations	.147	.181	.173	.187
Institutions, Values, Beliefs and Behaviour	.215	.199	.145	.141
The Human Mind and Its Complexity	.148	.159	.164	.132
The Social World, Diversity, Population	.118	.125	.101	.146
The Study of the Human Past	.151	.126	.138	.101

Taking Humanities as an example, Area 10 (Languages) and Area 11 (other Humanities) have been underfunded compared to this benchmark. A possible explanation for this is that research in Humanities is carried out by multidisciplinary research groups (Cantone, 2024). This is rather evident noticing the dominance of the most funded Area 5 (Biology), also the most outperforming the benchmark. By contrast, Area 12 (Law) receives a low share of PRIN funds, with exception of International Law. While there are also differences in the proportion of allocated funds between the ordinary rounds and the NRRP-labelled one, none is particularly sharp. Table 3 covers allocations within areas, as well. Some cases are noteworthy. For example, while in Italy there are more tenured professors in Business (42% of Area 13), 55% of the funds went to recipients affiliated with Economics.

Table 3 – *Relative allocations of funds among the Italian academic population.*

Area	Bench. (Ar.)	PRIN	NRRP	Macro Disciplinary Group	Bench.	PRIN	NRRP
01	**.058**	**.037**	**.033**	Mathematics	.713	.622	.562
01				Computer Science	.287	.378	.438
02	**.044**	**.06**	**.042**	Fundamental Interactions	.38	.312	.152
02				Physics of Matter	.37	.441	.515
02				Astrophysics	.107	.153	.078
02				Applied Physics	.143	.095	.255
03	**.052**	**.062**	**.079**	Analytical Chemistry	.258	.279	.232
03				Inorganic Chemistry	.256	.25	.31
03				Organic Chemistry	.229	.244	.252
03				Pharmaceutic Chemistry	.257	.227	.206
04	**.019**	**.015**	**.016**	Geosciences	1	1	1
05	**.08**	**.139**	**.137**	Plant Biology	.105	.074	.131
05				Animal Biology	.097	.07	.089
05				Ecology	.044	.037	.032
05				Physiology	.116	.141	.124
05				Biochemistry	.271	.242	.197
05				Applied Biology	.072	.085	.102
05				Pharmacy	.122	.194	.147
05				Human Anatomy	.102	.062	.082
05				Genetics	.071	.095	.095
06	**.144**	**.119**	**.11**	Pathology Diagnostics	.151	.311	.298
06				Internal Medicine	.065	.065	.054
06				General Surgery	.062	.015	.014
06				Specialist Medicine	.24	.333	.349
06				Specialist Surgery	.08	.018	.029
06				Integrated Surgery	.12	.037	.051
06				Pediatrics	.04	.045	.026
06				Gynecology	.028	.016	.005
06				Radiology	.042	.016	.016
06				Anesthesiology	.022	.013	.004
06				Public Health	.092	.049	.051
06				Health Professions	.059	.082	.103
07	**.055**	**.052**	**.052**	Agricultural Economics	.107	.073	.082
07				Agriculture and Forestry	.146	.187	.118
07				Biosystems Engineering	.072	.098	.085
07				Plant Pathology	.078	.119	.135
07				Agro-Chemistry	.085	.113	.146
07				Food Technology	.068	.075	.058
07				Animal Technology	.114	.118	.116
07				Veterinary	.279	.146	.204

Table 3 (cont.) – *Relative allocations of funds among the Italian academic population.*

Area	Bench. (Ar.)	PRIN	NRRP	Macro Disciplinary Group	Bench.	PRIN	NRRP
07				Agricultural Microbiology	.053	.071	.055
08	**.066**	**.042**	**.053**	Infrastructures and Territory	.246	.311	.235
08				Geotechnical Engineering	.226	.383	.379
08				Architectural Technology	.161	.085	.085
08				Architectural Design	.127	.052	.07
08				Restoration	.157	.095	.106
08				Urban Planning	.082	.074	.125
09	**.107**	**.106**	**.116**	Mechanical Engineering	.158	.1	.161
09				Industrial Engineering	.104	.07	.091
09				Energetic Engineering	.12	.122	.104
09				Materials Engineering	.104	.132	.162
09				Electrical Engineering	.179	.168	.167
09				Telecommunications	.091	.101	.072
09				Bioengineering	.102	.148	.115
09				Software Engineering	.142	.159	.129
10	**.086**	**.076**	**.051**	Archaeology	.084	.12	.171
10				Art History	.082	.061	.069
10				Media Studies	.082	.141	.128
10				Studies of Antiquity	.117	.093	.129
10				Middle-age Literatures	.038	.047	.037
10				Italian Studies	.146	.186	.124
10				Linguistics	.064	.115	.114
10				French Studies	.049	.017	.025
10				Spanish Studies	.058	.033	.027
10				English Studies	.109	.042	.026
10				Germanic & Slavic Studies	.086	.036	.056
10				Oriental Studies	.085	.11	.095
11	**.08**	**.068**	**.071**	History	.249	.294	.159
11				Geography	.073	.068	.057
11				Philosophy	.22	.165	.217
11				Pedagogy	.158	.063	.072
11				Psychology	.299	.411	.495
12	**.084**	**.035**	**.031**	Private Law	.137	.076	.093
12				Commercial Law	.143	.115	.111
12				Public Law	.125	.117	.208
12				Administrative Law	.132	.124	.074
12				International Law	.199	.29	.28
12				Civil Procedures	.042	.012	.012
12				Criminal Law	.088	.054	.108
12				History of Law	.134	.21	.115
13	**.094**	**.053**	**.065**	Economics	.315	.56	.554
13				Business	.425	.182	.184
13				Economic History	.035	.053	.036
13				Statistics	.225	.205	.225
14	**.031**	**.027**	**.03**	Political Theory	.205	.268	.267
14				Political History	.205	.114	.018
14				Sociology & Applied Soc.	.59	.618	.715

3.1. Inequalities among research institutions

In the PRINWINNERS dataset there are 129 research institutions that received funding in at least one round of the PRIN programme. In Table 4, key indicators are reported for the 74 institutions that have at least 30 recipients across the four rounds of funding.

Among these, the number of recipients is indicative of the size of the institutions in terms of tenured professors, but also of the capacity to attract competitive funds. For example, even if the largest university in Italy is "Sapienza" in Rome, with 1,818 tenured professors in 2022, "Federico II" University of Naples (1,423 tenures) gained the same amount of funds from the ordinary PRIN programme rounds and more than twice as much the extraordinary NRRP-labelled round, with a median allocation per unit that is even higher.

Indeed, variability in the median allocation of euros for unit calls for closer examinations. A first attempt is to correlate the allocation with the specialisation in the ERC macro-labels of the funded projects, that is the percentage of projects under these labels (Table 4). A correlation analysis finds weak positive correlations of the median allocation with the specialisation in PE (Kendall = 0.04), and with LS (Kendall = 0.16). A stronger negative correlation is with SH (Kendall = -0.26). This latter correlation is expected given the allocation among the three macro-labels (see Section 3.1), and yet it does not explain how the university with the highest median allocation per unit – "Bocconi" University in Milan – is specialised in SH. Indeed, all the universities located in Milan receive more funding per unit, not only compared to the universities located in the rest of Italy, but also compared to those in the rest of the Lombardy region. In fact, striking is the difference in median allocation per unit between Bocconi and the University of Bergamo or between "Humanitas" and "Insubria" in Table 4.

The NRRP round fulfilled its mission to balance the territorial inequalities between Southern areas and the rest of the country (46% of allocations vs. previous 27%, see Table 5), also considering that for the three large universities "Federico II" in Naples, Palermo. and Catania, the ordinary round of 2022 underperformed compared to previous rounds. By contrast, Bari has registered a stable allocation and very good performance with NRRP funds. Small Southern universities overall benefited from the two 2022 rounds.

Table 4 – *Recipients, allocation per unit, and specialisation in research institutions*.

Institution	PRIN	NRRP	Med. €	LS	PE	SH
N. Inst. Astrophysics	65	6	116432	0	1	0
"Bocconi" Uni., Milan	65	14	103116	.01	.08	.91
"Humanitas" Uni., Milan	44	9	102653	.98	.01	.01
"Gran Sasso" Inst., Aquila	33	10	98033	.01	.8	.19
"Anton Dohrn" Inst., Naples	23	14	95768	.96	.04	0
Milan Uni.	621	132	95000	.57	.18	.24

N. Inst. Nuclear Physics.	123	22	93369	.04	.96	0
“Campus Bio-Medico”, Rome	32	11	92855	.5	.4	.1
“Sant'Anna” Inst., Pisa	67	21	91167	.21	.29	.5
“Federico II” Uni., Naples	715	346	91036	.45	.37	.18
“A. Avogadro” Uni., Vercelli	86	31	90666	.56	.19	.25
“MagnaGraecia” Uni., Catanzaro	71	47	90000	.8	.1	.1
Trento Uni.	248	63	89893	.19	.38	.43
“Cattolica” Uni., Milan	234	54	89074	.47	.07	.46
“SISSA” Institute, Trieste	57	4	88613	.1	.71	.18
Milan Polytech.	338	78	88043	.06	.73	.2
“Bicocca” Uni., Milan	285	73	87937	.29	.36	.35
Padua Uni.	630	152	87327	.41	.32	.27
Nat. Research Council (CNR)	1140	432	87063	.41	.49	.09
Perugia Uni.	211	53	87051	.38	.38	.24
Salerno Uni.	203	121	86100	.28	.49	.24
“Sapienza” Uni., Rome	706	153	85500	.37	.29	.34
N. Inst. Geophysics	30	13	85499	0	.97	.03
Tuscia Uni., Viterbo	86	26	84587	.57	.13	.3
“San Raffaele” Institute, Milan	54	17	84281	.75	.04	.22
Camerino, Uni.	62	21	83987	.53	.37	.09
Verona Uni.	190	52	83945	.41	.12	.46
“Alma Mater” Uni., Bologna	711	161	83634	.28	.3	.42
“Normale” Uni., Pisa	56	6	83601	.1	.47	.43
Pavia Uni.	268	68	83358	.37	.34	.29
Ferrara Uni.	142	43	83077	.48	.3	.22
“Tor Vergata” Uni., Rome	276	67	83000	.42	.36	.21
Modena Reggio Uni.	190	45	82384	.33	.42	.24
Florence Uni.	425	111	81993	.36	.3	.34
“Aldo Moro” Uni, Bari	238	152	81620	.44	.27	.29
Pisa Uni.	377	80	81510	.33	.41	.27
Parma Uni.	194	51	81335	.4	.35	.26
“d'Annunzio” U., Chieti Pescara	105	62	80790	.38	.14	.48
Teramo Uni.	34	21	80687	.48	.11	.41
Brescia Uni.	123	34	80000	.43	.41	.17
Salento Uni., Lecce	120	76	79951	.17	.35	.47
Turin Uni.	504	119	79438	.44	.18	.37
Siena Uni.	157	46	79343	.42	.15	.44
“L. Vanvitelli” Uni., Caserta	193	119	78557	.57	.25	.18
“L'Orientale” Uni., Naples	40	11	78500	0	.02	.98
Cagliari Uni.	166	85	78477	.27	.43	.3
“Carlo Bo” Uni., Urbino	72	22	78456	.3	.22	.47
Marche Polytech.	114	21	78407	.55	.27	.18
Udine Uni.	148	26	78181	.32	.23	.45
“Roma Tre” Uni., Rome	203	45	78085	.13	.33	.55
Catania Uni.	243	111	78050	.37	.33	.3
Torino Polytech.	217	64	78047	.08	.78	.14
Calabria Uni., Cosenza	147	91	77976	.24	.51	.25
Bari Polytech.	62	51	77864	.01	.85	.14
Genova Uni.	284	69	77000	.27	.43	.3

Table 4 (cont.) – *Recipients, allocation per unit, and specialisation in research institutions.*

Institution	PRIN	NRRP	Med. €	LS	PE	SH
Sassari Uni.	79	44	76819	.41	.23	.36
“Parthenope” Uni, Naples	69	41	76520	.13	.59	.28
Palermo Uni.	256	117	75956	.28	.36	.36
“Kore” Uni., Enna	21	9	75683	.2	.24	.57
Foggia Uni.	62	35	75099	.42	.06	.52
Trieste Uni.	153	34	75001	.35	.41	.24
Aquila Uni.	104	56	74964	.34	.51	.15
Messina Uni.	173	105	74700	.39	.34	.27
“Ca' Foscari”, Venice	132	38	74258	.04	.19	.77
Sannio Uni., Benevento	46	31	73430	.24	.66	.09
Insubria Uni., Varese Como	72	12	72661	.57	.24	.19
Molise Uni., Campobasso	51	27	72075	.53	.15	.32
“LUISS G. Carli” Uni., Rome	40	6	70732	0	.04	.96
Bolzano/Bozen Uni.	56	9	69498	.26	.28	.46
Basilicata Uni., Potenza	46	38	68740	.46	.3	.24
“Mediterranea” Uni., R. Calabria	48	28	68167	.26	.4	.33
Cassino Uni.	44	14	67811	.02	.63	.35
Bergamo Uni.	69	19	66467	0	.17	.83
Macerata Uni.	44	10	58047	.01	.03	.96

Table 5 – *% of received € over top 74 institutions in Southern Universities.*

Universities	Region	2017	2020	2022	NRRP
Federico II, Naples	Campania	.07	.056	.057	.089
Bari	Apulia	.02	.017	.018	.037
Palermo	Sicily	.023	.013	.019	.026
Catania	Sicily	.021	.023	.015	.026
ALL of Calabria	Calabria	.026	.014	.019	.037
ALL of Sardinia	Sardinia	.021	.011	.018	.032
ALL of Abruzzo	Abruzzo	.015	.016	.018	.036
Salerno	Campania	.017	.021	.014	.031
Vanvitelli, Caserta	Campania	.014	.023	.014	.028
Messina	Sicily	.013	.006	.012	.022
Salento, Lecce	Apulia	.007	.009	.009	.019
Others in South Italy		.028	.032	.029	.053
Others not in South Italy		.724	.756	.759	.561

4. Conclusions and future developments

Results of the study are consistent with international findings presented in the Section 1: Life Sciences and élite institutions perform well in the Italian competitive model of the PRIN programme. A key insight may lie in the inconsistency of the SH label in representing a unitary area of research. There are signals that, while competitive funding naturally cannot favor all institutions, the competition within LS and PE is healthy with many well-performing actors, but this is not the case for SH, which appears to be dominated by Bocconi, indeed. Future studies could track the state of competition with advanced measures of concentration (Cantone, 2024). Such studies would be also useful to help policy-makers account for this difference between macro-areas in the design of the next generation of funding programmes.

In general, the low correlations found in this study justify the adoption of complex models to explain the inequalities in the allocation of funds. There is also a correlation between the number of funded projects (PRIN + NRRP in Table 4) and the median allocation (Kendall = 0.13). Evidence does not support attributing this to a lack of large institutions specialised in SH because the Kendall correlation between specialisation in SH and the number of funded projects is null instead. Hypothetically, while large and prestigious institutions may simply employ a greater number of highly qualified researchers, it is also possible that potential recipients affiliated with these institutions have access to better opportunities for negotiating richer projects, as the case of Milan discussed in Section 3.2 appears to illustrate.

Furthermore, the PRINWINNERS dataset allows for an assessment of whether the requirement that restricts access to reserved funding to projects presented by researchers all affiliated with Southern universities should be reconsidered. Indeed, the clause mandating that a share of the PRIN budget be allocated to projects in which all recipients are affiliated with research institutions in Southern Italy may restrict opportunities for researchers in Southern universities willing to collaborate with qualified colleagues from institutions outside the South. This could both generate unexpected consequences on scientific quality and affect the creation of networks. Current evidence shows that, in the long run, it could be appropriate to move beyond purely geographical criteria and focus on identifying the specific structural issues faced by disadvantaged universities, irrespective of the territorial area where they are located.

References

CANTONE G.G. 2024. How to measure interdisciplinary research? A systemic design for the model of measurement. *Scientometrics*, Vol. 129, No. 8, pp. 4937–4982.

LAUDEL G. 2006. The art of getting funded: how scientists adapt to their funding conditions, *Science and Public Policy*, Vol. 33, No. 7, pp. 489–504.

LEPORI B., VAN DEN BESSELAAR P., DINGES M., VAN DER MEULEN B., POTÌ B., REALE E., SLIPERSAETER S., THEVES J. 2007. Indicators for comparative analysis of public project funding: concepts, implementation and evaluation, *Research Evaluation*, 16, Vol. 4, pp. 243–255.

LIYANAGE C.L., VILLALBA-ROMERO F., CARMICHAEL A. 2024. Fairness in Higher Education Research and Innovation Funding in the UK, *Trends in Higher Education*, Vol. 3, No. 4, pp. 1031–1052.

REALE E., 2018. Il finanziamento pubblico per Ricerca e Sviluppo, Le politiche dell'Italia in prospettiva comparata, In: CONSIGLIO NAZIONALE DELLE

RICERCHE (Ed.): *Relazione sulla ricerca e l'innovazione in Italia, Analisi e dati di politica della scienza e della tecnologia (quarta edizione)*, Rome: CNR Edizioni, pp. 225–253.

REALE E., ZINILLI A. 2020. Ex-ante evaluation of research project funding: what is the importance of scientific quality? *Welfare e Ergonomia*, Vol. 2019, No. 1, pp. 29-41.

SPINELLO A.O., VARINETTI, E., ZINILLI A., REALE E. 2023. Il finanziamento competitivo della Ricerca e Sviluppo in Italia per le sfide sociali e tecnologiche. In: CONSIGLIO NAZIONALE DELLE RICERCHE (Ed.): *Relazione sulla ricerca e l'innovazione in Italia, Analisi e dati di politica della scienza e della tecnologia (quarta edizione)*, Rome: CNR Edizioni, pp. 23–58.

TIAN W., CAI R., FANG Z., XIE Q., HU Z., WANG X. 2024. Research funding in different SCI disciplines: A comparison analysis based on Web of Science, *Quantitative Science Studies*, Vol. 5, No. 3, pp. 757–777.

VAN DEN BESSELAAR P., HEYMAN U., SANDSTROM U. 2017. Perverse effects of output-based research funding? Butler's Australian case revisited, *Journal of Informetrics*, Vol. 11, No. 3, pp. 905–918.

WANG J., LEE Y., WALSH J. P. 2018. Funding model and creativity in science: Competitive versus block funding and status contingency effects, *Research Policy*, Vol. 47, No. 6, pp.1070–1083.

Venera TOMASELLI, University of Catania venera.tomaselli@unict.it
Andrea Orazio SPINELLO, CNR-IRCrES, andreaorazio.spinello@cnr.it
Giulio Giacomo CANTONE, Magna Græcia University of Catanzaro, g.cantone@unicz.it

EMOTIONS AND THE MEANING OF ASSESSMENT IN SCHOOL SETTINGS

Nidia Batic

Abstract. Emotions shape our relationship with reality and influence decisions across social, relational, and professional domains (Goleman, 1995). In the context of schooling, emotions specifically linked to students' academic experiences are referred to as "academic emotions" (Pekrun, 2006), and they may be either positive or negative. Positive emotions can foster learning (Poggi *et al*., 2004), and those experienced during assessment are closely related to students' self-confidence and self-efficacy (Tomasi, 2023). Emotional responses significantly affect assessment outcomes (Camacho-Morales *et al*., 2021), reinforcing the idea that emotion and cognition are deeply intertwined (Perla, 2002). In primary education, assessment serves primarily a formative function: it enables teachers to understand students' learning processes and to guide their academic development (Benvenuto in Nigris and Agrusti, 2021). Consequently, recognizing children's emotional states during assessment is essential. This exploratory study, based on a non-probabilistic sample, investigates primary school students' perceptions regarding the purpose and meaning of assessment, as well as their emotional responses during evaluative tasks. Results show that students assign considerable importance to assessment and clearly understand its primary purposes: to measure their learning, track their progress, and identify mistakes. Fewer students view assessment as a means of demonstrating competence to the teacher or achieving high report card grades. Emotional and behavioral reactions—both in written and oral assessments—are strongly correlated with the level of preparation students report. When they are aware of being unprepared, they frequently experience fear or anxiety, although they still attempt to remain focused during the test. Nonetheless, physiological signs along with the use of distracting or copensative strategies indicate emotional tension. The findings also highlight a generally healthy level of self-esteem: students demonstrate an awareness that their self-worth is not solely determined by assessment outcomes. Moreover, they recognize a link between consistent academic effort and positive test results.

1. Introduction

Emotions are complex processes that arise in response to internal or external stimuli—referred to as emotion-eliciting events—through physiological changes, thoughts, and behaviors (Lewis *et al.*, 2008). They shape how we relate to reality and influence our decision-making processes (Goleman, 1995). As Campana (2017) argues, there is no cognitive act without emotional significance, and emotions are "considered fundamental and indispensable elements for education" (Cagol, 2019, p. 111). Goleman (1995) proposes a model of emotional intelligence in which the recognition and regulation of one's own emotions contribute to improved social, professional, and overall personal well-being. This framework is also relevant in educational contexts, where the emotional dimension is deeply embedded throughout the entire learning process.

The emotions experienced by children throughout their educational journey are referred to as "academic emotions", which can manifest with varying intensity, carry either positive or negative valence, and directly influence the level of well-being or distress perceived in the classroom (Pekrun, 2006). Academic emotions have been shown to impact motivational processes related to learning (Pekrun and Schutz, 2007), as well as cognitive functions such as attention (Vuilleumier, 2005), memory consolidation, and learning acquisition (Phelps, 2004). These emotional processes, in turn, influence students' responses during assessment (Hopwood *et al.*, 2024). Positive emotional states have been shown to stimulate motivation for learning (Poggi *et al.*, 2004). Emotions experienced during assessment situations are significantly correlated with self-confidence, perceived self-efficacy, and mastery goal orientation (Tomasi, 2023), and they can exert a meaningful influence on students' performance in evaluative contexts (Camacho-Morales *et al.*, 2021). Among the key factors contributing to the emergence of emotions within educational settings are the significance attributed to assessment and the fear of failure (Zeidner, 2014). Assessment, in turn, plays a critical role in shaping students' behavior and influencing their self-concept (Carless and Lam, 2014, p. 315).

When students approach *assessment-for-learning* tasks with strong emotions such as anxiety or fear, this can lead to situations of discomfort or even tension. Such emotional climates may negatively affect the collective well-being that is essential for fostering a positive and supportive learning environment. The impact of the school climate on learning is mediated by students' well-being (Fatou and Kubiszewski, 2018). As noted by Hossain *et al.* (2023, p. 448), "students with a higher sense of well-being perform better at school." Furthermore, empirical evidence suggests that the experience of well-being influences students' engagement in school activities beyond the effects attributable to their perception of the school climate (Lombardi *et al.*, 2019).

The literature is therefore in agreement in assigning emotions a very important role in children's school life; however empirical studies contextualized within primary school that address students' conceptions of assessment are still quite rare (Imperio and Seitz[1], 2023). The present research aims to highlight the emotional

[1] As an example, Imperio and Seitz (2023, p. 4) cite the following works: HARRIS L. R., BROWN G. T. L., HARNETT J. A. 2009. "Drawing" Out Student Conceptions: Using Pupils' Pictures to Examine

experiences of primary school students themselves, who are conceptualized as "competent actors," namely individuals recognized as "experts and principal informants of their own lived experiences" (Imperio and Seitz, 2023, p. 3). When children are regarded as competent witnesses of their educational experiences and can recognize the emotions linked to assessment, they are more likely to take an active role in learning; however, when assessment is viewed merely as accreditation, engagement declines (Remesal, 2009, in Imperio & Seitz, 2023). This understanding enables teachers to better support students in managing these emotions and fostering a positive classroom climate.

2. Objectives, tool and method

This research project is framed within an action research paradigm (Lewin, 1946) with a non-inferential aim. In this context, research aims to identify and understand a problem (Research) and then gather information to address it (Action), specifically focusing on strategies to foster a calm climate during assessment in order to enhance well-being in primary school classrooms (Batic, 2025).

The study explored various aspects of assessment; however, this paper focuses on the meaning assessment holds for children and its emotional dimension, aiming to test two hypotheses: 1) children perceive assessment as important or useful; 2) children experience varying levels of worry and anxiety before and during tests, depending on whether they have studied. These emotional states influence their performance (as already demonstrated by Hopwood et al., 2024), and assessment outcomes may, in turn, affect their self-esteem. To explore these aspects, a questionnaire was developed and refined through two preliminary cognitive pre-tests, aimed at evaluating item clarity, simplicity, and administration time. The final instrument consisted of 25 items, including structured questions, Likert-type scales from 0 to 5, and two open-ended questions. Emotions related to assessment were measured through both explicit self-reports (e.g., "I feel calm," "I feel anxious," "I feel scared") and indirect indicators of emotional states. The questionnaire included a list of behaviors, and children were asked to indicate those they exhibit when they have studied and when they have not. Kinesics was used to interpret nonverbal cues

---

Their Conceptions of Assessment. In MCINERNEY D. M., BROWN G. T.L., LIEM G.A.D. (Eds.) *Student perspectives on assessment: What students can tell us about assessment for learning*. Charlotte, NC: Information Age Pub, pp. 321-330; MONTEIRO V., MATA L., SANTOS N. N. 2021. Assessment Conceptions and Practices: Perspectives of Primary School Teachers and Students, *Frontiers in Education*, Vol. 6, Article 631185, pp. 1-15; MURPHY C., LUNDY L., EMERSON L., KERR K. 2013. Children's perceptions of primary science assessment in England and Wales, *British Educational Research Journal*, Vol. 39, No. 3, pp. 585–606.

(De Carlo & Perfetti, 2022), while physical symptoms such as stomach aches or crying were considered signs of anxiety (Mazzocco, 2017). The use of proxies for children's emotions and Likert-type scales implies an approximate measurement, potentially leading to over- or underestimation. Nonetheless, research indicates that children aged six and above can reliably report on their own health and emotional experiences (Riley, 2004); therefore, self-reporting was preferred. Another potential limitation concerns social desirability bias (Wang & Zang, 2025). To minimize this risk, the questionnaires were administered anonymously and completed independently by the children, following standardized instructions. A single administration was conducted, during which participants were asked to report the emotions they remembered experiencing when they were aware of having studied or not.

The data collection[2] took place between March and April 2025 and involved 29 primary school classes across the provinces of Gorizia, Pordenone, Trieste, and Udine. A non-probability sample of 472 children aged between 8 and 11 years was interviewed. The distribution across grade levels was as follows: 25.4% in third grade, 26.1% in fourth grade, and 48.5% in fifth grade. The sample consisted of 47% boys and 53% girls.

3. Data analysis and hypothesis testing

3.1. The purpose of assessment

One of the main findings of the study is that children attribute high importance to assessment: in response to a direct question, both the median and the mode were 5, with no significant gender[3] differences. When responses were dichotomized, only 6.8% rated between 0 and 2. Students clearly recognize tests as a means to gauge learning, improvement, and mistakes (Mdn = 5; Mo = 5). They are also aware—albeit to a slightly lesser extent—that assessment serves to determine the grade recorded in the report card or to demonstrate their competence to the teacher (Mdn

[2] The questionnaires were administered by trainee students from the Educational Sciences program at the University of Udine, in the presence of classroom teachers. Prior to administration, the project was approved by school principals, and families provided informed consent. Data were collected anonymously and processed in aggregate form, in full compliance with GDPR (2016/679) and Legislative Decree 101/2018. To test gender differences, the chi-square test was used, applying Yates' correction when necessary.

[3] To examine the statistical significance of gender differences, chi-square tests were performed. When appropriate, Yates' continuity correction was applied to adjust for small sample sizes.

= 4, Mo = 5) (tab. 1). The normalized Shannon index (H')[4] indicates substantial heterogeneity in the responses, except for the first two items and the last one, which show a moderate heterogeneity. The apparent discrepancy between the mean indices and the high H' values can be explained by the fact that, when categories are associated with low frequencies, the Shannon index tends to "reward" their presence, thereby potentially overestimating the heterogeneity of the distribution. Students' awareness of the meaning of assessment reflects both intrinsic motivation—linked to the importance attributed to testing, such as understanding how much they have learned, whether they have improved, and identifying mistakes—and extrinsic motivation, such as obtaining good grades and demonstrating their ability to the teacher. Assessment is not perceived by students as a means of competing with peers, such as receiving public praise, making comparisons with classmates, or determining who is the best or worst in the class. Statistically significant gender differences (p-value = 0.000) were observed in only two items, although the associations were weak. Boys, more than girls, were inclined to view assessment as a way to establish who is the best or worst in the class (Cramér's V = 0.179) and to receive praise in front of others (Cramér's V = 0.213).

Table 1 – *Children's perceived purposes of assessment (Mean[5], Median, Mode and Shannon Index).*

	Mean	Median	Mode	H'
To understand how much they have learned	4.3	5	5	0.622
To assess progress compared to previous evaluations	4.3	5	5	0.641
To identify their mistakes	4.1	5	5	0.719
To obtain a good report card	3.8	4	5	0.809
To identify which students did not understand	3.8	4	5	0.811
To determine the grades on the report card	3.8	4	5	0.811
To demonstrate their competence to the teacher	3.5	4	5	0.899
For the teacher to express a judgment on students	3.1	3	5	0.929
To reward those who perform well	2.2	2	0	0.972
To receive praise in front of the class	1.4	1	0	0.848
To make comparisons with classmates	1.3	1	0	0.830
To establish who is the best (and the worst) in the class	0.7	0	0	0.552

[4] The normalized Shannon index is $H' = -[\sum p_i \ln(p_i)]/\ln(k)$; where k is the total number of categories and p_i is the proportion of category i. It ranges from 0 to 1, reaching 1 when all categories are equally represented (maximum heterogeneity).

[5] The mean of ordinal variables is not strictly appropriate; however, it is reported here and in Table 4 for descriptive purposes.

3.2. Emotional and behavioral aspects related to assessment tasks

Assessment inevitably elicits emotional responses, and children were able to recognize and differentiate these across contexts—before tests, during written and oral exams, and depending on whether they had studied or felt unprepared (Tab. 2). Emotional states before a test differ depending on whether children have studied: calm predominates among the prepared (64.0%), while anxiety (56.8%) and fear (48%) are higher among the unprepared. Even prepared students experience anxiety (41.3%), with 36% not feeling calm and 53.4% not looking forward to the test. Boys tend to be calmer than girls, regardless of preparation. Additionally, among those who feel prepared, boys are more likely than girls to try not to think about the up-

Table 2 – *Children's Emotional States Before an Assessment (Percentage Values)*[6].

	Have studied	Have not studied
They feel calm	64.0 *	15.4 **
They are afraid of forgetting what they studied	52.6 *	38.4
They look forward to the test	46.6	9.4
They experience anxiety	41.3 *	56.8 **
They believe they do not know the material	22.2	38.1
They feel fear	17.2	48.0 **
They try not to think about it	12.6 *	23.1
They have stomach aches	7.2	17.1 **
They avoid going to school	0.6	2.8

coming test (16.0% vs. 9.7%). Girls show greater emotional reactivity, reporting higher anxiety than boys regardless of preparation, and are more likely to fear forgetting studied material (58.7% vs. 45.5%). Girls who do not feel adequately prepared report greater fear of taking the assessment compared to boys (66.0% vs. 46.0%) and exhibit higher levels of somatization, such as experiencing stomach aches (22.9% vs. 10.5%). McNemar's test, applied to all items, yielded statistically significant results ($p < 0.001$).

Children's emotional experiences during assessment tasks are detailed in Table 3, which compares emotional and behavioral reactions in written and oral tests under two conditions: when students had studied and when they had not. Items are presented in descending order based on the written test "studied = Yes" condition. Results show that children are aware of the need to focus, and most manage to do

[6] * Statistically significant differences between males and females were found using the chi-square test with $\alpha = 0.05$ and d.f. = 1, and Cramer's V values were, respectively, 0.096; 0.192; 0.099; 0.094.
** Statistically significant differences between males and females were found using the chi-square test with $\alpha = 0.05$ and d.f. = 1, and Cramer's V values were, respectively, 0.157; 0.201; 0.103; 0.165.

so. Comparisons indicate that the level of preparation, rather than the type of test, strongly influences emotional and behavioral responses. Statistically significant differences were observed across reactions, analyzed using McNemar's test with continuity correction for paired data. Specifically, for each type of assessment—written and oral—the responses given by the same students when they had studied were compared with those provided when they had not. For all comparisons, the null hypothesis was rejected in favor of the alternative hypothesis (H_1[7]), statistically confirming that attitudes, behaviors, and emotional states during both written and oral assessments are significantly influenced by preparation level.

Table 3 –*Self-Reported Emotional and Behavioral Reactions During Written and Oral Tests, Based on Whether Children Stated They Had Studied (Yes) or Not Studied (No) (Percentage Values).*

	Written tests		Oral tests	
	Yes	No	Yes	No
Ability to focus	94.2	80.6	93.5	79.2
Sustained attention on the task	88.7	60.5	90.7	65.2
Feeling calm	84.7	14.2	78.9	19.2
Remaining physically still	69.8	28.7	59.4	29.6
Body movements (e.g., fidgeting)	50.8	72.4	57.1	69.5
Sweaty hands	44.9	62.7	40.1	54.5
Requesting clarification from the teacher	33.1	60.7	15.5	36.3
Playing with hair	26.5	34.1	24.7	29.5
Difficulty recalling studied information	24.6	66.8	23.1	63.3
Manipulating objects on the desk	23.9	35.6	20.7	27.1
Comparing answers with classmates	17.6	25.1	--	--
Asking to use the bathroom	17.3	34.5	14.8	26.8
Trembling	17.0	46.7	19.7	44.8
Nail biting	15.8	26.2	13.7	23.0
Blushing	9.9	25.6	15.5	28.9
Headache	9.8	28.8	11.8	24.7
Drawing or doodling	9.3	19.9	8.2	14.2
Stomach ache	9.2	29.2	10.5	25.1
Copying from classmates	4.6	21.8	_--	--
Crying	2.4	12.6	1.7	9.7
Stammering or stumbling over words	--	--	33.5	60.5

'--' indicates that the item was not applicable or not included for that test type

From a descriptive standpoint, it emerges that, irrespective of their level of preparation, a greater number of children tend to seek guidance and suggestions from the teacher during written assessments than during oral ones. Furthermore, a significant proportion of students exhibit stuttering behaviors during oral evaluations, even when they report feeling adequately prepared. Almost all students

[7] All tests were statistically significant, with p-values < 0.001.

report being focused during the assessment; however, the most significant difference—as expected—is observed in the degree of calmness with which they approach these tests. Specifically, this disparity reaches 70.5% in written assessments (84.7% when prepared versus 14.2% when unprepared) and 59.7% in oral assessments (78.9% when prepared versus 19.2% when unprepared). The habit of playing with one's hair, which varies significantly across different situations, shows the least variation. Physiological indicators of anxiety (such as sweaty palms, trembling, flushing, headaches, stomachaches, crying, and stammering) and kinetic indicators (including movement during the test, playing with objects, and playing with hair) are more prevalent among children who are aware that they have not studied. Similarly, so-called distracting strategies (e.g., asking to go to the bathroom, nail-biting, drawing, and doodling) and surrogate strategies (such as checking with peers or copying from them during the test) are more frequently observed in these children. Regardless of children's emotional state during assessment, it is important to understand their relationship with evaluation and its influence on self-esteem (tab. 4). Students recognize the correlation between study time and performance (Mdn = 5 and Mo = 5), although 9% rated 2 or lower. They are less likely to see studying as aimed at obtaining a reward (Mdn = 2, Mo = 0), with 45.4% rating this statement 3–5. Overall, assessment outcomes appear to have little impact on self-esteem (Mdn = 0, Mo = 0), yet 14.5% report increased self-worth after success, and 16.5% feel negatively affected by poor performance. A gender difference emerges only regarding motivation by rewards: boys (53.1%) more than girls (38.7%) report higher motivation when a reward is expected ($p = 0.000$, Cramér's $V = 0.225$). The normalized Shannon index (H') shows high response heterogeneity, particularly regarding views on effort driven by rewards.

Table 4 – *Opinions On the Effects of Assessment (Mean, Median, Mode and Shannon Index).*

	Mean	Median	Mode	H'
The more time I study, the better the result I achieve in assessment	4.3	5	5	0.613
I put in more effort if I receive a reward afterward	2.1	2	0	0.935
I feel more valuable if I perform well on the assessment	1.4	0	0	0.816
I feel less valuable if I perform poorly on the assessment	1.0	0	0	0.712

4. Conclusions: from research to action

The findings of this study have practical implications. Although it does not pursue inferential objectives, the two initial hypotheses were confirmed. Importantly, the results align with existing literature, indicating that assessment can be considered an activating situation to which students attribute highly subjective interpretations.

These interpretations, in turn, shape the emotional responses that influence subsequent behaviors. One may refer to a triadic model of emotions, encompassing a physiological component (bodily manifestations), a mental focus, and an internal dialogue with oneself. Effective regulation of these components can contribute to enhanced emotional well-being. Given the correlation between emotional states and quality of life, learning to recognize and accept one's emotions can lead to improvements in overall well-being (De Carlo and Perfetti, 2022).

Awareness of being unprepared strongly affects students' pre-test emotions and behaviors, regardless of test type, with unprepared students finding it harder to maintain focus despite generally trying to concentrate. Children tend to experience emotional unrest and express their anxiety through a range of physiological indicators (such as sweating hands, trembling, headaches, stomachaches, stuttering, or even crying) and kinetic behaviors (including restlessness, playing with their hair, or manipulating objects). These manifestations are often accompanied by cognitive difficulties (e.g., forgetting information, seeking guidance from the teacher) and the use of distracting strategies (such as asking to go to the restroom) or surrogate strategies (e.g., copying or checking answers with peers).

All these reactions highlight the need for teachers to Act by fostering in children the awareness that assessment for learning should not be a source of stress, but rather— as some children have already recognized— an important opportunity for self-awareness and personal growth. In particular, most children already understand that assessment provides an opportunity to evaluate how much they have learned and to identify their mistakes. However, this view is not yet universally shared, and thus it may serve as one of the key focal points within the Action Project. Indeed, error plays a significant role in the assessment process, particularly in formative assessment (Vannini, 2019), and should be regarded as an opportunity for improvement (Benvenuto in Nigris and Agrusti, 2021, p. 15). Furthermore, errors make the individual learning process visible, thereby enabling the teacher to offer personalized and targeted instructional support (Hattie, 2012). In this way, teacher feedback is not merely a correction of errors but takes the form of a "dialogic pedagogy" (Manca, 1996), involving mutual understanding of mistakes by both teacher and student, who is guided and encouraged throughout their learning journey. The student should not be taught to avoid errors, but rather to use the information they contain (Barth, 1973, p. 42); thus, errors become a "driving force of the educational process in which students are actively engaged" (Castaldi, 2016, p. 12). In conclusion, error serves as a tool for growth for the child within an antifragile system (Taleb, 2012) and is a source of learning. Similarly, assessment provides an opportunity for continuous feedback that supports students' development and fosters critical and creative thinking.

Contrary to what is indicated in the literature, which suggests that children are motivated to attend school by extrinsic factors (i.e., motivation driven by the promise of a reward upon achieving a goal, Campana, 2017), the present research reveals that children do not assign significant value to statements such as "an assessment can serve to reward those who work well" and "I put in more effort if I receive a reward." Instead, intrinsic motivations appear to prevail (Hattie, 2012), related to interest, self-affirmation, and social validation, which are also nurtured by teacher feedback. At the same time, the utility of assessment is not primarily associated with competitive dynamics or social comparison. There is a widespread confidence in personal abilities and effort, coupled with the awareness that a strong commitment to studying can lead to improved assessment outcomes. Self-esteem tends to be high, as indicated by the absence of a correlation between self-perceived value and assessment results. On the basis of this awareness, the teacher can enhance the Action research process by implementing a structured project aimed at fostering emotional literacy (Maggi and Ricci, 2022), aimed at helping all children to recognize their emotions and to become aware of the role these emotions play in the process of learning and in both cognitive and personal development.

References

BARTH R. S. 1973. *Open Education and the American School*. New York; Agathon Press, 2ª ed.

BATIC N. 2025. Well-being at school, *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXXIX, No. 1, pp. 55-66.

BENVENUTO G. 2021. La valutazione formativa, per una didattica inclusiva. In NIGRIS E. and AGRUSTI G. (Eds) *Valutare per apprendere. La nuova valutazione descrittiva nella scuola primaria*, Milano: Pearson Academy, pp. 7-18.

CAGOL M. 2019. Emozioni e pedagogia. Per un primo inquadramento scientifico, *Studi sulla Formazione*, Vol. 2, No. 22, pp. 111-121.

CAMACHO-MORALES, J., SLEMP G.R., PEKRUN R., LODERER K., HOU H., OADES L.G. 2021. Activity achievement emotions and academic performance: A meta-analysis, *Educational psychology review*, Vol. 33, pp. 1051–1095.

CAMPANA G. 2017. (consultato il 13 febbraio 2025). *Motivazione all'apprendimento: il ruolo delle emozioni,* https://blog.edises.it/wp-content/uploads/2017/07/La-motivazione-allapprendimento_il-ruolo-delle-emozoni. pdf

CARLESS D., LAM R. 2014. The examined life: perspectives of lower primary school students in Hong Kong, *Education 3-13: International Journal of Primary, Elementary and Early Years Education*, Vol. 42, No. 3, pp. 313-329.

CASTALDI A. 2016. L’importanza dell’errore nel processo di apprendimento, *Educare.it*, Vol.16, No.1, pp.10-14.
DE CARLO A., PERFETTI L. 2022. *Gestione e comunicazione delle emozioni*, https://archivio.iispoloamantea.edu.it/attachments/article/2946/Comunicazione_e_gestione_delle_emozioni%20(1).pdf
FATOU N., KUBISZEWSKI V. 2018. Are perceived school climate dimensions predictive of students’ engagement?, *Social Psychology of Education*, Vol. 21, pp. 427–446.
GOLEMAN D. 1995. *Emotional intelligence: Why it can matter more than IQ*. New York: Bantam Books.
HATTIE J. 2012. *Apprendimento visibile, insegnamento efficace* (trad. it. 2016). Trento: Erickson.
HOPWOOD N., PALMER T. A., KOH G. A., LAI M. Y., DONG Y., LOCH, S., YU K. 2024. Understanding student emotions when completing assessment: technological, teacher and student perspectives. *International Journal of Research & Method in Education, pp. 1-16.* https://doi.org/10.1080/1743727X.2024.2358792
HOSSAIN S., O’NEILL S., STRNADOVÁ I. 2023. What Constitutes Student Well-Being: A Scoping Review Of Students’ Perspectives, *Child Indicators Research*, Vol. 16, No. 2, pp. 447-483.
IMPERIO A., SEITZ S. 2023. Positioning of Children in Research on Assessment Practices in Primary School. In SEITZ S., AUER P., BELLACICCO R. (Eds) *International perspectives on inclusive education. In the light of educational justice*, Opladen, Berlin, Toronto: Verlag Barbara Budrich, pp. 47-66.
LEWIN K. 1946. Action Research and Minority Problems, *Journal of Social Iusses*, Vol. 2, pp. 34-46.
LEWIS M., HAVILAND-JONES J. M., FELDMAN BARRETT L. (Eds.) 2008. *Handbook of Emotions* (3rd ed.). New York: The Guilford Press.
LOMBARDI E., TRAFICANTE D., BETTONI R., OFFREDI I., GIORGETTI M., VERNICE M. 2019. The Impact of School Climate on Well-Being Experience and School Engagement: A Study WithHigh-School Students, *Frontiers in Psychology*, online, doi: 10.3389/fpsyg.2019.02482.
MAGGI M., RICCI A. 2022. *L'educazione emozionale. Strategie e strumenti operativi per promuovere lo sviluppo delle competenze emotive a scuola e in famiglia*. Milano: Franco Angeli.
MANCA G. 1996. Le possibilità pedagogiche dell’errore, *Pedagogia e Vita*, Vol. 54, No. 2, pp. 60-83.
MAZZOCCO F. 2017. *Ansia scolastica: sintomatologia, manifestazioni e trattamento*. https://www.stateofmind.it/2017/05/ansia-scolastica/

PEKRUN R. 2006. The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice, *Educational Psychology Review*. Vol. 18. pp. 315-341.

PEKRUN R., SCHUTZ P.A. 2007. Where do we go from here? implications and future directions for inquiry on emotions in education. In SCHUTZ P.A. and PEKRUN R, (Eds) *Emotion in education*. Amsterdam: Elsevier Academic Press, pp. 313–331.

PERLA L. 2002. *Educazione e sentimenti. Interpretazioni e modulazioni*. Brescia: La Scuola.

PHELPS, E. A. 2004. Human emotion and memory: interactions of the amygdala and hippocampal complex, *Current opinion in neurobiology*, Vol. 14, pp. 198-202.

POGGI I., BARTOLUCCI L., VIOLINI S. 2004. Emozioni: un'arma per l'apprendimento. In SUSI F., CIPRIANI R., MEGHNAGI D. (Eds) *Antinomie dell'educazione nel XXI secolo*, ROMA: Armando Editore, pp. 258-266.

REMESAL A. 2009. Accessing Primary Pupils' Conceptions of Daily Classroom Assessment Practices. In MCINERNEY D. M., BROWN G.T.L., LIEM G. A.D. (Eds.) *Students Perspectives on Assessment. What students can tell us about assessment for learning*, Charlotte, NC: Information Age Pub, pp. 25-53.

RILEY A. W. 2004. Evidence That School-Age Children Can Self-Report on Their Health, *Ambulatory Pediatrics*, Vol. 4, No. 4, pp. 371-376.

TALEB N. N. 2012. *Antifragile: Things That Gain from Disorder*. New York: Random House.

TOMASI M. 2022/23. *Il ruolo delle emozioni nell'apprendimento scolastico: relazioni con strategie, motivazione e successo in studenti di scuola secondaria di primo grado*. Undergraduate thesis, Università degli studi di Padova.

VANNINI I. 2019. Valutare per apprendere e progettare. In NIGRIS E., BALCONI B., ZECCA L. (Eds) *Dalla progettazione alla valutazione didattica*, Milano: Pearson.

VUILLEUMIER P. 2005. How Brains Beware: Neural Mechanisms of Emotional Attention, *Trends in Cognitive Sciences*, Vol. 9, pp. 585-594.

WANG, M. D., ZANG, X. 2025. Can we trust children's self-reports? Examining socially desirable responses in elementary school surveys, *International Journal of Educational Methodology*, Vol. 11, No. 3, pp. 349-357.

ZEIDNER M. 2014. Anxiety in education. In PEKRUN R. and LINNENBRINK-GARCIA L. (Eds) *International handbook of emotions in education*. London: Routledge, 265–288.

Nidia BATIC, Dipartimento di Scienze Economiche e Statistiche, Università di Udine, nidia.batic@uniud.it

HOUSING CONDITIONS AMONG ELDERLY PEOPLE USING 2021 ITALIAN PERMANENT POPULATION AND HOUSING CENSUS DATA: URBAN AND RURAL PATTERNS[1]

Angela Chieppa, Simona Mastroluca, Alessandro Sasso

Abstract. Urban and rural perspectives provide a vital framework for analysing ageing in contemporary societies. Urban areas often highlight inequalities arising from demographic diversity, while rural regions are typically perceived as more stable, yet remain under-researched (Marcellini et al., 2007). This study investigates housing conditions among elderly people in Italy - including floor space, density standard, and tenure status - according to household composition.The analysis adopts a multidimensional approach that considers territorial disparities by evaluating the influence of both regional contexts and degree of urbanization.

The aim is to explore whether housing conditions for the elderly differ significantly between urban and rural areas. Urban contexts are expected to show greater inequalities due to demographic and social factors, while rural areas may exhibit more uniform patterns. The analysis relies on data from the Permanent Population and Housing Census (PPHC), which integrates administrative sources and surveys. This rich database allows for a direct analysis of multivariate distributions, such as the joint distribution of housing conditions and type of households. The availability of detailed information over time and at a fine territorial level is particularly valuable for examining regional disparities and demographic trends.

1. Elderly people in Italy

In 2021, the reference year of the latest decennial Population and Housing Census(PPHC), the share of the population of working age (15-64) stands at 63.5 per cent, while individuals aged 65 or over make up 23.8 per cent of the total, up 3 percentage points from ten years earlier. This dynamic is common to all Member States in the EU27, but Italy has the highest share of elderly people.

According to demographic projections drawn up by ISTAT on the basis of data as of 1 January 2023, by 2050 people at least 65 years of age could account for 34.5% of the total population, with the incidence of the over-80s rising from 7.6% in 2023 to 13.6% by 2050. This is what comes out of the latest Istat Annual Report (Istat, 2025), which offers alarming food for thought. The ageing of the population leads to

[1] Sections are attributed as follows: sections 3 and 4 to Angela Chieppa, section 1 to Simona Mastroluca, section 2 to Alessandro Sasso.

significant challenges for society as a whole, which must address complex issues including those related to housing conditions and, consequently, the quality of life of older people. This paragraph provides the contextual background for the analysis presented in the following sections, offering a synthetic overview of the demographic ageing process, household composition of the elderly population, and their housing conditions in Italy.

Based on the 2021 PPHC data, households in Italy with at least one member aged 65 or older are 10.276.199 representing 39.2 percent of total households. In recent years there has also been a significant increase in the elderly living alone. In particular, the share of women who, by choice or need, form one person household reached 56% in 2021, almost double that of men (27%). This phenomenon appears especially marked among older women, peaking between the ages of 75 and 84 (AISP Report 2023). More than 1 in 3 elderly people (36.9%, about 5.2 million) live in households composed exclusively of people in their own age group, probably with their spouse or other family members, with no children or other young adults. Another significant share, 32.7 percent (about 4.6 million), are in multigenerational households. This means that nearly one-third of people aged 65 and over cohabit with members of other generations, such as children or grandchildren, a situation that can foster intergenerational exchange, responsibility sharing, and more direct social and practical support.

As for housing conditions, the majority of those over 64 live in conventional dwellings, a small share, approximately 154,000 (1.1%), in collective living quarters such as retirement homes or care facilities, while an even smaller proportion (0.2%, about 29,000) reside in other housing units[2] or are homeless. Although it is a small percentage, it represents a critical situation that requires urgent efforts for social inclusion, access to primary care services, and policies to deal with poverty and marginalization.

To better describe the housing condition of the elderly living in conventional dwellings, the following variables were considered in this paper: period of construction, useful floor space, density standard[3], tenure status and type of building.

Conventional dwellings that were built before 1919 make up about 9.5% of the total housing stock. On the other hand, conventional dwellings constructed between 1961 and 2000 are much more numerous, totalling nearly 20 million units. They represent the majority of the housing stock, accounting for over half, specifically 56.3%, which highlights the significant growth and development of residential

[2] ‘Other housing units’ are huts, cabins, shacks, shanties, caravans, houseboats, barns, mills, caves or any other shelter used for human habitation at the time of the census, irrespective if it was designed for human habitation.
[3] ‘Density standard’ relates the useful floor space in square meters or the number of rooms to the number of occupants

buildings during the latter half of the 20th century. The majority of occupied conventional dwellings in Italy fall within the size range of 80 to 99 square meters, making up about 26.7% of the total. This is followed by homes sized between 60 and 79 square meters, which account for 20.6%. Additionally, a significant portion of occupied conventional dwellings, around 18.1%, are slightly larger, ranging from 100 to 119 square meters. Average density standard per person is 44.3 m².

In 2021, nearly 72.5 percent of households reside in a home they own, while 20.5 percent live in rented housing. A residual share, on the other hand, occupies the dwelling under some other form of tenure, representing a tiny fraction compared to the prevailing forms of housing enjoyment.

With respect to the territorial distribution, the housing pattern of the elderly by degree of urbanization[4] shows that this segment of the population is more likely to live in towns and suburbs. Specifically, of the more than 14 million people aged 65 and older residing in Italy as of December 31, 2021, almost half live in intermediate density areas, more than a third in densely populated areas (cities) and 18.4 percent in areas with the lowest level of urbanization (rural areas).

In this paper, the results of new and targeted analyses conducted on data from the 2021 Permanent Population and Housing Census (PPHC) are presented. The aim is to explore the housing conditions of the elderly across different levels of urbanization, following a progressive analytical approach, from univariate descriptions to multilevel and multivariate models. First, paragraph 2 provides an initial exploratory assessment of elderly people's housing conditions, focusing, through univariate analyses, on associations between selected housing topics and the degree of urbanization. Next, paragraph 3 presents the results of multivariate and multilevel analyses, aimed at investigating in greater depth the relationships between housing conditions, the urban or rural setting, household structure (e.g., elderly-only or multigenerational households), and individual characteristics such as gender, age, and educational attainment. Particular attention is paid to the selection of statistical techniques suited to the heterogeneous nature of the available data, whether estimating probabilities or modelling discrete outcomes.

[4] The degree of urbanisation methodology classifies *local administrative units* (LAU) as cities, towns and suburbs, and rural areas based on a combination of geographical contiguity and population density. The basis for the classification is the data for 1 km² population grid cells. Each cell has the same shape and surface area, thereby avoiding distortions caused by using territorial units varying in size.
Cities: densely populated areas where at least 50% of the population lives in one or more urban centres
Towns and suburbs: intermediate density areas where less than 50% of the population lives in an urban centre and at least 50% of the population lives in an urban cluster
Rural areas: thinly populated areas where more than 50% of the population lives in rural grid cells.

2. Housing Conditions by Degree of Urbanization: Age of Building, Tenure Status and Floor Space

The housing conditions of elderly people in Italy differ significantly depending on the degree of urbanization, with urban-rural contrasts evident across several dimensions. The first aspect concerns the period of construction of dwellings occupied by the elderly. Data show that buildings constructed before 1946~~)~~ are more prevalent in rural areas (22.7%) than in towns and suburbs (13.1%) and cities (14.7%). This reflects a slower pace of housing renewal in rural municipalities and suggests the persistence of ageing housing stock, potentially associated with lower energy performance or accessibility standards. Conversely, dwellings built between 1961 and 1980 dominate across all areas, reaching peaks in cities (42.5%) and towns (40.4%), corresponding to the economic and urban expansion period. Rural areas show a more balanced distribution over the decades. At the regional level, the North-West stands out for a particularly high proportion of elderly living in rural areas in dwellings built before 1946, totalling over 31%. This value is substantially higher than the national average for rural areas (22.7%) and indicates a strong presence of older housing stock in this part of Italy. In contrast, the Center registers a lower share of pre-1946 rural dwellings (24.8%), while the North-East, South, and Islands have even smaller proportions, ranging between approximately 14% and 20%. These differences suggest that the persistence of historic housing in rural areas is more pronounced in the North-West compared to other regions.

Tenure status further highlights territorial inequalities. While homeownership is widespread among elderly households across the country, its incidence increases with decreasing urban density: 81.4% of older adults in cities are homeowners, compared to 86.1% in towns and 87.9% in rural areas. Rental tenure is significantly more common in cities (15%) than in towns (9.7%) or rural areas (6.8%), suggesting greater housing market pressures and possibly more frequent economic vulnerability among urban elderly. Interestingly, the “other” category, comprising free use, service benefits, etc., shows a modest increase in rural areas (5.2%). This may reflect more frequent forms of extended family living or informal agreements for housing use among relatives, which are more typical in less urbanized contexts, where family and property ties tend to be stronger and more deeply rooted in the local area.

Significant territorial differences also emerge. In the South, rental tenure among elderly in cities reaches 18.9%, well above the national urban average, highlighting the fragility of tenure security in this area. Similarly, the South and Islands record the highest shares of elderly in rural areas living under “other” forms of tenure (7.4% and 5.6% respectively), possibly reflecting informal cohabitation patterns. Meanwhile, the North-East presents the most consolidated pattern of ownership

(89.2% in rural areas), suggesting stronger homeowner stability in this part of the country.

Floor space data adds another layer to this territorial narrative. Smaller dwellings (under 60 square meters) are more frequent in cities (10.8%) than in towns (6.8%) or rural areas (6.9%). In contrast, the share of elderly people living in dwellings over 100 square meters rises sharply moving from urban to rural areas: from 39.1% in cities to 49.8% in towns and 53.9% in rural zones. This reflects both different housing market dynamics and possibly differing household structures, with rural elderly more likely to occupy spacious, older homes once shared with larger families. While larger floor space may suggest better living standards, it can also indicate potential underuse of space and higher maintenance burdens for ageing individuals.

The North-East rural areas stand out with over 60% of elderly residents living in dwellings larger than 100 square meters, the highest national share, confirming the region’s legacy of larger single-household homes and a tradition of multi-generational rural dwellings. Conversely, cities in the North-West register the highest proportions of elderly living in dwellings under 60 square meters, with 13.4%, well above the national average of 10.8%. In contrast, the Islands show a considerably lower share of elderly in such small dwellings, at only 7.6%, indicating a different spatial housing pattern in this regions. Taken together, these three dimensions point to a coherent territorial pattern. Elderly residents in rural areas tend to live in older, larger, and owner-occupied homes, which may offer stability but also present challenges in terms of accessibility and energy efficiency. On the contrary, urban elderly are more exposed to smaller and newer dwellings but face higher rates of renting and greater market pressures. These contrasts underscore the need for differentiated housing policies that take into account the interaction among dwelling characteristics, tenure security and demographic ageing across diverse territorial settings. Moreover, regional disparities within urban and rural areas signal the importance of considering not only the urban-rural divide but also context-specific factors such as historical housing development and local housing markets.

3. Uncovering Territorial and Individual Patterns in Elderly Housing Conditions

Understanding housing conditions among the elderly requires an analytical framework capable of capturing the multidimensional and territorially embedded nature of such phenomena. To this end, the analysis presented in this paragraph adopts a multilevel analytical design, combining analyses at the municipal level, which reflect contextual and structural characteristics with evaluations at the individual level, taking into account specific personal and household factors.

A distinctive feature of the PPHC is that many housing-related variables, such as floor space, tenure status, and type of building, are produced as model-based estimates. These are highly reliable at aggregated levels, such as municipalities or population subgroups, but less suitable for granular, individual-level inference. As a result, a multilevel strategy is not only methodologically appropriate but also dictated by the architecture of the data.

The first analytical step focuses on municipalities as units of analysis and proxies for the broader territorial context, using the degree of urbanization as a key structural dimension. The aim is to identify homogeneous clusters of municipalities based on demographic and housing characteristics of the elderly population. This allows us to delineate territorial profiles that reflect different combinations of population ageing, housing supply, tenure patterns and settlement structures. The second analytical step shifts to the individual level, investigating whether specific vulnerability profiles—such as elderly individuals living alone and with low education—are subject to differentiated housing conditions depending on the municipal cluster in which they reside. This approach enables a more nuanced analysis of housing inequality by considering the interaction between individual characteristics and territorial context.

The target population of the study considers all individuals aged 65 and over, residing in conventional dwellings. The dataset includes a rich list of variables that could be grouped into three main domains:

- *housing variables*: tenure status (owner, renter, or other); useful floor space and density standard: total and per capita surface; type of building: single-family house or apartment block; period of construction: dwellings built before 1919 considered as ‘older buildings’ and those built after 2001 as ‘recent buildings.
- *sociodemographic variables*: age, sex, educational attainment, citizenship.
- *household composition variables*: household composition: living alone, elderly-only or multigenerational;
- *territorial classification*: degree of urbanization (urban, suburban, rural) and geographical macro-area (North, Centre, South, Islands).

The degree of urbanization serves as a fundamental classification axis in this framework. Defined according to population density and settlement characteristics, it distinguishes densely populated areas (*cities*), with high population density and significant urban centres; intermediate areas (*towns and suburbs*): combining urban and rural traits; thinly populated areas (rural areas): marked by low density and dispersed settlements (Eurostat, 2021).

By combining territorial profiles with individual indicators, this framework enables a refined reading of housing inequalities, helping to identify who is vulnerable, where vulnerabilities concentrate, and how personal and contextual factors jointly shape housing outcomes in later life. This multilevel perspective is also intended to inform place-sensitive housing policies that move beyond binary urban-rural distinctions.

3.1. Municipal-Level Cluster Analysis

The first analytical step involved constructing a typology of Italian municipalities based on a set of housing and demographic indicators referring to the elderly population (aged 65 or over). The selected variables are: the proportion of elderly residents and those aged 85 or more, the share of women, the share of people (aged 65 or over) living alone, the share of elderly-only and multigenerational households, the proportion of households in rented dwellings, dwelling characteristics: useful floor space, type of building (single vs. multi-unit) and period of construction, the educational attainment (share with low education).

It is important to emphasize that territorial variables—such as the degree of urbanization and geographical macro-area—were not used in the clustering process as the first aim was to identify clusters based solely on housing and demographic characteristics, potentially transversal across territorial classes. However, in all subsequent multivariate and multilevel analyses, territorial variables are included, and the resulting clusters result indeed strongly associated with territorial features.

To address redundancy and ensure parsimony, a Principal Component Analysis (PCA) was applied to the standardized indicators and the resulting components were fed into a Ward's hierarchical clustering algorithm. This procedure led to the identification of eight distinct clusters of municipalities, each representing a typical configuration of housing and demographic characteristics for the elderly population.

Once defined, the clusters were examined in relation to territorial attributes to support interpretation. Table 1 shows the percentage distribution of municipalities within each cluster by urbanization degree (cities, towns/suburbs, rural areas) and geographical macro-area.

The results reveal a strong association between the eight clusters and territorial attributes, especially the degree of urbanization, although this variable was not used as input in the clustering procedure. Urban municipalities are entirely concentrated in a single cluster (Cluster 1). Suburban areas are split across two distinct clusters: Cluster 2, with a marked concentration in the North-West, and Cluster 3, covering a broader range of territories across the Centre, South, and Islands. Even more diversified are the rural areas, which are distributed among five separate clusters. Each rural cluster shows a distinct geographical imprint: Cluster 4 includes rural municipalities predominantly from the South and Islands; Cluster 5 represents a North-West and Central rural mix; Cluster 6 gathers rural areas in Northern regions; Cluster 7 spans the North-East and the Islands; finally, Cluster 8 is almost exclusively composed of North-Western rural municipalities. These patterns confirm the territorial specificity of the resulting clusters and the greater internal differentiation observed in non-urban contexts, particularly in rural settings.

Table 1 - *Territorial profile of Clusters (Urbanization and Macro-regions).*

Cluster	N.	Cities	Suburbs/Towns	Rural Areas	North West	North East	Centr.Italy	South	Main Islands
1	255	100,0%	0.0%	0.0%.	43.5%	8.2%	5.5%	38.4%	4.3%
2	1.919	0.0%	100.0%	0.0%	54.1%	11.6%	13.0%	15.6%	5.7%
3	687	0.0%	100.0%	0.0%	6.7%	41.3%	4.2%	29.4%	18.3%
4	1.669	0.0%	0.0%	100.0%	8.3%	7.3%	18.2%	50.1%	16.1%
5	1.380	0.0%	0.1%	99.9%	45.8%	9.3%	19.2%	22.6%	3.0%
6	751	0.0%	0.0%	100.0%	50.1%	37.3%	10.9%	1.5%	0.3%
7	598	0.0%	0.0%	100.0%	16.2%	44.6%	1.0%	3.5%	34.6%
8	645	0.0%	0.0%	100.0%	86.2%	9.9%	2.8%	0.6%	0.5%
Total	*7.904*	*3%*	*33%*	*64%*	*38%*	*18%*	*12%*	*23%*	*10%*

Source: Elaborations on Permanent Population and Housing Census (PPHC) data, Istat.

Figure 1 presents a heatmap that synthesizes the distribution of key demographic and housing indicators across the eight municipal clusters identified in the analysis. Each row corresponds to a cluster, while columns represent standardized values of selected variables, including age structure, household composition, tenure status, dwelling characteristics and education levels. The colour gradients highlight deviations from the overall average, allowing for a rapid visual comparison and identification of distinctive cluster profiles.

Figure 1- *Heatmap of key elderly demographic and housing indicators by cluster.*

CLUS.	% Elderly *	% Women	%Over 85	%Low Attein.	% Living Alone	% Elderly-only Households	% Multi-gen. Households	%Rentals	%Low FloorP	%Single Housing	%Rece Buildin
1	21.69	**56.01**	13.68	41.38	26.66	38.07	35.27	**14.18**	**5.66**	12.15	33.04
2	23.19	54.90	14.27	44.88	28.19	**38.82**	32.99	10.05	3.71	22.79	34.37
3	22.77	54.78	14.52	**52.34**	25.48	37.79	**36.73**	6.96	2.85	43.02	**36.01**
4	26.73	54.41	**17.78**	**55.26**	31.17	34.33	34.51	5.85	3.62	40.57	29.36
5	**30.08**	53.47	**18.43**	49.22	**41.28**	31.88	26.85	7.24	3.78	32.61	21.85
6	22.94	53.53	14.16	45.22	28.55	**38.58**	32.87	8.40	2.69	33.67	34.91
7	26.58	54.24	16.30	**54.04**	27.30	35.60	**37.10**	5.15	1.68	**61.92**	32.69
8	**29.16**	53.48	16.88	45.58	33.90	35.05	31.05	6.69	1.53	**61.19**	18.37
Total Avg	25.77	54.28	16.10	49.15	31.25	35.97	32.78	7.73	3.26	36.80	29.85

Source: Elaborations on PPHC data, Istat.

The main findings can be summarized as follows:

- *Cluster 1 (Urban)* is typified by moderate elderly shares, a higher female proportion, predominantly elderly-only households, high education levels and a tenure status dominated by small rental flats in multi-unit buildings.
- *Cluster 2 (Suburban NW)* exhibits similar demographic features with slightly elevated elderly presence and living alone rates, combined with recent buildings and a high share of rented flats.
- *Cluster 3 (Suburban Elderly Households)* stands out with lower educational attainment but better housing conditions: larger floor space, home ownership and prevalence of one-person households.
- *Cluster 4 (Southern Rural, Multigenerational)* reflects the highest elderly and oldest-old proportions, lower education, predominance of multigenerational households, smaller, older, owner-occupied dwellings.
- *Cluster 5 (Rural NW and Centre, Elderly Living Alone)* records the oldest population, highest living-alone rates, mixed tenure, and small dwellings, indicating elevated vulnerability.
- *Cluster 6 (Northern Rural)* shows moderate elderly shares, medium-sized dwellings, and higher prevalence of recent buildings.
- *Cluster 7 (Rural with Good Housing)* combines high elderly presence with multigenerational households, low education but favourable housing conditions with large-size, one-person households and low rental rates.
- *Cluster 8 (Elderly Rural NW, Living Alone)* features high elderly and living alone shares, mixed tenure with many one-person households but older building stock, indicating potential maintenance issues.

These clusters highlight the marked heterogeneity in elderly housing conditions, shaped by both the urban–rural gradient and regional specificities. This typology reveals not only a clear urban–rural divide, but also significant intra-category disparities—particularly within rural areas—underscoring the limitations of dichotomous classifications. Rural clusters display wide internal variation: some are marked by isolation and limited space, others by better housing conditions but lower levels of educational attainment.

3.2. Individual-Level Vulnerability Profiles across Clusters

The second stage of the analysis focuses on housing vulnerability among older individuals, examined in light of the territorial clusters identified in the previous section. The objective is to assess whether elderly persons with similar sociodemographic characteristics are exposed to different housing risks, depending on the territorial features of the municipality in which they reside.

Two vulnerability profiles were examined, based on potential risk-enhancing covariates:

- education[5]: elderly aged 75+ living alone with high education (upper secondary or tertiary) versus those with low education (primary or lower secondary);
- household-composition: elderly living alone compared to elderly-only households.

The housing risk indicators considered include tenure status (share of renters), dwelling size (share living in units smaller than 60 m²), floor space per person below the national median, and type of building (single-unit vs. multi-unit buildings).

The results of these preliminary exploratory analyses are summarized in Tables 3 and 4 showing, for each profile, how the values of housing risk indicators vary with changes in the covariates and across territorial clusters. Only clusters with significant differences are included. Both the observed percentages and the percentage-point differences among profiles within each cluster are reported, allowing for a clear comparison of housing risks by educational attainment (Table 3) and household composition (Table 4) across territorial contexts.

Table 3 - *Housing Risk Indicators by Education Level for Elderly Living Alone, by Cluster. Values and percentage point differences by profile and cluster*.

		Risk Housing Indicators (Educated/Less-Educated)			
	Cluster	Rent (%)	Single Housing (%)	Low Density (%)	Small Dwellings (%)
1	Urban Elderly Conditions	28 / 26	2.6 / 3.6	13.6 / 15.0	0.1 / 0.2
3	Suburban Elderly Households	60 / 55	1.2 / 1.5	3.3 / 4.8	0.04 / 0.14
6	Rural North Seniors	4 / 7	2.3 / 4.2	11.6 / 15.3	0.03 / 0.13
8	Elderly Rural North-West Living Alone	48 / 60	1.2 / 1.8	5.5 / 6.5	0.35 / 0.57
		Difference Between Educated vs Less-Educated (pp)			
	Cluster	Rent (%)	Single Housing (%)	Low Density (%)	Small Dwellings (%)
1	Urban Elderly Conditions	▲ + 2	▼ – 1.0	▼ – 1.4	▼ – 0.1
3	Suburban Elderly Households	▲ + 5	▼ – 0.3	▼ – 1.5	▼ – 0.10
6	Rural North Seniors	▼ – 3	▼ – 1.9	▼ – 3.7	▼ – 0.10
8	Elderly Rural North-West Living Alone	▼ – 12	▼ – 0.6	▼ – 1.0	▼ – 0.22

Source: *Elaborations on PPHC data, Istat.*

Results in Table 3 highlight substantial variation across clusters in housing outcomes by educational profile. In rural clusters (e.g. 6 and 8), individuals with low education exhibit higher vulnerability, especially in rental rates and housing size. In Cluster 8 (rural north-west), the share of renters among the less educated reaches 60%, compared to 48% among the more educated. Conversely, in highly urbanised areas such as Cluster 1, the gap among education groups narrows considerably, suggesting that housing market constraints may limit the protective role of higher education.

Table 4 compares older individuals living alone to those in elderly-only households. Across most clusters, those living alone experience higher risk,

[5] Education is used to identify potential vulnerability profiles among older adults, as it serves as a robust socioeconomic proxy and was readily available in the first dataset, derived from census surveys and estimates. Future analyses, integrating additional data sources, could consider variables more directly linked to income

particularly in terms of dwelling size and rental incidence. These differences are more pronounced in clusters with greater urban pressure. For example, in Cluster 8, the share of small dwellings is nearly half a percentage point higher for those living alone (0.57% vs. 0.10%). The evidence indicates that individual vulnerability does not translate into uniform housing disadvantage across space. Territorial characteristics modulate the extent to which education and household composition impact housing outcomes.

Table 4 - *Housing Risk Indicators by Household Composition among Elderly, by Cluster. Values and percentage point differences by profile and cluster*.

		Risk Housing Indicators (Alone/ Elderly-only Households)			
	Cluster	Rent (%)	Single Unit (%)	Low Density (%)	Small Dwellings (%)
1	Urban Elderly Conditions	26 / 25	3.6 / 3.4	15.0 / 12.5	0.20 / 0.14
4	Southern Rural Multigenerational	23 / 29	2.5 / 1.9	8.9 / 5.3	0.23 / 0.07
6	Rural North Seniors	7 / 8	4.2 / 3.9	15.3 / 10.5	0.13 / 0.06
8	Elderly Rural North-West Living Alone	60 / 57	1.8 / 1.4	6.5 / 5.0	0.57 / 0.10
		Difference Between Alone vs Elderly-only Households (pp) (positive = higher in Educated profile)			
	Cluster	Rent	Single Unit	Low Density	Small Dwellings
1	Urban Elderly Conditions	▲ +1	▲ +0.2	▲ +2.5	▲ +0.06
4	Southern Rural Multigenerational	▼ –6	▲ +0.6	▲ +3.6	▲ +0.16
6	Rural North Seniors	▼ –1	▲ +0.3	▲ +4.8	▲ +0.07
8	Elderly Rural North-West Living Alone	▲ +3	▲ +0.4	▲ +1.5	▲ +0.47

Source: Elaborations on PPHC data, Istat.

These results support the relevance of a multilevel perspective, which enables a more granular understanding of how structural (municipal-level) and compositional (individual-level) factors jointly shape housing inequalities among older adults, and suggest that further research should include statistical modelling of interactions.

4. Conclusions

This study addresses the complexity of jointly considering territorial, household, and individual-level factors in investigating housing conditions among the elderly, illustrating a significant advantage of the Permanent Population and Housing Census data. Using a multilevel framework, the analysis integrates contextual municipal classifications with individual vulnerability profiles, shedding light on how structural housing inequalities and personal characteristics interact.

The reliance on model-based housing estimates in the 2021 Permanent Census requires analytical approaches that acknowledge the hierarchical nature of data and its varying precision. Aggregated municipal-level data offer a robust basis for territorial classification, while individual-level analysis enables the identification of vulnerable subpopulations within their broader territorial contexts.

The cluster analysis confirms urbanization as a key stratifying dimension. Urban municipalities form a distinct cluster characterized by high rental rates, limited floor space and a concentration of elderly living alone. Suburban areas split into contrasting profiles, such as a North-West cluster with more recent housing and higher renter prevalence, versus a more heterogeneous suburban type with different household and housing conditions. Rural areas prove highly diversified, dividing into multiple clusters ranging from multigenerational households in modest owned homes to aged populations living alone in smaller dwellings. Integrating these territorial clusters with individual vulnerability profiles reveals important interaction effects. Education appears protective against adverse housing outcomes in several rural and suburban clusters but shows limited influence in densely populated urban settings where structural housing constraints dominate. Such findings highlight the need for housing policies sensitive to both individual vulnerabilities and territorial contexts. Building on these results, future research will focus on validating and refining the typology across demographic and housing subgroups, modelling interaction effects through multivariate approaches and enhancing territorial classifications by introducing finer stratifications at municipal, household and individual levels.

References

ITALIAN ASSOCIATION FOR POPULATION STUDIES (AISP) 2023. *Population Report. Families in Italy: Types, Obstacles, and Challenges*. Bologna: il Mulino.

ISTAT 2025. *Annual Report – The State of the Nation*. Rome: ISTAT.

MARCELLINI F., RIVA A., CECCHI F. 2007. Aging in Italy: Urban–rural differences. *Archives of Gerontology and Geriatrics*, Vol. 44, No. 2, pp. 213-227.

EUROSTAT 2021. Applying the degree of urbanisation – A methodological manual to define cities, towns and rural areas for international comparisons – 2021 edition. Luxembourg: Publications Office of the European Union.

ISTAT – ITALIAN NATIONAL INSTITUTE OF STATISTICS 2024. Abitazioni in Italia: dati e analisi aggiornate. Rome: ISTAT. Available at: https://www.istat.it/wp-content/uploads/2024/08/Today-Abitazioni_01_08-2024.pdf

Angela CHIEPPA, Istat. chieppa@istat.it
Simona MASTROLUCA, Istat. mastrolu@istat.it
Alessandro SASSO, Istat, alesasso@istat.it

ASSESSING ENERGY POVERTY IN ITALY: AN ENDOGENOUS CUT-OFF DETERMINATION

Gloria Polinesi, Maria Cristina Recchioni, Mariateresa Ciommi, Francesca Mariani, Alfonso Carfora

Abstract. In recent years, the key role of energy in individual well-being has drawn the attention of policymakers, institutions, and researchers to the issue of energy poverty. Energy poverty refers to a situation in which households struggle to access or afford essential energy services. The vulnerability of several households has worsened due to the reduction of disposable incomes resulting from the labor market crisis, escalating unemployment triggered by the COVID-19 pandemic, and the increase in domestic energy demand brought about by the extended periods of lockdown.
The expenditure-based approach defines energy poverty as the inability to afford adequate energy services while keeping expenditures within a reasonable level arbitrarily defined. A commonly used cut-off considers household energy-poor if the ratio between energy expenditure - in heating and electricity consumption - and the total expenditure is greater than 10%.
With the aim of investigating the theoretical foundation of this ratio we implement a stratification approach identifying sub-groups of cut-offs at each iteration by approximating their distribution with a sequence of two-component log-normal mixtures. Thus, different cut-offs are not fixed a priori but rather endogenously by the iterative procedure.
Using data from the Household Budget Survey (HBS) provided by the Italian National Institute of Statistics in the years 2019-2022, the stratification algorithm supports the assessment of the percentage of Italian households in energy poverty highlighting differences pre and post COVID-19.

1. Introduction

Energy poverty has emerged as one of the most pressing challenges of the 21st century, representing a critical intersection between energy security, social equity, and sustainable development. This multifaceted phenomenon affects households across both developing and developed countries, transcending traditional economic boundaries and highlighting the universal nature of energy access efforts in contemporary society.

Energy poverty is broadly defined as the lack of access to affordable, reliable, and adequate energy services essential for human well-being and socio-economic development (Bouzarovski, 2014; Zarghami, 2025). This definition encompasses not only the absolute absence of energy access but also situations where households face disproportionate energy costs relative to their income, inadequate energy services that fail to meet basic needs, or unreliable energy supply that disrupts daily activities and economic opportunities.

In Europe, the proportion of the population unable to adequately heat their homes rose dramatically from 6.9% in 2021 to 10.6% in 2023 (EC, 2025), largely driven by the energy crisis following geopolitical tensions and supply chain disruptions. The Italian context provides a particularly compelling case study for examining energy poverty dynamics in developed economies. By the end of 2023, approximately 2.36 million Italian households - representing 9% of the national population - were experiencing energy poverty conditions, marking one of the highest prevalence rates recorded since the inception of systematic data collection in 1997 (OIPE, 2024).

The complexity of energy poverty extends beyond simple access metrics to encompass multiple dimensions including affordability, reliability, quality, and sustainability of energy services. This multidimensional nature has significant implications for measurement approaches, as traditional single-indicator metrics often fail to capture the full spectrum of energy poverty experiences (Kashour and Jaber, 2024). Given the urgency and complexity of energy poverty, there is a critical need for robust, comprehensive, and context-sensitive measurement approaches that can inform effective policy interventions and track progress toward energy equity goals. Accurate measurement of energy poverty is essential for identifying affected populations, understanding the underlying drivers, evaluating the effectiveness of interventions, and ensuring that energy transition policies do not inadvertently exacerbate existing inequalities.

Energy poverty measurement has evolved through three distinct methodological frameworks, each offering unique perspectives on household energy deprivation. The expenditure-based approach employs quantitative indicators examining the relationship between household energy expenditure and income levels, including the widely-used 10% threshold rule, the Low Income High Cost (LIHC) indicator combining high energy costs with low disposable income (Hills, 2011), and adaptations incorporating vulnerability components (Faiella and Lavecchia, 2015). Recent applications have explored hidden energy poverty across EU countries (Menyh'ert, 2024) and micro-level elasticity relationships between energy prices and poverty status (Bardazzi *et al.*, 2024). The consensus-based approach captures subjective experiences through household surveys, addressing energy poverty's multidimensional nature by incorporating indicators such as heating difficulties, bill payment delays, and perceived energy discomfort. This methodology can adopt either a union approach (considering at least one indicator) or an intersection approach requiring simultaneous deprivation across multiple dimensions, with scholars increasingly favoring multidimensional composite indicators (Marchand, 2019; Simionescu, 2024). The direct measurement approach evaluates energy services against established benchmarks using objective temperature and consumption data (Okushima, 2019), providing empirical evidence while potentially overlooking contextual variations in energy needs (Sy and Mokaddem, 2022). This methodological diversity reflects the scholars' ongoing debates regarding optimal quantification approaches. Expenditure-based methods offer objectivity and income poverty connections but may overlook energy rationing behaviors. Conversely, consensus-based indicators capture multiple deprivation dimensions but can be influenced by demographic and cultural perception biases. This methodological complexity underscores energy poverty's multidimensional nature, situated at the intersection of economic, social, and housing well-being dimensions

This paper addresses this methodological challenge by proposing a novel analytical framework that advances beyond traditional energy poverty measurement approaches. Our research introduces a hybrid methodology that integrates statistical modelling techniques with socioeconomic analysis to overcome the limitations of arbitrary threshold selection that characterizes existing literature.

The core innovation lies in the development of a two-phase analytical process. Initially, we employ advanced statistical procedures to identify natural breakpoints within energy spending patterns, moving away from conventional fixed percentage rules. Subsequently, we incorporate broader household consumption patterns to construct a comprehensive poverty assessment that reflects both energy-specific vulnerabilities and general economic constraints.

This integrated approach represents a methodological advancement in energy poverty research, offering a more empirically grounded and contextually sensitive measurement tool.

2. Methodology

The first step in our approach involves the determination of endogenous cut-off points for energy expenditure. Rather than relying on a fixed threshold - such as the widely used 10% rule - we adopt a data-driven method that accounts for heterogeneity in household spending behavior. Specifically, we implement an iterative stratification procedure based on two-component log-normal mixture models, which allows for the identification of distinct segments within the energy expenditure distribution. These captures underlying subpopulations and provide a theoretically grounded means of determining cut-off values that are internally consistent with the observed data.

In the second stage, we use these endogenously derived thresholds to estimate the incidence of energy poverty. This estimation is conducted in conjunction with information on total household expenditure, focusing on households whose overall consumption falls below 60% of the national median. By combining relative economic constraints with energy expenditure cut-offs, our method offers a nuanced and regionally sensitive measure of energy poverty. This dual-criteria approach ensures that both affordability and social inclusion considerations are reflected in the final energy poverty indicator.

2.1. Determination of the endogenous cut-offs

The recent iterative stratification procedure applied to household incomes by Mariani *et al.* (2022) can be adapted to analyze the left-hand tail distribution of the ratio between energy expenditure and total expenditure (see Polinesi *et al.*, 2025).

Under the assumption that this ratio can be approximated by a univariate log-normal mixture, in the first iteration, the stratification procedure identifies the so-called change point and divides the ratios into two distinct groups: ratios that are smaller than or equal to the change point and ratios that are larger than the change point.

The change point is the threshold where the leftmost component of the mixture dominates the rightmost component to its left and is dominated by it to its right. This indicates that a structural change in the distribution occurs. In each subsequent iteration, the procedure approximates with a lognormal mixture the right group of returns from the previous iteration, appropriately shifted. It identifies the change point associated with this right group and then, it splits this group into two sub-groups.

The procedure stops when it fails to find a new change point or when the new change point is larger than the median of the ratios.

In its first iteration, the procedure identifies the first change point a^1 to split the set of all ratios $\mathcal{S}_n = \{y_1, y_2, \dots, y_n\}$ into two disjoint groups: the left group $\mathcal{K}_1 = \{y \in \mathcal{S}_n \ \wedge \ y \in (0, a^1]\}$, composed of ratios smaller than or equal to the threshold value, and a right group $\mathcal{R}_1 = \mathcal{S}_n \setminus \mathcal{K}_1$, composed of ratios larger than the threshold value.

In the second iteration, the procedure considers the subset $\mathcal{R}_1$, obtained in the first iteration. It identifies a new threshold value $a^2 > a^1$, and splits $\mathcal{R}_1$ into two disjoint groups: the left group $\mathcal{K}_2 = \{y \in \mathcal{S}_n \ \wedge \ y \in (a^1, a^2]\}$ and the right group $\mathcal{R}_2 = \{y \in \mathcal{S}_n \ \wedge \ y \in (a^2, +\infty)\}$. In the k-th iteration, the algorithm proceeds similarly to the first two iterations, by identifying the threshold a^k and dividing the set $\mathcal{R}_{k-1}$ into two groups: $\mathcal{K}_k = \{y \in \mathcal{S}_n \ \wedge \ y \in (a^{k-1}, a^k]\}$ and $\mathcal{R}_k = \{y \in \mathcal{S}_n \ \wedge \ y \in \left(a^k, +\infty\right)\}$.

The vector of unknown parameters for the density functions associated with the two mixture components $\Theta_k = (\pi_k, \mu_{1,k}, \mu_{2,k}, \sigma_{1,k}, \sigma_{2,k})'$ is estimated using the return in the set $\mathcal{R}_{k-1}$ through the expectation maximization (EM) algorithm (see Dempster *et al*., 1977). Note that $\pi_k \in [0,1]$ is the mixing weight representing the a priori probability that the point $x = y - a^{k-1}$, $y \in \mathcal{K}_{k-1}$, for $k = 1,2,\ldots$, belongs to the first component[1]. The change point a^k of the mixture is determined using the following rule:

$$a^k = \min\{y \in \mathcal{R}_{k-1} \wedge \pi_k f_{1,k}(y - a^{k-1}) = (1-\pi_k) f_{2,k}(y - a^{k-1})\}, \qquad (1)$$

where $f_{1,k}(x)$ and $f_{2,k}(x)$, $x \in \mathbb{R}_+$, are the log-normal densities of parameters $\mu_{1,k}$, $\mu_{2,k}$, $\sigma_{1,k}$, $\sigma_{2,k} \in \mathbb{R}$ associated with the two mixture components. As already specified, at step k the change point represents the smallest point where the leftmost component is equal to the rightmost component.

The change point a^k is the frontier of the two groups $\mathcal{K}_k$ and $\mathcal{R}_k$ at the k-th iteration, and, broadly speaking, a^k divides the sample into two subsamples with non-homogeneous distributions. The procedure stops when a new a^k cannot be determined (i.e., Eq. (1)) does not admit any solution.

2.2. *Estimation of the energy poverty indicator*

In the second stage of the analysis, we estimate energy poverty using a joint condition that reflects both relative spending effort and overall economic constraint. The indicator is defined in terms of two key components of household consumption:

- X, representing energy expenditure, and
- Z, representing all other expenditures.

Specifically, we consider the share of energy expenditure relative to total consumption, $\frac{X}{X+Z}$, and compare it to an endogenously determined threshold, identified as the complement to one of the first change point a^1 (i.e., k=*1*) of the stratification procedure. At the same time, it assesses the household's total expenditure $X + Z$ relative to the national median *m*.

A household is classified as energy poor if it satisfies the double condition:

$$\frac{X}{X+Z} > a^1 \wedge X + Z < m \qquad (2)$$

[1] Hereafter we assume $a^0 = 0$.

This definition captures households that are simultaneously burdened by disproportionately high energy costs and have limited overall spending capacity. By combining these two criteria - excessive energy burden and low total expenditure - the indicator provides a refined, behaviorally informed measure of energy poverty.

Specifically, our indicator accounts for how people adjust their energy consumption in response to their financial constraints, rather than simply assuming that all households have similar energy needs or consumption patterns.

3. Data and results

This study investigates energy poverty of Italian households at regional level by drawing on data from the Household Budget Survey (HBS) conducted by ISTAT over the period 2019-2022. The HBS provides detailed information on household consumption patterns across Italy, capturing changes in both the level and composition of spending. It incorporates key social, economic, and geographic variables, allowing for a detailed analysis of household behavior. The survey plays a central role in producing official estimates of relative and absolute poverty, as well as in calculating inflation indicators based on expenditure categories. By combining consumption data with socio-demographic characteristics, the HBS serves as a vital source for understanding economic conditions and informing both public policy and market strategies.

In this section, we also present the results for the energy poverty indicator defined in Eq. (2), separately for the four years considered. First, following Eq. (1), we compute the endogenous cut-offs at the regional level (Table 1). Then, we illustrate the values of the headcount ratio by highlighting differences between pre and post COVID-19 (Figure 1) and its spatial distribution across regions (Figure 2).

Table 1 shows that the endogenous cut-offs are generally consistent with the ten percent rule (Boardman, 1991). This rule defines a household as being in energy poverty if its energy expenditure exceeds 10% of its income (or total expenditure), where the 10% threshold represents the minimum energy required to achieve a basic level of comfort. However, the endogenous cut-offs derived from Eq. (1) vary both across regions and within the same region, highlighting that policy guidelines should be tailored to the specific characteristics of each area.

Table 1 - *Endogenous cut-offs according with Eq. (1) for each region by year.*

	2019	2020	2021	2022
Piemonte	0.117	0.134	0.12	0.157
Valle d'Aosta	0.106	0.135	0.106	0.123
Lombardia	0.08	0.092	0.086	0.087
Trentino-Alto Adige	0.07	0.094	0.08	0.112
Veneto	0.122	0.094	0.083	0.112
Friuli-Venezia Giulia	0.122	0.08	0.086	0.093
Liguria	0.063	0.078	0.077	0.069
Emilia-Romagna	0.078	0.091	0.086	0.119
Toscana	0.08	0.082	0.082	0.092
Umbria	0.085	0.092	0.10	0.103
Marche	0.093	0.08	0.098	0.119
Lazio	0.092	0.098	0.093	0.103
Abruzzo	0.104	0.111	0.111	0.136
Molise	0.127	0.112	0.098	0.223
Campania	0.089	0.083	0.089	0.102
Puglia	0.102	0.109	0.124	0.116
Basilicata	0.128	0.154	0.095	0.139
Calabria	0.098	0.099	0.121	0.157
Sicilia	0.104	0.10	0.11	0.114
Sardegna	0.135	0.072	0.093	0.109

Figure 1 shows the fraction of households in energy poverty, as defined by Eq. (2). Notably, the Southern regions exhibit higher values. Overall, the share of households in energy poverty has decreased - the national average declined from 11.2% in 2019 to 10.7% in 2022 - highlighting regional disparities before and after the COVID-19 pandemic. This trend is consistent with the estimates presented in the latest energy poverty report for Italy (OIPE, 2024), which indicate that the proportion of energy-poor households fell from 8.5% in 2019 to 7.7% in 2022.

Figure 1 – *Fraction of households in energy poverty according with Eq. (2) for the years 2019 (red circle), 2020 (green circle), 2021 (cyan circle) and 2022 (purple circle). The graph shows the difference between the values of 2022 and those of 2019.*

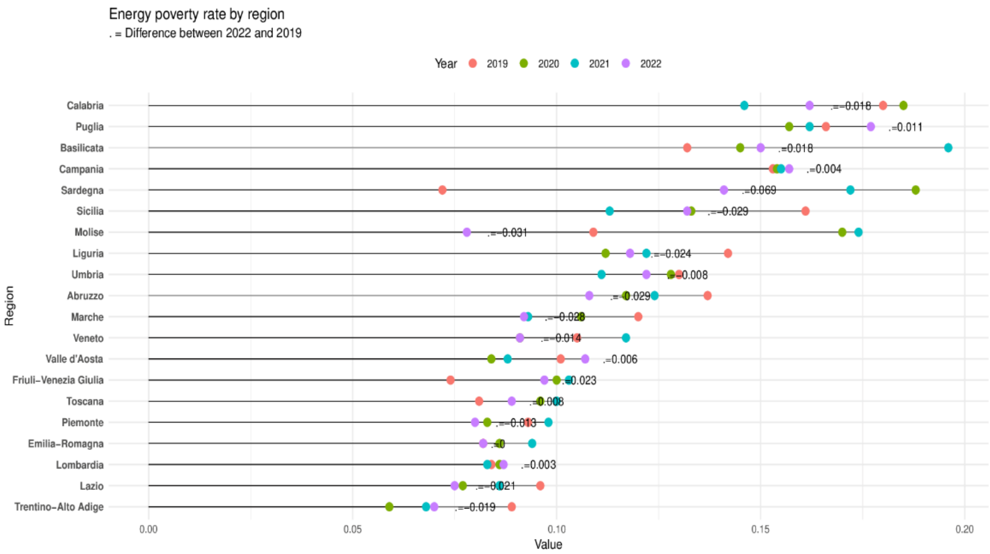
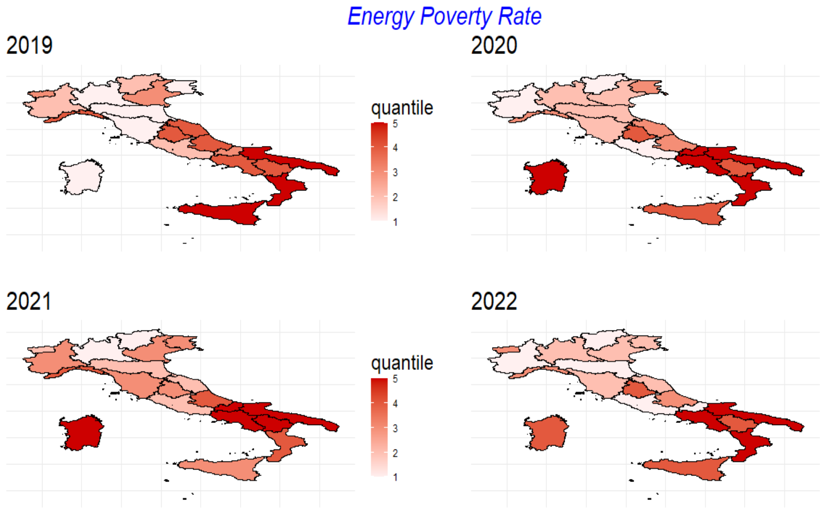


Figure 2 illustrates a geographical representation of how the energy poverty indicator is distributed across the Italian regions over the years 2019-2022. Its values are plotted following a colour scale with darker colours representing higher levels of energy poverty.

A distinct divide between northern and southern regions is evident, supporting the need to develop specific energy poverty reduction strategies for more vulnerable groups. The availability of a local area-level indicator can be a first step towards a national energy poverty dashboard that supports the design of targeted policies to combat energy poverty at the local level (Lavecchia *et al*., 2024).

Figure 2 - *Spatial distribution of households in energy poverty for the years 2019-2022. Darker colours indicate higher level of energy poverty.*



4. Conclusions

Energy poverty continues to represent a significant and persistent social issue, particularly in contexts marked by economic vulnerability and unequal access to resources. Traditional approaches to measuring energy poverty - often based on fixed expenditure thresholds - risk oversimplifying the diverse and complex realities experienced by households.

In response, this study proposes a stratification-based methodology that offers a data-driven alternative to standard benchmarks. By identifying endogenous cut-offs in energy spending behavior, the approach better reflects the heterogeneity in household circumstances and provides more nuanced estimates of energy deprivation.

Our findings reveal that energy poverty levels are consistently higher in Southern Italy, underscoring the need for geographically targeted interventions aimed to reduction strategies.

References

BARDAZZI R., GASTALDI F., IAFRATE F., PANSINI R. V., PAZIENZA, M. G., POLLASTRI C. 2024. Inflation and Distributional Impacts: Have Mitigation

Policies been Successful for Vulnerable and Energy Poor Households? *Energy Policy*, Vol. 188, pp. 114082.
BOARDMAN B. 1991. Fuel Poverty: From Cold Homes to Affordable Warmth. Pinter Pub Limited.
BOUZAROVSKI, S., 2014. Energy Poverty in the European Union: Landscapes of vulnerability, WIREs *Energy Environ*., Vol 3, pp. 276-289.
DEMPSTER AP, LAIRD NM, RUBIN DB. 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society Series B*, Vol. 39, No. 1, pp. 1-38.
EC-European Commission. 2025. Framing Summer Energy Poverty - Insights and Recommendations for a Resilient Future - *Final Report,* Office of the European Union.
FAIELLA I., LAVECCHIA L. 2015. La Povertà Energetica in Italia, *Politica Economica*, Vol. 31, No. 1, pp. 27-76.
HILLS J. 2011. Fuel poverty: the problem and its measurement. Department for Energy and Climate Change.
KASHOUR M., JABER M. M. 2024. Revisiting Energy Poverty Measurement for the European Union, *Energy Research & Social Science*, Vol. 109, pp. 103420.
KHALID U., SHAFIULLAH M., CHAUDHRY S. M. 2024. Does Conflict Aggravate Energy Poverty? *Energy Policy*, Vol. 194, pp. 114317.
LAVECCHIA L., MINIACI R., VALBONESI P., VENKATESWARAN G. 2024. Energy Poverty Risk: A Spatial Index Based on Energy Efficiency. *Occasional Paper 864*. Banca d’Italia, Roma.
MARCHAND R., GENOVESE A., KOH S.C.L., BRENNAN A. 2019. Examining the Relationship between Energy Poverty and Measures of Deprivation, *Energy Policy*, Vol. 130, pp. 206-217.
MARIANI F., CIOMMI M., CHELLI F.M., RECCHIONI M.C. 2020. An Iterative Approach to Stratification: Poverty at Regional Level in Italy, *Social Indicators Research*, Vol. 31, pp. 1.
MENYH´ERT, B. 2024. Energy poverty in the European Union. The Art of Kaleidoscopic Measurement, *Energy Policy*, Vol. 190, pp. 114160.
POLINESI G., CARFORA A., RECCHIONI M.C., MARIANI F., CIOMMI M. 2025. A New Distributional Based Indicator to Measure Energy Poverty in Italy, *Mimeo*.
OIPE 2024. In FAIELLA I., LAVECCHIA L., MINIACI R., VALBONESI P. (Eds.) La povertà energetica in Italia nel 2022, OIPE.
OKUSHIMA, S. 2019. Understanding Regional Energy Poverty in Japan: A direct Measurement Approach, *Energy and Buildings*, Vol. 193, pp. 174-184.

SIMIONESCU M., CIFUENTES-FAURA J. 2024. Evaluating the Relationship between Income Inequality, Renewable Energy and Energy Poverty in the V4 Countries, *Energy Research & Social Science*, Vol. 115, pp. 103640.

SOTO G., NGHIEM X.-H., MARTINEZ-COBAS X. 2025. Insufficiency of Renewable Energy. How do the Transition to Green Energy Economies and Foreign Direct Investment Affect Energy Poverty in Europe*? Energy*, pp. 135350.

´SMIECH S., KARPINSKA L., BOUZAROVSKI S. 2025. Impact of Energy Transitions on Energy Poverty in the European Union, *Renewable and Sustainable Energy Reviews*, Vol. 211, pp. 115311.

SY S. A., MOKADDEM L. 2022. Energy Poverty in Developing Countries: A Review of the Concept and its Measurements, *Energy Research and Social Science*, Vol. 89.

ZARGHAMI S. A. 2025. The Role of Economic Policies in Achieving Sustainable Development Goal 7: Insights from OECD and European Countries, *Applied Energy*, Vol. 377, pp. 124558.

Gloria POLINESI, Università Politecnica delle Marche, g.polinesi@staff.univpm.it
Maria Cristina RECCHIONI, Università Politecnica delle Marche, m.c.recchioni@univpm.it
Mariateresa CIOMMI, Università Politecnica delle Marche m.ciommi@univpm.it
Francesca MARIANI, Università Politecnica delle Marche, f.mariani@univpm.it
Alfonso CARFORA, Università di Macerata, alfonso.carfora@unimc.it

ENERGY POVERTY AND THE IMPACT OF ENERGY SUBSIDIES: RECENT TRENDS IN ITALY[1]

Elisabetta Segre, Paola Tanda

Abstract. In 2008, Italy introduced a system of means-tested energy subsidies which are deducted directly from electricity and gas bills. Since 2021, the deduction has been automatic for households with an ISEE certificate on file, and additional financial resources have been allocated to help households cope with the rise in energy prices. This has significantly increased both the number of households benefiting from the measure and the amount of the bonus. This analysis aims to assess recent trends in the impact of energy subsidies on reducing the number of energy-poor households.
The ISTAT's household microsimulation model, which is based on the IT-SILC survey matched with administrative data, allows us to estimate energy poverty at household level using the Low Income High Costs approach. In addition, the model allows us to identify which households received the subsidies and to estimate the amount of bonus received by each household.
Our findings show that energy subsidies have been effective in offsetting the impact of rising energy prices for energy-poor households. However, enhancing the targeting mechanisms could improve the effectiveness of the bonuses in reducing energy poverty.

1. Introduction

Energy poverty—defined as the condition in which households are unable to secure adequate energy services at an affordable cost[2]— is driven by a confluence of factors including energy inefficient housing, volatile energy prices, reliance on fossil fuels, and structural socio-economic disadvantages. Its consequences are profound, affecting physical and mental health outcomes, social inclusion, and

[1] The views and opinions expressed in this paper are those of the authors and do not necessarily reflect the official position of ISTAT.
[2] According to the Energy Efficiency Directive ((EU) 2023/1791) "energy poverty means a household's lack of access to essential energy services, where such services provide basic levels and decent standards of living and health, including adequate heating, hot water, cooling, lighting, and energy to power appliances, in the relevant national context, existing national social policy and other relevant national policies, caused by a combination of factors, including at least non-affordability, insufficient disposable income, high energy expenditure and poor energy efficiency of homes".

economic opportunities (Marmot Review Team 2011, Thomson *et al* 2017, European Commission and Cornelis, 2025).

At the European level, a suite of initiatives has been developed to tackle this issue like the Social Climate Fund and the Clean Energy for All Europeans legislative package and in its 2024 updated National Energy and Climate Plans (NECP), Italy acknowledges energy poverty as a major socio-political challenge. Several national initiatives have been implemented to mitigate the impact of rising energy costs on vulnerable populations: the strengthening of social energy bonuses; economic incentives to support the installation of solar PV systems for low-income families; and the promotion of Renewable Energy Communities (RECs); fiscal measures such as VAT reductions on energy bills and temporary removal of system charges.

In this paper, we focus our attention on energy subsidies (from here on social bonuses). This welfare measure have existed in Italy since 2008, providing discounts on electricity and gas bills for low-income households. The eligibility is based on the Equivalent Economic Situation Indicator (ISEE) and household composition. The ISEE certificate is issued by INPS (The National Institute for Social Security) upon demand. Since 2021, the system of social bonuses has been revised. First, the discount became automatic, households are no longer required to apply if ISEE data is already on file. Higher discounts for eligible households have been provided, partially funded by general tax revenues, and income thresholds have been raised several times.[3] Using ISTAT' microsimulation model (FaMiMod) we are able to identify beneficiaries of energy social bonuses.[4] This information combined with household characteristics like ISEE, size of the households, number of children, and climatic zone of residence, is used to assess the amount of the benefit a household is entitled to receive. [5]

Once beneficiaries and social bonuses are estimated at household level, we assess the impact of social bonuses on energy poverty.

[3] *Eligibility criteria*

	ISEE threshold	
	Less than 4 children	*More than 4 children*
2021	ISEE < 8.265	ISEE < 20.000
2022	ISEE < 12.000	ISEE < 20.000
2023	ISEE < 15.000	ISEE < 30.000
2024	ISEE < 9.530	ISEE < 20.000

[4] For details on the FaMiMod model refer to https://www.istat.it/it/files//2015/10/rsu_2_2015.pdf

[5] Less than 1 out of 2 households applies for an ISEE certificate, in 2023 they were 10.4 million out 26 million households (INPS, ISEE Observatory). To identify households receiving the bonus, a take-up rate was applied to the measure, in order to align the estimates of the number of beneficiaries with administrative information on number of bonuses paid (ARERA 2021, 2022, 2023).

Microsimulation models developed by Public Institutions are usually static and short-run oriented (Colombino 2016) and FaMiMod is no exception. The non behavioral assumption might induce a bias in our estimates, since a change in the price of energy induces a change in the level of energy demand. Nevertheless we have reasons to believe that the size of the bias might be negligible. First of all, we focus our analysis on the impact on energy poverty of an automatic discount in the bill. Unlike social tariffs, this kind of measure has a low or null impact on the price signal (Faiella e Lavecchia 2014). Another aspect that should be taken into account is that in Italy a high share of the energy costs are not linked to energy prices (Faiella and Lavecchia 2021).

An additional reason for us to believe that the bias could be negligible comes from the literature on the effect of energy prices dynamic on energy demand (see Priesmann and Praktiknjo 2025 for a review). Estimations of the price elasticities are rather heterogenous and inconsistent. Nevertheless, there is a strong consensus on the fact that energy demand is rather inelastic to price change in the short run (Espey and Espey 2004, Faiella and Lavecchia 2021). Since behavioral response need time to materialize (Colombino 2016), a short-run oriented policy assessment like ours should produce rather accurate estimates.

To the best of our knowledge, there are few examples in literature of behavioral microsimulation models used to study energy poverty (Tovar Reaños and Lynch 2022, Colabella *et al.* 2023) and we did not find evidence of behavioral model used to assess the effect of welfare measures on energy poverty. The Italian Ministry of Economy and Finance has an on-going, not yet published, work where a non behavioral micro-simulation model is used to evaluate the effect of energy inflation on energy expenditure and energy poverty (De Sario *et al.* 2025).

2. Data and methods

The measurement of energy poverty ranges from simple expenditure-based thresholds to complex, multidimensional frameworks that incorporate income, energy needs, housing quality, and subjective experience (for a review see Thomson *et al* 2017a, Gouveia *et al* 2022, Faiella and Lavecchia 2014, Tovar Reaños and Lynch 2022). Energy poverty accounts for several driving factors, summarized by Bouzarovski and Petrova (2017) in seven categories: access, affordability, flexibility, energy efficiency, need and cultural practices. Thomson *et al* (2017) classifies methods of measurement in three groups: expenditure approaches, which provide a proxy of energy deprivation by comparing actual energy costs to a threshold (absolute or relative); consensual approaches, which accounts for self-reported subjective measures such as the ability to afford an adequate level of heating or cooling; direct measurements, which assess the adequacy of energy services (like heating, cooling, lightning), for example, by taking the internal temperature of the dwelling.

We rely on the expenditure approach to assess the prevalence of energy poverty in Italy. In particular, we estimate the headcount of energy poor households using a Low Income High Costs (LIHC) type of measure. The LIHC measure defines a household as fuel poor if its required energy costs are above the national median and if its income net of the energy expenditure would fall below the poverty line (UK Department of Energy and Climate Change, 2013). To assess required energy costs, energy needs are modelled accounting for household composition, dwelling characteristics, and regional climatic conditions. This approach emphasized the structural dimensions of energy poverty, highlighting how factors like housing quality and energy efficiency play a role alongside income constraints.[6]

The same type of approach is used by the Italian Observatory for Energy Poverty (OIPE) to produce a measure of energy poverty often reported in official documents as the NECP. To assess energy poverty OIPE uses data from the Household Budget Survey carried out by ISTAT. This data source provides the official measure of energy expenditure faced by households according to the COICOP classification.

Nevertheless, in order to keep our analysis within the data environment of the micro-simulation model FaMiMod and thereby being able to carry out a policy evaluation analysis, we use an alternative data source. FaMiMod runs using the survey data collected by ISTAT in order to provide Eurostat with the set of Eu-Silc variables, matched with administrative data. The dataset contains detailed information on several housing expenditures including energy costs faced by the household[7]. This allows us to apply the LIHC approach, defining a household as energy poor when its energy expenditure exceeds the national median expenditure and its income, net of energy costs, falls below the "At risk of poverty line" (defined as 60% of the national median equivalized disposable income). According to Faiella and Lavecchia (2014), we added to this first group also At risk of poverty households reporting a null energy expenditure. As Faiella and Lavecchia, we are not able to estimate the household energy needs, as would be required by the LIHC method developed in the UK, due to lack of information on the energy efficiency of the

[6] In 2025, the UK Government introduced a revised framework - Low Income Low Energy Efficiency (LILEE) - that further refined the measurement of energy poverty. According to the LILEE metric, a household is considered energy poor if it has a low income (defined as below the poverty threshold after accounting for housing costs) and lives in a home with an energy efficiency rating of band D or below. This metric shifts the focus more explicitly toward energy efficiency, aligning energy poverty policy with broader environmental and decarbonization goals.

[7] Information on housing expenditures is collected in order to provide Eurostat with the Eu-Silc variable "Housing cost overburden rate". Energy expenditure driven from this source are overall coherent with official estimates coming from HBS.

dwelling. The headcount ratio of energy poverty (H), given the income (Y_i) and the energy expenditure (EE_i) of household, is obtained as follows:

$$H = \frac{1}{n}\sum_{i=1}^{n} E_i \tag{1}$$

where

$$E_i \begin{cases} 1 \text{ if } EE_i > \text{median } (EE) \text{ and } Y_i - EE_i < 0.6 \text{ median } (Y) \\ 1 \text{ if } EE_i = 0 \text{ and } Y_i < 0.6 \text{ median } (Y) \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

In our work EE_i is evaluated according to two scenarios. In the first one, we use energy expenditure as collected by the survey that is the actual cost faced by households. This cost is net of the energy bonuses received by a recipient household as a discount on the bills and allow us to estimate energy poverty after the subsidies (H_a) by applying equation (1) and (2).

In the second scenario, we calculate the cost a recipient household would have faced if it had not received the subsidy. We do that by adding the amount of the bonus to the actual energy cost and thereby obtaining the energy poverty headcount before the subsidies (H_b).[8]

Being our model a static non-behavioral one, the level of energy expenditure is held constant throughout the two scenarios, i.e. we assume that the policy does not affect the level of demand (see paragraph 2 for a discussion).

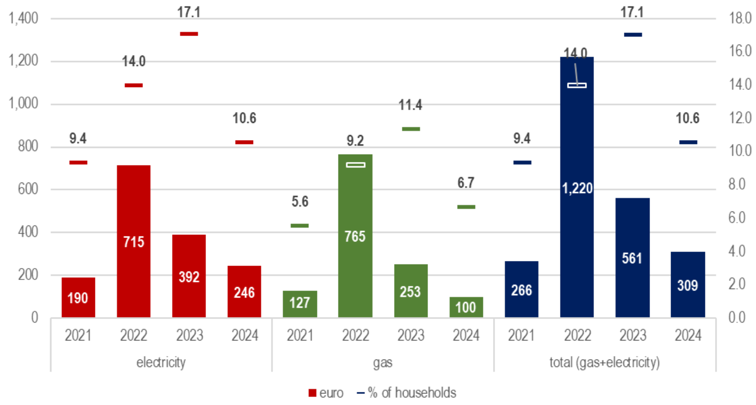
Although our estimates share the same theoretical approach as those performed by Faiella and Lavecchia and by the Italian Observatory on Energy poverty (OIPE), results should not be compared. Along with the fact that we use different data sources, the lack of comparability is due to important differences in equation (2). First of all, we compare EE_i with the median level of EE, while Faiella and Lavecchia and OIPE compare it with two times the average EE. Another difference relies on the fact that in the second condition of equation (2) we use data on disposable household income (Y_i) and the Eurostat' At risk of poverty threshold, whereas Faiella and Lavecchia and OIPE use data on the overall level of household expenditure (E_i) and their threshold is the average per-capita consumption expenditure equivalized by the number of household components.

[8] Households living in buildings with central heating system need to file a demand to obtain gas bonuses. Hence, for these households only, the information on the actual gas expenditure collected in the survey is not net of the value of the bonus. The two scenarios are computed accordingly.

3. Energy subsidies in Italy: recipients and amount

In 2021, almost 1 out of 10 households received a discount on the electricity bill of €190 on average, and 60% of these households received also a cut in the gas bill of €127(Figure 1). The average total amount of energy bonuses is estimated at €266. As mentioned above, the system became more generous to help household face the increase in energy prices and, in 2022 the average total amount peaked to €1,220, reaching 14.0% of the total resident households. This share kept growing in 2023, reaching its highest level (17.1%) thanks to the extension of the ISEE threshold to €15,000. In 2024, the ISEE threshold has been set back to 9.530 euro, resulting in a significant reduction in the proportion of eligible recipient households.

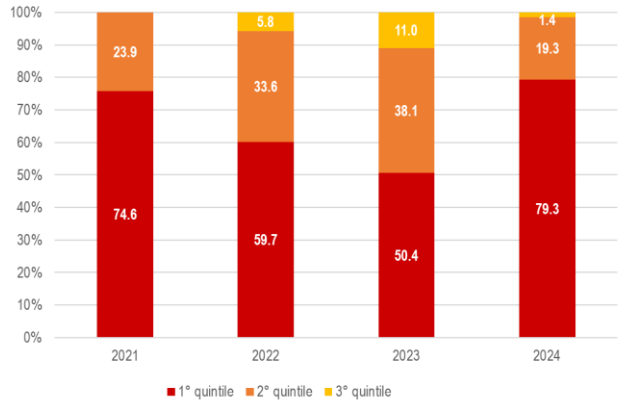
Figure 1 – *Household-based energy subsidies by type – Years 2021-2024 (euro and share of households).*



Source: Estimates based on Istat' microsimulation model FaMiMod.

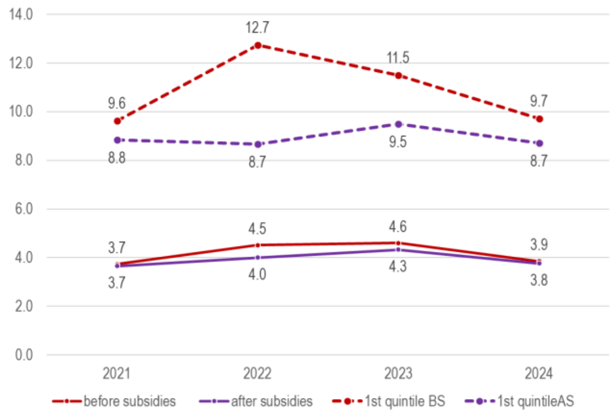
Energy subsidies target the poorest households. Over the 2021-2024 period, approximately 94% of beneficiary households belonged to the bottom two income quintiles, receiving more than 93% of the total subsidy expenditure (Figure 2). Starting from 2022, thanks to the increase in the ISEE threshold, among the beneficiaries we find a small share of households belonging to the central quintile (5.8% in 2022, 11.0% in 2023). This share decreased substantially in 2024, when the ISEE threshold has been set back to €9.530.

Figure 2 – *Household-based energy subsidies by disposable income quintile - Years 2021-2024 (composition).*



Source: Estimates based on the Istat'household microsimulation model (FaMiMod).

Figure 3 *Equivalent energy expenditure as a share of equivalent disposable income before and after energy subsidies – Years 2021-2024 (% of total household expenditure).*



Source: Estimates based on Istat' microsimulation model FaMiMod.

Energy subsidies helped offset the impact of rising energy prices on poor households. Figure 3 shows that, without the subsidies, the average share of equivalent energy expenditure over the first quintile household equivalent disposable income would have grown in 2022 up to 12.7%. Thanks to the generous bonus system in place that year, the share decreased to 8.7%, slightly lower than the share observed in 2021 (8.8%).

4. The impact of energy subsidies on energy poverty

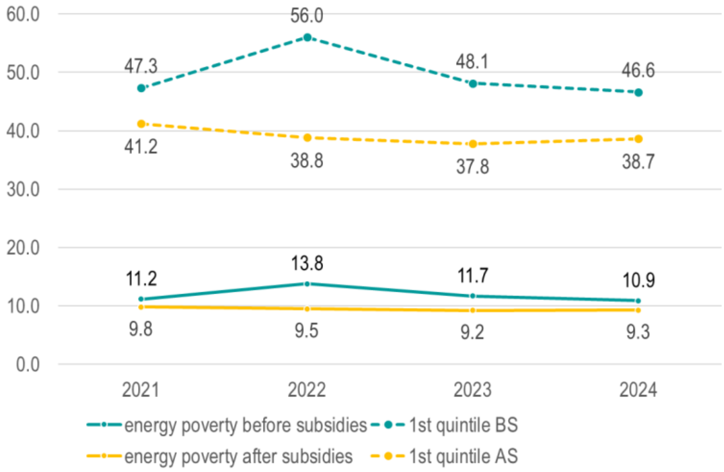
As explained in paragraph 2, to assess the impact of social bonuses between 2021 and 2024, we identify households in energy poverty according to two scenarios: one with bonuses (H_a) and one without (H_b). The latter was constructed by adding the estimated value of the bonuses to actual household expenditure. This provides an estimate of the energy expenditure the households would have incurred in the absence of such measures.

Simulation results show that in 2021, 11.2% of households in Italy were in energy poverty before social bonuses were paid (Figure 4). This percentage fell to 9.8% after the bonuses were received, a decrease of 0.4 percentage points. The impact of the bonus energy poverty reached a peak in 2022. With the rise in energy prices, the share of households in energy poverty before subsidies H_b increased to 13.8%. Yet, the amount of the bonuses was so high in 2022, that the resulting incidence of households in energy poverty after subsidies, 9.5%, is lower than the one observed in 2021. In 2023, 11.7% of households experienced energy poverty before subsidies. This figure decline to 9.2% afterwards (-2.5%). In 2024, social bonuses helped reduce energy poverty by 1.6 percentage points (from 10.9 to 9.3%).[9]

Households in the first quintile of the income distribution are more likely to be energy poor. In 2022 energy poverty before energy subsidies reached a peak of 56.0% among the poorest households, moving from the 47.3% observed in 2021. Again, energy subsidies helped off-setting the impact of the peak in energy prices. Without the social bonus measure, energy poverty would have increased by 8.7 p.p. Instead, once social bonuses are accounted for, the share of families experiencing energy poverty is lower in 2022 than in 2021 by 6 p.p. (41.2 vs 38.8%).

[9] The time series of energy poverty after social bonuses estimated in this work differs from the one provided by OIPE (OIPE, 2024). In particular, we observe a rather steady trend while OIPE's estimates show a sharp increase between 2022 and 2023 (from 7.7 to 9.0). Nevertheless, estimates are not comparable, refer to paragraph 2 for details.

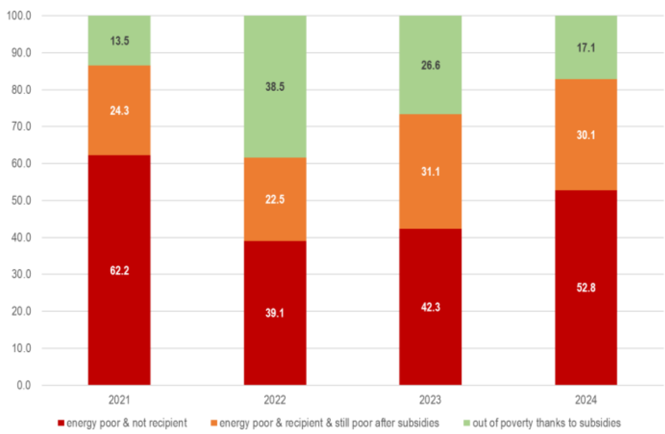
Figure 4 *Energy poverty before and after energy subsidies - Years 2021-2024 (% of households).*



Source: Estimates based on Istat' microsimulation model FaMiMod.

Focusing on households below the energy poverty threshold in 2021, 62.2% did not receive the bonuses, 24.3% received bonuses while remaining in energy poverty, and 13.5% escaped poverty thanks to the subsidies (Fig. 5). Target efficiency improved in 2022, when energy poor households not reached by the social bonus decreased to 39.1% and 38.5% of beneficiary households escaped energy poverty.

Figure 5 *Target efficiency of energy subsidies - Years 2021-2024 (composition).*



Source: Estimates based on Istat' microsimulation model FaMiMod.

In 2023, when the number of beneficiary households increased and the amount granted became less generous, we observe a rise in the share of households receiving the social bonus and yet remaining energy poor (31.1%). In 2024, more than one energy poor household out of 2 is not among the beneficiaries, and we observe the highest share of households receiving the bonus yet remaining poor.

An energy-poor family may not be eligible for the subsidy because it does not have a valid ISEE certificate, or an electricity account registered in the name of a family member (for example, the family's electricity account is registered in the name of the homeowner), or has an ISEE higher than the threshold to be eligible for the bonus.

The model allows studying subsidy effectiveness by breakdown variables such as household size, geographical and climate area. These more detailed analysis are a priority for future research.

5. Conclusions

Between 2021 and 2024, Italy's reform of the social energy bonuses helped significantly containing the rise in energy poverty, particularly during the surge in energy prices in 2022. Using a Low Income High Costs (LIHC) approach and ISTAT's FaMiMod microsimulation model, we examined the effectiveness of these subsidies in reducing energy poverty.

The results confirm the role of subsidies in mitigating the impact of energy price shocks and their long-term effects on energy poverty. Without social bonuses, the proportion of households experiencing energy poverty would have increased significantly in 2022. However, the same year saw a proportion of energy-poor households after subsidies that was lower than that observed in 2021, which more than offset the impact of rising energy prices. A fairly large mitigating effect was also observed in 2023.

Overall, the social energy bonus system has been effective in mitigating the impact of rising energy costs on household welfare. In terms of targeting efficiency the largest proportion of households lifted out of energy poverty thanks to the bonuses is observed in 2022. Enhancing the targeting mechanisms could improve the system's effectiveness in reducing energy poverty.

Appendix

The ISTAT's Microsimulation Model: FaMiMod is based on administrative data from the Ministry of Finance, matched to ISTAT survey data from EU-SILC.

Although the model is static, it is regularly updated to the most recent year by: 1) projecting monetary variables forward using either National Accounts or MeMo-It forecasts; 2) reweighting the survey sample based on the most recent populations breakdown by age, sex, and employment status (i.e. employed, dependent, self-employed or unemployed), and 3) updating the model's legislative framework to ensure the baseline accurately reflects current legislation. Once updated, the model can simulate the effects of new policies by comparing income level and income distribution under different scenarios (baseline vs reform or alternative scenario). For more details see https://www.istat.it/it/files//2015/10/rsu_2_2015.pdf.

Acknowledgements

Authors would like to thank the anonymous referee for useful comments and suggestions.

References

ARERA. 2021, 2022, and 2023. *I bonus sociali elettrico e gas: stato di attuazione*. Annual report by the Italian Regulatory Authority for Energy Networks and Environment.

COLABELLA A., LAVECCHIA L., MICHELANGELI V., PICO R. 2023. To eat or to heat: are energy bills squeezing people's spending? *Occasional Papers* , No. 800, Bank of Italy.

COLOMBINO U. 2015. The Istat Microsimulation Models. *Rivista di statistica ufficiale* Vol. 2, pp. 5-15. https://www.istat.it/it/files//2015/10/rsu_2_2015.pdf

DE SARIO G., IMPERIOLI L., CASTALDI G., CIRILLO C. 2025. The effect of energy inflation on energy expenditure and energy poverty: evidence from microsimulations in ECOP model. Work presented at the *Bank of Italy Workshop on Microsimulation Modelling* 4th of July 2025.

ESPEY J. A., ESPEY M. 2004. Turning on the Lights: A Meta-Analysis of Residential Electricity Demand Elasticities. *Journal of Agricultural and Applied Economics* Vol. 36 , No. 1, pp. 1–17.

EUROPEAN COMMISSION: DG FOR ENERGY, CORNELIS M. 2025. *Framing summer energy poverty – Insights and recommendations for a resilient future – Final report*. Publications Office of the European Union.

FAIELLA I., LAVECCHIA L. 2014. La povertà energetica in Italia. *Occasional Papers,* No. 240, Bank of Italy.

FAIELLA I., LAVECCHIA L. 2021. Households' energy demand and the effects of carbon pricing in Italy. *Occasional Papers*, No. 614, Bank of Italy.

GOUVEIA J.P., PALMA P., BESSA S., MAHONEY K., SEQUEIRA M. 2022. *Energy Poverty National Indicators: Insights for a more effective measuring.* Published by the Energy Poverty Advisory Hub of the European Commission.

HILLS J. 2012. Getting the Measure of Fuel Poverty. Final report of the Hills Independent Fuel Poverty Review. *Centre for analysis of social exclusion Report* No. 72.

MARMOT REVIEW TEAM. 2011. *The health impacts of cold homes and fuel poverty*. Friends of the Earth & Marmot Review.

OIPE. 2024. La povertà energetica in Italia nel 2023. *Osservatorio italiano sulla povertà energetica*, Centro Levi-Cases, Università di Padova

PRIESMANN J., PRAKTIKNJO A. 2025. Estimating short- and long-run price and income elasticities of final energy demand as a function of household income. *Energy Policy*, Vol. 207, 114850.

THOMSON H., BOUZAROVSKI S., SNELL C. 2017a. Rethinking the measurement of energy poverty in Europe: A critical analysis of indicators and data. *Indoor and Built Environment*, Vol. 26, No. 7, pp. 879-901.

THOMSON H., SNELL C., BOUZAROVSKI S. 2017b. Health, well-being and energy poverty in Europe: A comparative study of 32 European countries. *International Journal of Environmental Research and Public Health*, Vol.14, No. 6, p. 584.

TOVAR REAÑOS M. A., LYNCH M. A. 2022. Are fuel poverty metrics fit for purpose? An assessment using behavioural microsimulation. *Sustainable Cities and Society*, No. 81, 103817.

UK DEPARTMENT OF ENERGY AND CLIMATE CHANGE. 2013. *Fuel Poverty: a Framework for Future Action*.

Elisabetta SEGRE, Istat, elisabetta.segre@istat.it
Paola TANDA, Istat, paola.tanda@istat.it

POVERTY, INEQUALITIES AND EDUCATIONAL OUTCOMES: A TERRITORIAL ANALYSIS IN THE AGE OF TRANSITIONS[1]

Simona Cafieri, Gianmarco Borrata, Manuela Barba, Paola Bianco

Abstract. This paper provides a multidimensional spatial analysis of educational inequality in Italy by integrating regional data (2019–2023) with K-means clustering and Geographically Weighted Regression. The results show that structural poverty, digital exclusion, youth vulnerability, and parental educational background produce markedly uneven literacy and numeracy outcomes, with the magnitude and direction of these effects varying across space. By examining how demographic, digital, and ecological transitions interact with existing vulnerabilities, the study identifies emerging territorial pressures that intensify educational disparities. Overall, the findings offer updated evidence to support more targeted, place-based strategies aimed at reducing persistent and evolving forms of educational disadvantage.

1. Introduction

In Italy, territorial disparities remain among the most persistent forms of social inequality. Education—one of the main channels of upward mobility—is strongly shaped by structural and territorial imbalances, as clearly revealed during the COVID-19 pandemic. The transition to distance learning exposed how unequal access to infrastructure and resources constrained educational opportunities in disadvantaged areas, further widening existing divides *(Borgonovi and Ferrara, 2023).*

While prior research has documented regional educational inequalities, limited attention has been devoted to how structural vulnerabilities and transition-related processes jointly produce spatially heterogeneous effects. This study addresses this gap by analysing how demographic, digital, and ecological transitions interact with multidimensional forms of disadvantage to shape literacy and numeracy outcomes *(Cantalini et al., 2025).*

The analysis is conducted at the regional level, the most appropriate scale for integrating harmonised socioeconomic, digital, and educational indicators and for capturing the cumulative effects associated with long-term structural transitions.

[1] This study represents the result of a collaborative effort between the authors.

While the abstract emphasises the need to move beyond the traditional North–South divide, this should not be interpreted as an attempt to analyse sub-regional patterns such as centre–periphery or urban–rural gradients, which cannot be adequately captured with available data. Rather, the paper advances beyond the binary macro-regional classification by adopting a multidimensional and spatially sensitive regional perspective that uncovers differentiated mechanisms within and across regions.

Although territorial inequalities in education represent a well-established field of research in Italy, existing analyses have predominantly relied on macro-regional comparisons or global regression models that implicitly assume spatial stationarity. Such approaches tend to overlook how the determinants of educational disadvantage vary across space and how structural vulnerabilities interact with transition-related pressures in differentiated ways. This study advances the literature by integrating a multidimensional framework of territorial vulnerability with spatially explicit modelling at the regional scale. By combining cluster analysis with Geographically Weighted Regression, the paper reveals mechanisms that remain hidden in conventional models and shows that socioeconomic, digital and demographic factors do not exert uniform effects across Italian regions. In doing so, it offers updated evidence on the geography of educational disadvantage in the post-pandemic and transition-oriented context and provides a more nuanced understanding of how cumulative vulnerabilities shape regional outcomes. This perspective contributes to bridging the gap between descriptive territorial disparities and the need for analytically grounded, place-sensitive policy insights.

2. Background

The understanding of poverty has evolved from a narrow focus on income deprivation to a multidimensional perspective. Contemporary frameworks emphasise the interplay of cognitive, social, and cultural dimensions in shaping educational trajectories *(Spicker et al., 2008).* In this view, educational poverty refers to limited access to quality learning environments, essential skills, and opportunities for cognitive development. Educational disadvantage is not solely rooted in individual or family conditions but is also shaped by contextual and territorial factors *(OECD, 201*2). In Italy, persistent spatial inequalities have produced regional clusters of deprivation where economic, social, and institutional vulnerabilities intersect and reinforce one another *(Ballarino, 2009).*

Despite extensive research on regional disparities in Italy, existing studies rarely adopt a multidimensional and transition-oriented perspective, nor do they examine how the digital, demographic, and ecological transitions translate into spatially differentiated educational outcomes. This leaves a gap in understanding the mechanisms through which structural territorial vulnerabilities shape learning inequalities—a gap that the present study addresses by integrating these transitions into a spatial analytical framework.

Recent literature highlights the need to interpret educational inequalities in light of three major transitions currently reshaping territorial dynamics: digital, demographic, and ecological *(Parsons et al., 2024; Rosário and Disa, 2022)*.

• Digital transition: limited broadband coverage, low device availability, and weak digital skills constrain access to technology-enhanced learning and exacerbate existing gaps.

• Demographic transition: population ageing and shrinking youth cohorts reduce school network density and continuity of provision *(Muttarak and Lutz, 2014)*, particularly in shrinking regions.

• Ecological transition: exposure to environmental risks—such as heatwaves or hydrogeological instability—affects the resilience and quality of school infrastructure and disrupts learning conditions.

In this perspective, each transition is expected to generate territorially differentiated impacts: digital constraints amplify learning gaps most acutely in low-connectivity regions; demographic decline heightens educational vulnerability in ageing and shrinking areas where school provision becomes fragile; and ecological pressures disproportionately affect territories exposed to environmental risks that compromise the continuity and quality of learning conditions. Taken together, these transitions do not simply coexist but operate as structural amplifiers of territorial inequality, reinforcing context-specific mechanisms that shape geographically differentiated educational outcomes.

By integrating the three structural transitions into a spatially sensitive analytical framework and combining cluster analysis, OLS, and GWR, this study moves beyond traditional accounts of the North–South divide and provides a novel multidimensional reading of educational inequality in Italy. This approach not only identifies territorial disparities but explains the mechanisms through which digital, demographic, and ecological vulnerabilities translate into regionally differentiated educational outcomes—an aspect largely overlooked in previous research.

3. Data and methodology

The empirical analysis integrates data from a broad range of official statistical sources for the period 2019–2023. We combined EU-SILC *(Istat, 2022–2023)*, LFS *(Istat, 2023)*, INVALSI and PNRR information to capture socioeconomic, educational and policy dimensions across Italian regions. The regional scale allows the integration of harmonised indicators and provides sufficient territorial variation to analyse spatial heterogeneity through GWR, even though it does not permit an explicit investigation of urban–rural or centre–periphery gradients.

The methodological strategy comprises two steps. First, we conduct a descriptive territorial analysis and apply K-means clustering to classify regions according to structural determinants, using INVALSI proficiency distributions and indicators of socioeconomic and digital vulnerability. The cluster typologies serve as a structural lens to contextualise the subsequent regression and GWR results, helping to interpret how different configurations of vulnerability shape spatially varying relationships between socioeconomic factors and educational outcomes.

In the second step, we estimate the relationship between educational outcomes and socioeconomic conditions using multiple linear regression models. Since these models do not account for spatial dependence, we adopt GWR , following recent contributions *(Sacco and Falzetti, 2021).* GWR allows the strength and direction of associations to vary across space, highlighting mechanisms that remain hidden in global models.

This combined exploratory–explanatory approach enables both the identification of structural regional profiles and the analysis of spatially heterogeneous relationships between vulnerabilities and educational outcomes. "Structural determinants" refer here to persistent socioeconomic and educational characteristics—such as youth vulnerability, digital skills, household income and parental education—selected in line with established literature on territorial inequalities.

3.1. Cluster Analysis of Regional Education Performance

To classify regional heterogeneity, we applied K-means clustering, using the regional distributions of student performance in INVALSI literacy and numeracy assessments for 2019 and 2023, grouped into five ordered proficiency levels across primary and secondary school grades. The elbow method indicated a three-cluster solution, supported by empirical interpretability and consistency with well-documented territorial disparities.

The clustering relies on the following INVALSI indicators (regional level) for years: 2019, 2023:

1) Literacy proficiency distributions (5 ordered levels: Grades 2 and 5 primary; Grades 3 and 5 secondary);
2) Numeracy proficiency distributions (5 ordered levels: Grades 2 and 5 primary; Grades 3 and 5 secondary).

The clustering captures structural dimensions that the literature consistently associates with territorial inequalities, including socioeconomic fragility, digital capital, household resources and parental education. These dimensions shape cumulative learning opportunities and help identify stable regional profiles of disadvantage. Rather than explaining educational outcomes directly, the cluster typology provides a theoretically grounded framework for interpreting the subsequent regression and GWR analyses. In particular, it offers a structural lens through which to understand how different configurations of vulnerability relate to the spatially varying relationships revealed by the GWR.

3.2. Cluster Composition and Interpretation

The cluster analysis identifies three distinct regional profiles, reflecting well-established territorial divides in Italy (Figures 1, and 2).

The first group includes regions characterised by stronger socioeconomic and educational endowments; the second captures intermediate and internally heterogeneous contexts; and the third comprises structurally disadvantaged regions with persistently lower proficiency levels. These profiles provide the structural backdrop for interpreting the regression and GWR results, highlighting how different combinations of vulnerabilities underpin the spatially varying mechanisms identified in the subsequent analysis.

Figure 1 – *Cluster of Italian regions for 2019: inadequate numerical and literacy competence.*

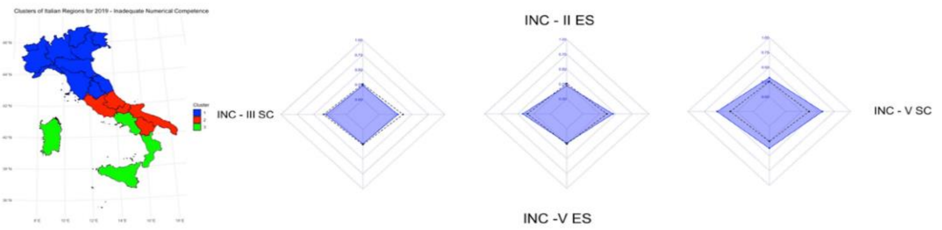
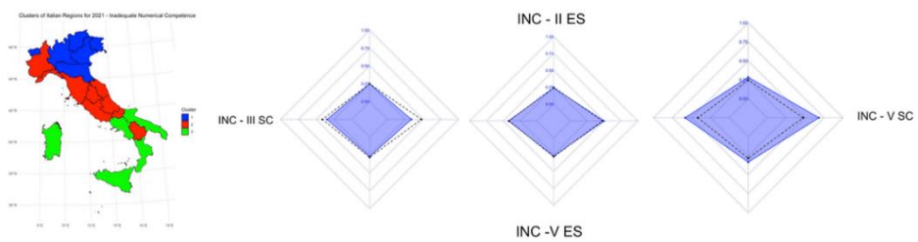


Figure 2 – *Cluster of Italian regions for 2021: inadequate numerical and literacy competence.*



3.3. Cluster Results to Territorial Patterns of Educational Disadvantage

The three clusters identified in Section 3.2 reveal distinct territorial patterns of educational disadvantage, distinguishing regions marked by persistent structural vulnerabilities from those characterised by more favourable socioeconomic and educational conditions. These profiles reflect long-standing territorial divides and illustrate how cumulative disadvantages shape regional differences in literacy and numeracy outcomes.

By linking the cluster results to the regional socioeconomic landscape, it becomes possible to interpret the spatial variation observed in the regression and GWR analyses. The clusters highlight how different territorial endowments condition the strength and direction of key determinants, thereby providing a coherent framework for understanding the spatially heterogeneous mechanisms identified through GWR.

4. Determinants of education deprivation: Geographically Weighted Regression

To examine how structural vulnerabilities relate to literacy and numeracy outcomes across space, we complement the cluster analysis with a set of Geographically Weighted Regression (GWR) models. Table 1 summarises the socioeconomic indicators used in the empirical analysis. The four dependent variables (Y1–Y4) reflect inadequate literacy and numeracy competences at the end of primary and upper secondary school.

The explanatory variables capture key dimensions of territorial vulnerability: digital proficiency (X1); youth disadvantage, including NEET rates (X2) and early school leaving (X3); household economic resources (X4); and school service provision (X5). Additional indicators (X6–X9) represent parental educational composition.

GWR models are estimated for each dependent variable to assess how the strength and direction of these relationships vary across regions.

GWR is used to capture spatial heterogeneity in the relationship between literacy and numeracy outcomes and their socioeconomic determinants, allowing coefficients to vary across regions *(Brunsdon et al., 1998; Sacco et al., 2021).* This approach complements the global regression by identifying territorial differences in the magnitude and direction of key associations.The spatial analyses have been performed using the R package spdep[2] e GWmodel[3].

[2] https://cran.r-project.org/web/packages/spdep/index.html
[3] https://cran.r-project.org/web/packages/GWmodel/index.html

Table 1 – *Socioeconomic Indicators used in the analysis.*

Variable	Indicator	Definition
Y1	Low literacy – Grade 5 (Primary)	Percentage of students in 5th grade of primary school with inadequate literacy skills
Y2	Low literacy – Grade 5 (Upper Secondary)	Percentage of students in final year of upper secondary school with inadequate literacy skills
Y3	Low numeracy – Grade 5 (Primary)	Percentage of students in 5th grade of primary school with inadequate numeracy skills
Y4	Low numeracy – Grade 5 (Upper Secondary)	Percentage of students in final year of upper secondary school with inadequate numeracy skills
X1	Digital skills	Share of individuals with at least basic digital skills
X2	NEET	Percentage of youth aged 15–24 not in employment, education or training
X3	Early leavers from education and training	Percentage of 18–24 year-olds with at most lower secondary education who are not in further education or training
X4	Average household income	Mean income per household
X5	Schools offering basic services	Percentage of schools providing basic infrastructure (internet, labs, accessibility, etc.)
X6	Husbands: upper secondary – Wives: upper secondary	Share of couples where both spouses have upper secondary education
X7	Husbands: tertiary – Wives: upper secondary	Share of couples where husbands have tertiary and wives upper secondary education
X8	Husbands: tertiary – Wives: tertiary	Share of couples where both spouses have tertiary education
X9	Husbands: upper secondary – Wives: tertiary	Share of couples where husbands have upper secondary and wives tertiary education

4.1. Results of Geographically Weighted Regression

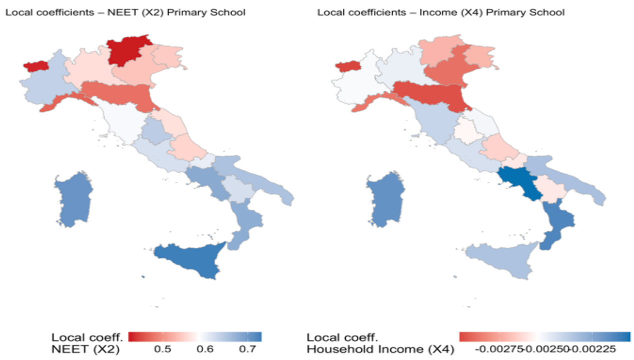
Table 2 presents the GWR estimates for inadequate literacy competence in Grade 5 of primary school. To support the interpretation of spatial heterogeneity, we include maps illustrating the geographical variation of selected coefficients. These visualisations show how the effects of specific determinants differ across regions, highlighting territorial patterns concealed in global models.

Household income (X4) is the only predictor consistently significant across all regions, showing a uniformly negative association with inadequate literacy outcomes. Early school leaving (X3) and the provision of basic school services (X5) also display negative effects, though their significance varies territorially (65% and 55% of regions, respectively). Digital skills (X1) and NEET rates (X2) exert more localised influences, emerging as significant only in specific areas, while parental education variables (X6–X9) show generally weak associations.

Table 2 – *Results of the geographically weighted regression (GWR) model for Inadequate Literacy Competence - Grade 5 (Primary School, 2021).*

	Coefficient Range						
	Min	1st Q	Mean	Median	3rd Q	Max	% Sig. Coef.
Intercept	-24.72	-22.31	-20.29	-19.47	-18.19	-16.98	85
X1	0.72	0.79	0.84	0.85	0.91	1.01	45
X2	0.42	0.55	0.60	0.61	0.66	0.74	30
X3	-0.34	-0.29	-0.26	-0.26	-0.21	-0.15	65
X4	-0.003	-0.003	-0.003	-0.003	-0.002	-0.002	100
X5	-0.52	-0.46	-0.45	-0.44	-0.40	-0.34	55
X6	-0.38	-0.30	-0.28	-0.27	-0.21	-0.18	20
X7	-0.15	-0.08	0.02	0.01	0.09	0.19	10
X8	-1.80	-1.72	-1.70	-1.68	-1.64	-1.60	0
X9	0.78	0.85	0.90	0.92	0.95	1.02	25
R^2	0.85	0.87	0.88	0.88	0.89	0.91	
	$R^2\ Global\ =\ 0.88$				$R^2\ OLS\ =\ 0.81$		

Figure 3 – *Local GWR coefficients for NEET (X2) and Household Income (X4) – Literacy Grade 5 (Primary School, 2021).*



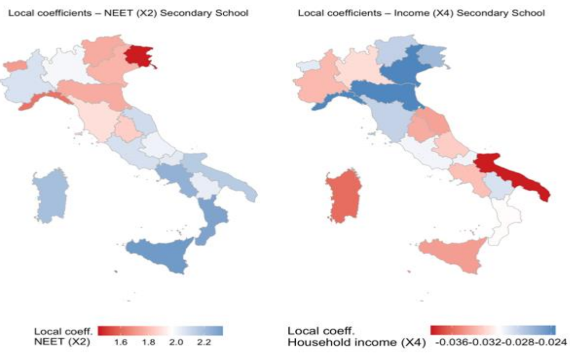
To assess robustness, we performed diagnostic checks: the GWR model outperformed the global OLS specification (lower AICc), Moran's I on residuals confirmed the absence of remaining spatial autocorrelation, and alternative bandwidth selection procedures yielded comparable coefficient surfaces, supporting the stability of the spatial patterns identified.

Table 3 presents the results of the GWR model estimating the determinants of inadequate literacy competence among students in the Grade 5 of secondary school.

Table 3 – *Results of the geographically weighted regression (GWR) model for Inadequate Literacy Competence - Grade 5 (Secondary School, 2021).*

	Coefficient Range						
	Min	1st Q	Mean	Median	3rd Q	Max	% Sig. Coef.
Intercept	65.2	70.8	74.5	74.1	78	82.6	80
X1	-1.85	-1.52	-1.39	-1.37	-1.21	-0.95	90
X2	1.42	1.70	1.94	1.95	2.15	2.36	100
X3	1.12	1.39	1.59	1.58	1.73	2.01	65
X4	-0.04	-0.036	-0.031	-0.030	-0.027	-0.022	100
X5	-1.54	-1.34	-1.22	-1.21	-1.12	-0.98	70
X6	-1.12	-0.95	-0.87	-0.86	-0.75	-0.63	35
X7	-2.35	-2.05	-1.88	-1.87	-1.69	-1.45	40
X8	-1.95	-1.66	-1.52	-1.51	-1.39	-1.20	50
X9	-1.10	-0.87	-0.74	-0.75	-0.62	-0.45	25
R^2	0.88	0.90	0.92	0.92	0.93	0.95	
	$R^2\ Global\ =\ 0.93$					$R^2\ OLS\ =\ 0.92$	

Figure 4 – *Local GWR coefficients for NEET (X2) and Household Income (X4) – Literacy Grade 5 (Secondary School, 2021).*



Compared to the global OLS estimates, the GWR results reveal marked spatial variation in several relationships. Predictors that appear weak in the global specification, such as digital skills and school service provision, display pronounced effects in some regions, whereas household income emerges as a consistently influential factor. These patterns show that global coefficients mask relevant territorial differences, underscoring the value of adopting a spatially explicit modelling strategy.

5. Education Policy and Public Investment

The discussion Public investment in education, including PNRR resources, has the potential to mitigate the territorial vulnerabilities highlighted by the empirical analysis. However, the current allocation of funds does not consistently reflect the spatial patterns identified by the GWR model. Regions where inadequate literacy outcomes are more strongly driven by digital constraints, socioeconomic disadvantage, or demographic decline do not always correspond to those receiving proportionally greater investment in digital infrastructure, school network consolidation, or ecological adaptation.

The findings suggest that PNRR measures could be more effective if aligned with the locally varying determinants of educational disadvantage. A more spatially targeted approach—sensitive to the mechanisms revealed by cluster analysis and GWR—would enhance the capacity of public investment to reduce structural inequalities rather than address needs in a uniform manner.

6. Discussions and conclusions

This study contributes to the analysis of territorial inequalities in Italy by integrating a multidimensional framework of vulnerability with a geographically weighted modelling strategy. By linking structural poverty, digital exclusion and youth disadvantage to spatially varying educational outcomes, it provides new evidence on mechanisms that remain invisible in conventional regional or macro-regional analyses. This spatial lens shows that inequalities are driven by locally differentiated processes rather than by uniform national patterns.

The empirical results confirm that the determinants of educational disadvantage vary substantially across space. These findings advance existing research by demonstrating pronounced spatial non-stationarity—an aspect largely overlooked in previous studies—and by clarifying how multidimensional vulnerabilities interact with demographic, digital and ecological transitions at the regional level.

The cluster analysis strengthens this interpretation by identifying distinct territorial profiles. Regions with strong socio-institutional endowments exhibit more resilient educational outcomes, whereas structurally fragile territories reveal compounded disadvantages that mirror the spatial patterns captured by the GWR. Rather than serving as explanatory variables, the clusters provide a structural lens through which the localised regression results can be interpreted, clarifying how different configurations of vulnerabilities shape spatially differentiated educational outcomes.

The policy implications follow directly from these spatially heterogeneous mechanisms. Regions in Cluster 1, where vulnerabilities are limited, may prioritise innovation and the consolidation of effective practices. Cluster 2 regions require targeted interventions to address internal fragmentation and mixed disadvantage profiles. Cluster 3 regions—marked by persistent deprivation and low performance—need long-term, integrated strategies combining income support, community-based programmes and investments in educational infrastructure. The GWR results underscore that policies designed without consideration of geographical variation risk misallocating resources, as identical measures may have markedly different effects depending on local conditions.

Overall, the study demonstrates that the geography of educational disadvantage is shaped by non-stationary mechanisms that can only be detected through spatially explicit models. Despite the constraints of regional data, the analysis provides updated insights into how structural vulnerabilities interact with demographic, digital and ecological transitions, offering a timely contribution to the post-pandemic and transition-oriented policy agenda.

Finally, the analysis points to several avenues for future research. Structural causal models and directed acyclic graphs (DAGs) could strengthen variable selection and reduce risks of collider bias. Additional factors—such as teacher quality, governance structures or school-level practices—could not be included due to data limitations but may play a substantive role in shaping territorial outcomes. Extending the framework to micro-data or panel designs would allow for a more precise identification of dynamic mechanisms. Evaluating the territorial effects of PNRR-funded interventions and examining how demographic, digital and ecological transitions reshape educational opportunities will be essential for informing future place-sensitive policies.

References

BALLARINO G., BERNARDI F., REQUENA M., SCHADEE H. 2009. Persistent inequalities? Expansion of education and class inequality in Italy and Spain. *European Sociological Review*, Vol. 25, No. 1, pp.123-138.

BORGONOVI F., FERRARA A. 2023. COVID-19 and inequalities in Educational achievement in Italy. *Research in social stratification and mobility*, Vol. 83:100760.

BRUNSDON C., FOTHERINGHAM S., CHARLTON M. 1998. Geographically weighted regression. *Journal of the Royal Statistical Society*, pp. 431-443.

CANTALINI S., PANICHELLA N., PIETROLUCCI A., TRIVENTI M. 2025 Regional disparities, geographical marginality, and educational pathways: A study on upper secondary education in Italy. *Social Inclusion*, Vol. 13.

INVALSI 2024. *Rilevazioni Nazionali degli Apprendimenti*. Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e Formazione.

ISTAT 2022-2023 EU-SILC – Statistics on Income and Living Conditions. Roma: Istituto Nazionale di Statistica.

ISTAT 2023. Labour Force Survey – LFS. Roma: Istituto Nazionale di Statistica.

ITALIADOMANI 2025. *PNRR Open Data Portal*. https://www.italiadomani.gov.it

MEF 2023. Ripartizione territoriale delle risorse del PNRR.

MUTTARAK R., LUTZ W. 2014. Is education a key to reducing vulnerability to natural disasters and hence climate change? *Ecology and society*, Vol. 19, No. 1.

OECD 2012. *Equity and quality in education: Supporting disadvantaged students.* Paris: Organisation for Economic Co-operation and Development.

OPENCOESIONE 2025. *Portale Open Data – Politiche di Coesione*. https://opencoesione.gov.it

ROSARIO A. T., DISA J. C. 2022. Sustainability and the digital transition: A literature review. *Sustainability*, Vol. 14, No. 7, pp. 4072.

SACCO C., FALZETTI P. 2021. Spatial variations of school-level determinants of reading achievement. *Large-scale Assessments in Education*, Vol. 9, No.1, No.12.

PARSONS E. S., JOWELL A., VEIDIS E., BARRY M., ISRANI S. T. 2024. Climate change and inequality. *Pediatric research*, pp. 1-8.

SPICKER P., LEGUIZAMON S. Á., GORDON D. (Eds.). 2008. *Poverty: an international glossary*. London: Bloomsbury Publishing.

Simona CAFIERI, Istat, simona.cafieri@istat.it
Gianmarco BORRATA, Università Federico II, gianmarco.borrata@unina.it
Manuela BARBA, Istat, manuela.barba@istat.it
Paola BIANCO, Istat, paola.bianco@istat.it

IN-WORK POVERTY AT THE LOCAL LEVEL: GENDER INEQUALITIES AND EMPIRICAL EVIDENCE FROM ADMINISTRATIVE SOURCES

Gianluca Truscello, Maria Chiara Zanarotti

Abstract. This study investigates in-work poverty within the local context of an Italy's province with two main objectives: (1) assessing whether the main individual and household-level risk factors identified by national and international research are also valid at subnational level, even in contexts where poverty is relatively less widespread; (2) evaluating the feasibility and adequacy of using administrative data sources to study socio-economic vulnerability. The broader goal is to contribute to the understanding of in-work poverty by exploring its determinants within a localized context and by testing the potential and added value of existing data infrastructures. The analysis focuses on the province of Reggio Emilia, using microdata from CAF-CGIL users. While the dataset is not statistically representative and lacks certain key variables - such as working hours, education, and industry - nonetheless it offers a rich, replicable data resource for local-level research. Empirical findings confirm the relevance of established risk factors such as household composition and labour market attachment. The incidence of in-work poverty is particularly high among foreign citizens, single-parent households, and large families. Furthermore, logistic regression models highlight the so-called "gender paradox": women living with partner are generally less likely to be working poor, but in single-earner households female workers face a significantly greater risk. This outcome reflects gendered economic dependence and household dynamics, contributing to the broader debate on in-work poverty and its determinants at the local level.

1. Introduction

In recent years, scholarly interest in the issue of in-work poverty, a phenomenon traditionally considered prevalent mainly in the United States, has significantly increased across Europe. Initially addressed primarily by economists, the topic has attracted growing attention from sociological literature as well, particularly since the end of the 2000s and in conjunction with the global economic crisis. Sociological contributions have explored the relationship between in-work poverty and various factors operating at both the micro level (socio-demographic characteristics and economic conditions of households) and the macro level (welfare regimes, wage policies, and labor market structures) (Brady *et al*., 2010; Crettaz, 2013). These studies have highlighted the emergence of a structural paradox that challenges the foundational assumptions of workfare policies: employment no longer guarantees automatic protection from poverty (Filandri, 2022; Saraceno, 2015).

The present contribution is part of a local welfare project promoted by the Municipality of Reggio Emilia, involving a network of public and private local stakeholders united by the objective of developing coordinated interventions to counter economic vulnerability. The focus on in-work poverty emerged as a shared concern among the stakeholders, despite the fact that the province of Reggio Emilia is among the most economically robust areas in Italy. As part of this project, the CGIL made available administrative data collected through its CAF service, allowing for an analysis of in-work poverty at the local level with a degree of territorial detail rarely accessible in other studies.

The aim of this analysis is to investigate how gender, family composition, and the socio-economic characteristics of households influence the risk of experiencing in-work poverty, in light of evidence already established in national and international literature on the topic.

2. Literature Review

The definition commonly adopted by Eurostat - now the European standard - identifies working poor as individuals (both full and part-time workers) who were employed for at least seven months in the reference year and who live in households with an equivalised disposable income below 60% of the national median. This measure, based on the modified OECD equivalence scale[1], incorporates both individual and household dimensions, distinguishing working poverty from the concept of low-wage employment, which is based solely on individual earnings (Salverda, 2018; Gautié and Ponthieux, 2017). As a result, low-paid workers may not be classified as poor, and conversely, higher earners may fall into poverty depending on household composition.

Criticism of the Eurostat indicator concerns the use of a relative poverty line - which may fluctuate with economic cycles - and the exclusion of workers with occasional (or seasonal) employment, a group particularly vulnerable during recessions (Crettaz, 2015; Horemans and Marx, 2013; Marx and Nolan, 2012). In response, scholars have proposed expanding the definition or complementing it with non-monetary indicators, such as material and social deprivation (Andress and Lohmann, 2008; Nolan and Whelan, 2010).

[1] This equivalence scale assigns a weight to each household member based on their age: 1 to the first adult, 0.5 to additional adults, and 0.3 to each child under the age of 14.

Despite these limitations, Eurostat's approach remains essential to capturing the interplay between income, work, and household structure. A purely individual perspective focused on wages risks distorting our understanding of inequalities and family economic well-being (Barbieri *et al*., 2018). In fact, low wages only partially explain working poverty: many working poor are not low-paid, and vice versa (Maitre *et al*., 2012). Household characteristics - such as number of earners, demographic structure, and shared expenses - and the role of national redistributive policies are key factors influencing poverty risk (Filandri and Struffolino, 2019).

Policy implications vary depending on the level of analysis. An individual wage-based approach emphasizes minimum wages, contract stability, and skills training. Conversely, a household-centred perspective highlights the importance of multiple earners and suggests policies that promote employment participation, parenting, work-life balance, and the inclusion of disadvantaged groups (Barbieri *et al*., 2018; Marx and Nolan, 2012).

The sociological literature on in-work poverty can be divided into three main strands. The first examines macroeconomic factors and labour market transformations. The so-called Unified Theory (Brady *et al*., 2010; Crettaz, 2013) sees in-work poverty as a consequence of institutional failure to absorb macroeconomic shocks. Theories of skill-biased technological change and labour market polarization link the phenomenon to structural changes in the economy (Barbieri *et al*., 2018), while empirical studies point to the effects of growth, unemployment, and sectoral composition (Albertini *et al*., 2020; Lohmann, 2009).

A second line of research focuses on welfare regimes and their impact on in-work poverty rates (Andress and Lohmann, 2008; Brady *et al*., 2010). Social-democratic countries generally show lower rates thanks to inclusive and universal systems, whereas Mediterranean and liberal regimes record higher rates due to weaker or more selective protections (Lohmann, 2009; Saraceno *et al*., 2022)

A third group of studies looks at socio-demographic risk factors. Single-parent families, large households, low education levels, and foreign citizenship are consistently associated with high in-work poverty risk (Crettaz, 2013; Crettaz, 2018; Polizzi *et al*., 2022). Furthermore, self-employed workers show a significantly elevated risk (Horemans and Marx, 2017).

The model proposed by Crettaz and Lohmann (2018) integrates macroeconomic, institutional, and demographic dimensions. It explains in-work poverty as the result of the interaction between low wages, weak work intensity, high family needs, and limited public redistribution, producing different outcomes depending on the balance between market dynamics, welfare systems, and family structures. A notable gap concerns the lack of micro-level, locally grounded analyses capable of capturing how individual socio-demographic characteristics, such as gender, and household configurations shape in-work poverty. Survey data rarely allow this level of detail.

Using rich local administrative records, this study addresses this limitation and offers a distinct empirical contribution.

In Italy, in-work poverty is a structural issue within the national labour market. The country's rate is nearly two percentage points above the EU average, with almost 10% of workers employed at least seven months per year living in poverty. Alongside Spain, Greece, and Portugal, Italy belongs to the so-called "Mediterranean poverty regime" (Saraceno *et al*., 2022), marked by weak labour market inclusion, high youth and female unemployment, limited active labour market policies, and a reliance on family-based welfare.

Labour market segmentation, exacerbated by deregulation since the 1990s and the 2008 economic crisis, has fostered a dual system: protected and stable workers coexist with an increasing share of precarious, involuntary part-time, and underpaid self-employed workers, often young people, women, and migrants. Marginal flexibility has weakened the bargaining power of low-skilled workers, increasing wage and regional inequalities, especially in Southern Italy, where most working poor are concentrated. Furthermore, in-work poverty in Italy often proves persistent rather than temporary, affecting the same individuals or families over long periods (Barbieri *et al*., 2018).

Weak public redistribution and fragmented, insufficient social services contribute to the vulnerability of many households, especially those with a single income, single parent, multiple dependents, or inactive members. Persistent gender asymmetries and the dominance of the male breadwinner model also increase working poverty risk, despite recent improvements in female employment.

International research has highlighted a paradoxical relationship between gender and in-work poverty. Although women are more exposed to precarious and low-paid employment, they often show a similar - or even lower - risk of in-work poverty than men. This "gender paradox" (Peña-Casas and Ghailani, 2011) reflects women's secondary economic role within households: their earnings commonly supplement those of a primary male breadwinner, thus protecting dual-earner families (Ponthieux, 2018). Yet this household-level shield can obscure women's individual vulnerability, including dependency and unequal intra-family resource allocation shaped by persistent gender norms (Kulic and Dotti Sani, 2020). Indeed, such disadvantages are particularly evident for single women or mothers lacking a second income (Gautié and Ponthieux, 2017).

Although in-work poverty can potentially be reduced by increasing the number of earners, typically through female employment, research warns against the exclusive reliance on this strategy. Policies aimed solely at raising women's labor market participation, without improving job quality, may even heighten individual poverty risks (Barbieri *et al*., 2018; Filandri and Struffolino, 2019).

These findings underscore the need to examine more closely how gender and household structure shape the likelihood of experiencing in-work poverty, the central focus of the present study.

3. Data and methods

This study uses microdata provided by the CGIL of Reggio Emilia, based on tax returns and ISEE (Equivalent Economic Situation Indicator) applications collected in 2021 by local CAF offices. Although these data are not an official statistical source, they allow for an in-depth analysis of in-work poverty at a fine-grained local level, such as municipalities, which are not typically covered by standard surveys.

The two datasets (tax returns and ISEE declarations) were linked using two identifiers: a personal ID and a family ISEE code. This linkage enabled the construction of family-level variables assigned to each individual.

The dataset includes: family ISEE code, individual ID, relationship to the ISEE applicant, age, gender, citizenship (Italian or foreign), employment status and contract type, days worked, individual and spousal earnings (if jointly declared), total income, municipality of residence, and net tax. From these, net individual income was calculated by subtracting net tax (including local add-ons) from gross income. Summing the net income of all household members and adjusting using an equivalence scale produced the net equivalent family income.

In accordance with the Eurostat definition, individuals were classified as working poor if they were employed for at least seven months in 2021 and their household equivalent income fell below 60% of the median household equivalent income. However, due to the non-random nature of the dataset, the poverty threshold was not calculated internally. This may have led to an underestimation of poverty, given the generally disadvantaged profile of the CAF sample.

To address this issue, a locally calibrated threshold was adopted in place of the national median income, in order to better reflect the socio-economic conditions of the population under study. Specifically, the poverty line was defined externally as 60% of the net median household income in Emilia-Romagna, based on Istat's "Income and Living Conditions" survey, and adjusted using a simplified equivalence scale (the square root of household size). The same scale was applied to the CAF data to ensure comparability. Households without any employed members (e.g., pensioners only) and cases with incomplete income information were excluded from the final dataset.

Using CAF data offers several strengths: it relies on official tax records, which reduces the bias from self-reported income, and provides detailed local information. However, it also has limitations. The data are not representative of the general population, as they come from individuals who voluntarily sought services from a union-affiliated CAF, likely overrepresenting socially vulnerable groups. Moreover, some important variables - such as educational level and weekly working hours - are missing, making it impossible to distinguish between full-time and part-time work. The datas*et al*so does not allow for analysis of post-transfer poverty, as information on social benefits is lacking. These limitations must be kept in mind, but the methodology remains robust and consistent with European standards.

The empirical analysis proceeds in two steps. First, descriptive statistics are used to measure the rate of in-work poverty in 2021, disaggregated by key individual and household characteristics. Second, logistic regression models are estimated to examine the relationship between gender and the likelihood of being working poor. These models are appropriate for binary outcomes and estimate the probability of in-work poverty based on categorical (e.g., gender) and continuous variables, using a logit function and maximum likelihood estimation.

To explore the gender effect in more detail, an interaction term is added between gender and a binary variable indicating whether the household has one or more earners. Two additional gender-specific models are also estimated. The first uses a classification of households based on the number and type of workers (following Barbieri *et al*., 2018). The second adopts a simplified Istat-based classification[2] that considers household structure and member relationships.

4. Results

As expected, the descriptive analysis reveals that the in-work poverty rate within the CAF-CGIL sample (23.7%) is higher than that of the Italian population as a whole (11.7%). The share of working poor in the sample is slightly more than double the national one reported by Eurostat in 2021. This suggests that CAF-CGIL clients in the province of Reggio Emilia experience a condition of economic disadvantage at the household level compared to the rest of the provincial population. Moreover, the in-work poverty rate for female workers (22.6%) is slightly lower than that observed for men (24.8%), a trend also seen at the national level and in several European countries (Eurostat, 2021).

To explore the association between gender and in-work poverty, an initial logistic regression model was estimated, to which individual demographic control variables and variables describing household and occupational characteristics were progressively added. The construction of the regression model is inspired by the work of Barbieri *et al*. (2018), adapted to the specific research needs and the data

[2] See: Famiglie per tipologia di famiglia

limitations, particularly regarding the selection of control variables included in the model.

The effect of gender remains robust across all model specifications: women exhibit a significantly lower likelihood of experiencing in-work poverty than men, even after controlling for age, citizenship, marital status, geographical area, household composition, and occupational characteristics. This result is consistent with the hypothesis of the so-called "gender paradox" (Peña-Casas and Ghailani, 2011; Ponthieux, 2018), according to which female workers are less exposed to the risk of in-work poverty than male workers, once individual socio-demographic and occupational characteristics are accounted for.

To further investigate gender differences and verify whether the effect of gender varies depending on the household's occupational structure, an interaction between gender and a dummy variable describing the number of workers in the household (0 = only one worker, 1 = more than one worker) was estimated. The gender effect varies substantially when the number of workers in the household is taken into account. As shown in Figure 1, the predicted probability of being working poor is significantly higher for women when they are the sole earner in the household (almost 65%), compared to men in the same situation (approximately 51%).

Figure 1 - *Predicted probabilities of being working poor by gender and number of earners in the household. 95% C.I.*

However, when more than one worker is present in the household, this probability drops drastically for both genders, settling at very similar levels: 6.4% for men and 6.8% for women.

In line with our research objectives, to further examine how the probability of being working poor varies depending on the interaction between gender and household characteristics, two additional logistic regression models were estimated.

The first model can be expressed as follows:

$$Logit\left(P(Y=1)\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_5 X_5 \qquad (1)$$

where the main independent explanatory variable of interest, X_1, is a categorical qualitative variable describing the household occupational characteristics[3]. Only individual-level control variables not strongly correlated with the two types of household structures were included in the model to limit multicollinearity issues.

Figure 2 - *Average Marginal Effects (AME) of Household Occupational Characteristics on the Probability of being Working Poor, by Gender. 95% C.I.*

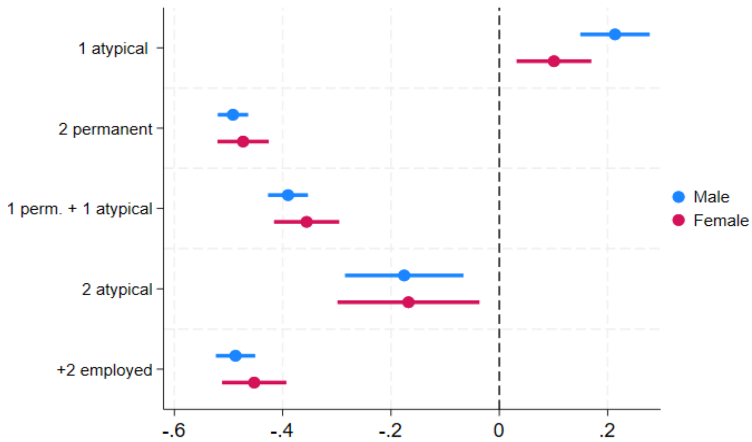


Figure notes: Reference category: only one permanent worker in the household

As shown in Figure 2, which reports the estimated AMEs (Average Marginal Effects), any combination of dual employment (whether with stable or atypical contracts) significantly reduces the risk of in-work poverty compared to the reference category of a single permanent worker. In particular, the presence of two permanent

[3] $\boldsymbol{X}_2 \dots \boldsymbol{X}_5$ = citizenship, age, marital status, municipality of residence
Models estimated separately for males (M) and females (F)

workers in the household reduces the probability of being working poor by over 40 percentage points (p.p.) for both genders. Other household configurations with at least two workers, such as the presence of one stable and one atypical contract (approximately -39 p.p. for men and -35 p.p. for women) or two atypical contracts (around -17 p.p. for both), are also associated with a lower probability of in-work poverty compared to the reference category. In households with only one atypical worker, the risk of poverty compared to those with a single permanent worker appears to be more pronounced among men (approximately +20 p.p.) than among women (about +10 p.p.). These results suggest that, even within the CAF-CGIL sample, the protective effects of dual employment are relatively similar for both men and women. Thus, even in a local context, the national-level findings of Barbieri and colleagues (2018) are confirmed: the presence of two or more workers in a household significantly reduces the risk of in-work poverty, regardless of contractual status.

The second model follows the same formulation as the previous one, but the main independent explanatory variable (X_1) is the household type.

Figure 3 - *Average Marginal Effects (AME) of household type on the probability of being working poor, by gender. 95% C.I.*

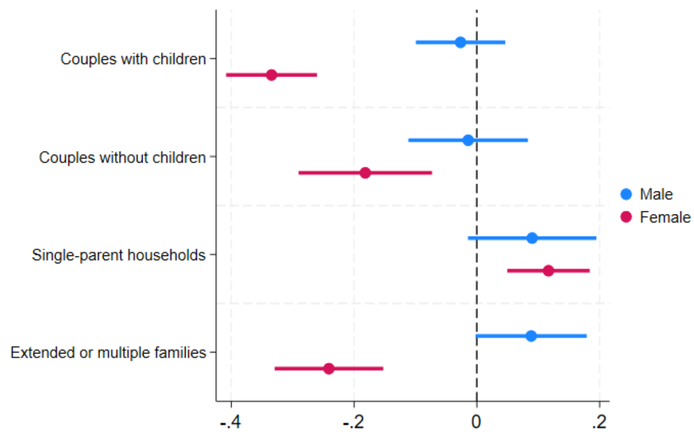


Figure notes: Reference category: solo member household

Looking at the AMEs reported in Figure 3, it can be seen that for women, living in a couple - whether with or without children - is associated with a substantial reduction in the absolute probability of being working poor compared to living alone: around 35 p.p. less for women living in a couple with children, and nearly 20 p.p. for those without children. Also, belonging to extended or multi-nuclear families represents a markedly protective economic condition for female workers compared to living alone (over 20 p.p. less). For men, by contrast, being in a couple (with or without children) leads to a much smaller reduction in the risk of in-work poverty, close to zero and statistically not significant. On the contrary, for men as well, living in a single-parent household or in extended/multi-nuclear families results in an absolute increase in the risk of being working poor of nearly 10 p.p. in both cases. However, in these two cases, the wide confidence intervals indicate greater uncertainty in the estimates.

These results show that for female workers, living in a couple or extended family is strongly protective against the risk of in-work poverty compared to living alone. This finding confirms what has emerged at the national level, where the presence of shared family resources represents an important safeguard, particularly for women (Ponthieux, 2018). For men, however, the relationship with family structure is less clear: even when the result is statistically significant, as in the case of extended or multi-nuclear families, the direction of the effect is opposite to that observed for women, indicating a higher risk of experiencing in-work poverty.

5. Conclusion

The analysis confirms that gender remains a crucial factor in determining the risk of in-work poverty: on average, women are less exposed than men when controlling for relevant characteristics, in line with the so-called “gender paradox” described in the literature. The findings also point to a weakening of the traditional single-earner model: being the sole income earner, also for men, is no longer a safeguard against economic vulnerability. In this regard, the role of the "male breadwinner" appears to be losing its effectiveness in today’s labour market.

These dynamics are clearly reflected within the CAF-CGIL Reggio Emilia sample, a population of formally employed workers who, nonetheless, are not exempt from economic hardship. The use of CAF data has proven especially valuable for analysing in-work poverty at the local level: these are actual administrative data, not self-reported, on both individual and household income, updated and detailed, providing a solid empirical basis for the analysis of economic inequality.

Looking ahead, integration with other administrative sources - such as INPS records and municipal population registries - could further enhance the analytical potential of the dataset, enriching the available information with additional variables (e.g., hours worked, social transfers, educational attainment, sector of employment, occupational status), and enabling the construction of statistically representative samples. The use of integrated datasets could open up new avenues of research aimed at exploring the relationships between individual and household conditions and poverty, including through the use of alternative indicators to the in-work poverty measure adopted here.

References

ALBERTINI M., BALLARINO G., DE LUCA D. 2020. Social Class, Work-Related Incomes, and Socio-Economic Polarization in Europe, 2005–2014. *European Sociological Review*, Vol. 36, No. 4, pp. 513–532.

ANDRESS H.J., LOHMANN H. 2008. *The working poor in Europe: Employment, poverty and globalisation*. Cheltenham: Edward Elgar Publishing.

BARBIERI P., CUTULI G., SCHERER S. 2018. In-work poverty in un mercato del lavoro duale: Individualizzazione riflessiva dei rischi sociali o stratificazione della diseguaglianza sociale? *Stato e mercato*, Vol. 3, pp. 419–460.

BRADY D., FULLERTON A.S., MOREN CROSS, J. 2010. More Than Just Nickels and Dimes: A Cross-National Analysis of Working Poverty in Affluent Democracies. *Social Problems*, Vol. 57, No. 4, pp. 559–585.

CRETTAZ E. 2013. A state-of-the-art review of working poverty in advanced economies: Theoretical models, measurement issues and risk groups. *Journal of European Social Policy*, Vol. 23, No. 4, pp. 347–362.

CRETTAZ E. 2015. Poverty and material deprivation among European workers in times of crisis. *International Journal of Social Welfare*, Vol. 24, No. 4, pp. 312–323.

CRETTAZ E. 2018. In-work poverty among migrants. In LOHMANN H., and MARX I. (Eds.) *Handbook on In-Work Poverty*, Cheltenham: Edward Elgar Publishing, pp. 89-105.

FILANDRI M. 2022. Lavorare non basta. Bari: GLF editori Laterza.

FILANDRI M., STRUFFOLINO E. 2019. Individual and household in-work poverty in Europe: Understanding the role of labor market characteristics. *European Societies*, Vol. 21, No. 1, pp. 130–157.

GAUTIÉ J., PONTHIEUX S. 2017. Employment and the Working Poor. In BRADY D. and BURTON L. M. (Eds.) *The Oxford Handbook of the Social Science of Poverty*, Vol. 1. Oxford University Press, pp. 486–504.

HOREMANS J., MARX I. 2013. In-work poverty in times of crisis: Do part-timers fare worse? *ImPRovE Discussion Paper*, *13/14*.

HOREMANS J., MARX I. 2017. Poverty and Material Deprivation Among the Self-Employed in Europe: An Exploration of a Relatively Uncharted Landscape. *IZA DP No. 11007.*

KULIC N., DOTTI SANI G. 2020. Within-couple distribution of economic resources. A critical review of extant studies and avenues for future research. *Rassegna Italiana Di Sociologia*, *EarlyAccess*, 0–0.

LOHMANN H. 2009. Welfare States, Labour Market Institutions and the Working Poor: A Comparative Analysis of 20 European Countries. *European Sociological Review*, Vol. 25, No. 4, pp. 489–504.

MAITRE B., NOLAN B., WHELAN C. T. 2012. Low pay, in-work poverty and economic vulnerability: A comparative analysis using EU-SILC. *The Manchester School*, Vol. 80, No. 1, pp. 99–116.

MARX I., NOLAN B. 2012. In-work poverty. *GINI Discussion Paper 51,* Amsterdam Institute for Advanced Labour Studies.

NOLAN B., WHELAN C. T. 2010. Using non-monetary deprivation indicators to analyze poverty and social exclusion: Lessons from Europe? *Journal of Policy Analysis and Management*, Vol. 29, No. 2, pp. 305–325.

PEÑA-CASAS R., GHAILANI D. 2011. Towards Individualizing Gender In-Work Poverty Risks. In FRASER N., GUTIÉRREZ R., PEÑA-CASAS R. (Eds.) *Working Poverty in Europe,* Palgrave Macmillan UK, pp. 202–231.

POLIZZI A., STRUFFOLINO E., VAN WINKLE, Z. 2022. Family demographic processes and in-work poverty: A systematic review. Advances in Life Course Research, Vol. 52.

PONTHIEUX S. 2018. Gender and in-work poverty. In LOHMANN H., and MARX I. (Eds.) *Handbook on In-Work Poverty*, Cheltenham: Edward Elgar Publishing, pp. 70-88.

SALVERDA W. 2018. Low earnings and their drivers in relation to in-work poverty. In LOHMANN H., and MARX I. (Eds.) *Handbook on In-Work Poverty*, Cheltenham: Edward Elgar Publishing, pp. 26-49.

SARACENO C. 2015. *Il lavoro non basta: La povertà in Europa negli anni della crisi*. Milano: Giangiacomo Feltrinelli Editore.

SARACENO C., BENASSI D., MORLICCHIO E. 2022. *La povertà in Italia*. Bologna: il Mulino.

Gianluca TRUSCELLO, Università Cattolica del Sacro Cuore Milano, gianluca.truscello@unicatt.it

Maria Chiara ZANAROTTI, Università Cattolica del Sacro Cuore Milano, maria.zanarotti@unicatt.it

HOUSEHOLD FINAL CONSUMPTION EXPENDITURE DISTRIBUTIONAL ACCOUNTS: HARMONISING MACRO AND MICRO DATA[1]

Sara Basso, Incoronata Donnarumma, Stefania Massari

Abstract. Well-being is a multidimensional concept including income, consumption and wealth and their measures are crucial to the design of economic and social policies. Aggregate measures and average values fail to capture the disparities existing among different types of households, which remain far from homogeneous.
Average values are meaningful statistics, but they do not tell the whole story about living standards.
The reconciliation of micro and macro data on households is essential. Micro data sources, can provide distributional information among households but they may not be consistent across the primary components of economic well-being and may not be comparable across countries. In contrast, the system of National Accounts provides comprehensive, consistent and internationally comparable information but it cannot provide any evidence on distribution of economic resources among groups of households.
To bridge macro and micro data, National Accounts values can be combined with distributional indicators from micro data sources, carefully accounting for potential differences in definitions and concepts between macro and micro aggregates.
This paper aims to harmonise micro and macro data as a first step toward developing experimental distributional estimates for household consumption, based on National Accounts data as well as on survey data for consumption (HBS).

1. Introduction

Measuring the level and evolution of economic inequality means attempting to assess people's living conditions and well-being of individuals and households (Blanchet, Chancel, Gethin, 2019). Well-being is a multidimensional concept including income, consumption and wealth and aggregates and average values are unable to capture disparities between different types of households (Lustig, 2018).

The reconciliation of micro and macro data on households is crucial. In fact micro data sources (surveys or administrative records) can provide distributional

[1] This paper is a joint effort of the authors. However, Sara Basso is the author of sections 1 and 2.1, Incoronata Donnarumma is the author of sections 2.3 and 4, Stefania Massari is the author of sections 2, 2.2 and 3.

information among households but they are not consistent across the primary components of economic well-being (e.g. income, consumption and wealth) and not comparable across countries. On the other hand, the System of National Accounts provides comprehensive, consistent and internationally comparable information but it cannot provide any evidence on distribution of economic resources among groups of households (Fesseau and Van De Ven2014).

The distributional accounts are a key issue: information in line with National accounts totals (Zwijnenburg, 2022) and able to provide data on the economic resources distribution across households (Coli *et al.*, 2022).

In order to bridge macro and micro data, National Accounts values can be read jointly with distribution indicators from micro data sources, paying attention to the fact that that macro aggregates may not fit the micro aggregates in terms of definitions and concepts. Adjustments are needed to integrate micro information into System of National Accounts framework.

This paper aims at harmonising micro and macro data that is the first step to compile experimental distributional estimates for household consumption. It is based on National Accounts Household Final Consumption Expentiture (HFCE) data as well as survey data for consumption (HBS).

2. The macro and micro perspective: similarities and differences

The first step to achieve this goal is to develop a good understanding of the differences between micro-sources (HBS) and macro-sources (HFCE). Such differences arise not only from the scope of what is considered consumption, but also they also from in the varying classifications and other adjustments that are specific to each source (OECD, 2020).

To ensure international comparability, both the HBS and the HFCE are based on the harmonized international classification of expenditure items, Classification of Individual COnsumption by Purpose (UNSD, 2000) – Coicop, but they follow two different regulations: HBS is based on Regulation (EU) 2019/1700, Integrated European Social Statistics; National Accounts are based on "European system of accounts ESA 2010" (Eurostat, 2013) that is an internationally compatible accounting framework for a systematic and detailed description of total economy (that is a region, country or group of countries), its components and its relations with other total economies.

The HBS focuses on consumption expenditure behaviours of households residing in Italy. The survey analyses the evolution of level and composition of household consumption expenditure and it represents the informative base for the official estimates of relative and absolute poverty in Italy and for the inflation measure by household expenditure classes.

The focus of the HBS is represented by all expenditures incurred by resident households to purchase goods and services exclusively devoted to household consumption (including self-consumptions and imputed rentals); every other expenditure for a different purpose is excluded from the data collection (e. g., payments of fees and business expenditures).

ESA 2010 defines final consumption expenditure as the expenditure incurred by resident institutional units on goods or services used for the direct satisfaction of individual needs or wants or the collective needs of members of the community.

In NA, final consumption expenditure is calculated as total expenditures made by all households, resident or not, within the economic territory and adjusted by adding the expenditures of residents abroad and subtracting the expenditures of non residents within the economic territory.

According to the ESA 2010 definition, household final consumption expenditure includes the following items that are not detected or differently treated in the HBS:

1. services of owner-occupied dwellings;
2. income in kind,
3. financial services directly charged and the part of FISIM used for final consumption purposes by households;
4. insurance services by the amount of the implicit service charge;
5. pension funding services by the amount of the implicit service charge;
6. illegal activities as narcotics, smuggling of tobacco and prostitution;
7. tips

Instead household final consumption expenditure excludes:

1. social transfers in kind,
2. items treated as intermediate consumption or gross capital formation

Table 1 highlights the differences between HFCE and HBS, along with all the items involved. This analysis needs to be conducted for the 41 Coicop groups (3-digit) to better understand and detail all the discrepancies that must be addressed. While not all categories are a, some groups in those involved require more in-depth analysis during the harmonization stage.

Own final consumption in agriculture is similar in terms of concept in both domains, but different in estimation method. Illegal activities and FISIM in CP02 and CP12 (smuggling of tobacco, narcotics and prostitution) are not detected in HBS, but included in HFCE. In CP04, imputed rents and major maintenance of dwelling need to be discussed: the estimation method is different between HBS and HFCE for the first item mentioned, while the second one is not included in HFCE. In CP07, second-hand cars exchanging between households are excluded in HFCE but included in HBS (considered as an expenditure as well). Definitions are different also for gambling (CP09) and insurance (CP12).

Table 1 – *Differences in definitions and concepts.*

Division		Group		NA	HBS
CP01	**Food and non-alcoholic beverages**	CP011	Food	Own final consumption of agricultural products: estimated on the basis of statistics on agricultural production	Own final consumption of agricultural products: quantities are detected by the HBS
		CP012	Non-alcoholic beverages		
CP02	**Alcoholic beverages, tobacco and narcotics**	CP021	Alcoholic beverages		
		CP022	Tobacco	Smuggling of cigarettes included	Smuggling of cigarettes not detected by the HBS
		CP023	Narcotics	Included in NA	Not detected by the HBS
CP03	Clothing and footwear				
CP04	**Rents, fuels and maintenance of the dwelling**	CP041	Actual rentals for housing		
		CP042	Imputed rentals for housing	Estimated by applying market rents to the housing stock	Imputed rents estimated by the households are detected by HBS
		CP043	Maintenance and repair of the dwelling	Major maintenance of the dwelling excluded from NA	Major maintenance of the dwelling included in the HBS
		CP044	Water supply and miscellaneous services related to the dwelling		
		CP045	Electricity, gas and other fuels		
CP05	Goods and services for the dwelling				
CP06	Health				
CP07	**Transport**	CP071	Purchase of vehicles	Second-hand cars excludes exchanges of cars between households	Second-hand cars includes exchanges of cars between households
CP08	Communication				
CP09	**Goods and services for recreation and culture**	CP091	Audio-visual, photographic and information processing equipment		
		CP092	Other major durables for recreation and culture		
		CP093	Other recreational items and equipment, gardens and pets		
		CP094	Recreational and cultural services	Gambling included in NA net of winnings	Gambling included in HBS gross of winnings
		CP095	Newspapers, books and stationery		
		CP096	Package holidays		
CP10	Education				
CP11	**Restaurants and hotels**	CP111	Catering services	Income in kind included in NA	Income in kind not detected by the HBS
		CP112	Accommodation services	Income in kind included in NA	Income in kind not detected by the HBS
CP12	**Miscellaneous goods and services**	CP121	Personal care		
		CP122	Prostitution	Included in NA	Not detected by the HBS
		CP123	Personal effects n.e.c.		
		CP124	Social protection		
		CP125	Insurance	Supplementary insurance premiums included in NA Only insurance services	Supplementary insurance premiums not detected by the HBS Expenditures on insurance are recorded gross of any reimbursements
		CP126	Financial services n.e.c.	FISIM Included in HFCE	FISIM Not detected by the HBS
		CP127	Other services n.e.c.		

Harmonising HBS and HFCE is a key step in allocating consumption expenditure among household groups taking into account differences in definitions and concepts but also in reference population.

2.1 The reference population

Following the EG DNA provided recommendations, first step is the correction for expenditures of non-resident households on the territory and of resident households abroad. The choice of 2019 as the reference year is related to the availability of the Tourism Satellite Account for that year and thus the possibility of using this data to adjust some consumption categories.

As mentioned above HFCE follows a domestic concept (expenditures of non-resident households on the territory are included while expenditures of resident households abroad are excluded) while the HBS follows a national concept (expenditures of non-resident households on the territory are excluded while expenditures of resident households abroad are included).

Moreover the population underlying HFCE differs from the population underlying the HBS: the survey covers the resident population with the exclusion of persons living permanently in institutions or without a registered place of residence while the reference population in HFCE is the present population on the national territory at a given date including households and persons living in institutions (convents, boarding schools, prisons, etc.). Reference population according to the HFCE concept is obtained by subtracting the number of residents temporarily abroad and adding foreigners present on the territory but not resident. As foreigners non-resident, tourists and present foreigners in Italy for one year or more (non-tourists) are taken into account. Non-resident foreigners include both foreigners with residence permit, but without a residence certificate, and unregistered foreigners without or expired residence permit. Stays in hotels and other accommodation structures for tourists collected by statistics on tourism are used to estimate the non-resident population on the Italian territory.

Table 2 shows the population underlying the HFCE.

Table 2 – *HFCE population, 2019 (thousands).*

	2019
Resident population (annual average)	59,729
Citizen temporarily resident abroad	-393
Non-resident foreigners present for at least one year	540
Foreign tourists	605
NA consumer population	**60,480**

Once defined the NA reference population, the first attempt to reconcile NA HFCE with HBS is a proportional “removing” the consumption of the population not covered in the micro source using the ratio between the two reference populations (table 3).

Table 3 – *HBS/NA population, 2019 (thousands).*

Reference population	2019
HBS (a)	59,211
NA (b)	60,480
Coefficient (b/a)	1.021

In fact, due to the lack of detailed information on expenditures of non-resident households on the territory and those of resident households abroad for each Coicop items, implicit coefficient derived is used as a correction coefficient at an aggregated level to move NA figures from domestic to national concept. Of course this leads to "rough" adjustment, because consumption expenditure by non-residents on territory and resident abroad may vary significantly across consumption items.

Table 4 shows the coverage rate (micro aggregate as a percentage of NA total) for all consumption items: first column ("raw data") is the ratio between HBS and NA without any adjustment. The total coverage rate is 72.3 percent: some items are well covered, some others show a rate less than 50 percent and in the case of the CP04 division the micro item is higher than NA estimate.

Table 4 – *Coverage rates for consumption items, 2019.*

Coicop (2-digit)	Coverage rate		
	Raw data	Adjusted data	
		P_1*	P_2**
Food and non-alcoholic beverages	93.0	94.4	95.2
Alcoholic beverages, tobacco and narcotics	31.2	31.9	31.7
Clothing and footwear	54.3	55.5	55.7
Rents, fuels and maintenance of the dwelling	110.4	112.8	111.8
Goods and services for the dwelling	52.9	54.0	54.2
Health	95.5	97.6	97.9
Transport	65.3	66.7	66.7
Communication	78.6	80.3	80.5
Goods and services for recreation and culture	51.9	53.0	53.2
Education	50.2	51.3	51.4
Restaurants and hotels	35.5	36.3	38.4
Miscellaneous goods and services	50.8	51.9	52.0
Total	**72.3**	**73.8**	**74.0**

** Proportional adjustment by the ratio between the two reference populations*
*** Proportional adjustment by the ratio between the two reference populations and tourism satellite account (for specific items)*

The correction for expenditures of non-resident households on the territory and of resident households abroad was made applying the population ratio (shown in

Table 3) to all Coicop categories leading to a better alignment and a reduction of the micro-macro gap (Table 4).

To adjust certain specific consumption categories at the detailed level (according to the national concept), information from Tourism Satellite Account are taken into account. The categories involved are imputed rents (CP04), transport services (CP07), recreational and cultural services (CP09), and restaurants and hotels (CP11). These categories are adjusted using the satellite account information, while the remaining categories are adjusted based on the population ratio (this explains the differing coverage rates between columns P1 and P2, even for items not covered by the satellite account).

2.2 Definitions and concepts

After the population adjustment, it is necessary to align NA totals to differences in definitions and concepts. Differences in definitions and concepts can be grouped into two types: treatment of items considered in both domains and types of expenditure covered by the survey but not by the NA, or vice-versa.

Table 5 – *Coverage rates for consumption items, 2019.*

Coicop (2-digit)	Coverage rate		
	Raw data	Adjusted data	
		P_2^{*}	P_3^{**}
Food and non-alcoholic beverages	93.0	95.2	96.5
Alcoholic beverages, tobacco and narcotics	31.2	31.7	51.7
Clothing and footwear	54.3	55.7	55.7
Rents, fuels and maintenance of the dwelling	110.4	111.8	111.8
Goods and services for the dwelling	52.9	54.2	54.2
Health	95.5	97.9	97.9
Transport	65.3	66.7	66.7
Communication	78.6	80.5	80.5
Goods and services for recreation and culture	51.9	53.2	71.3
Education	50.2	51.4	51.4
Restaurants and hotels	35.5	38.4	40.5
Miscellaneous goods and services	50.8	52.0	50.6
Total	**72.3**	**74.0**	**77.7**

* *Population adjustment*
** *Conceptual adjustment*

Table 5 presents a comparison between the HBS results and NA estimates. The comparison is made with not adjusted data (first column) and data after the related adjustments (second and third column). Adjustments made for population had already leading HBS/NA ratio from 72.3 to 74.0 for total consumption.

Conceptual adjustments consist in excluding from NA those items that are not detected by the survey (e.g. expenses related to illegal activities, FISIM, tips and income in kind) and also some items that, although considered household expenditure by the survey and by NA, are quantified in different ways (e.g. spending on gambling, insurance, etc.).

This second step further improves the coverage rate to 77.7 for total consumption, at a more disaggregated level of expenditure, the ratio varies greatly: from 40.5 for the expenditure on restaurants and hotels to 111.8 for the expenditure on housing.

It is worth emphasizing that although we have compared the data made homogeneous both for the underlying population and from a conceptual point of view, a rather high gap remains for some consumption divisions, such as clothing and footwear.

It is important to stress that the fit between NA and HBS depends not only on the conceptual differences listed above but also on the sources used in NA to estimate household consumption. Clothing and footwear division is a clear example of this, it has a very low fit even if it has no conceptual differences.

2.3 National accounts sources

NA are not intended to cover aspects of households' well-being and several sources are used to derive household consumption, including HBS; moreover balancing process of the National accounts may have relevant impact on consumption estimates.

Five main groups of sources and methods identified are the following: commodity flow method (CFM), Household Budget Survey (HBS), Multipurpose Survey (MS), other Istat surveys (OIS) and administrative and other sources (Admins). All sources contribute to define the household consumption estimations and refers to specific item. Table 6 shows the sources involved for each consumption item.

The use of surveys on the demand side and their integration with other sources of information ensure a good degree of coverage, since no source, taken individually, can be considered as appropriate for estimating the overall consumption by Coicop item.

Table 6 – *Sources in NA household consumption estimation, by Coicop.*

Division		Sources
CP01	Food and non-alcoholic beverages	HBS
CP02	Alcoholic beverages, tobacco and narcotics	CFM/HBS/Admins
CP03	Clothing and footwear	CFM
CP04	Rents, fuels and maintenance of the dwelling	HBS/Admins
CP05	Goods and services for the dwelling	HBS/CFM/Admins
CP06	Health	HBS/Admins
CP07	Transport	HBS/Admins
CP08	Communication	HBS/CFM/Admins
CP09	Goods and services for recreation and culture	CFM/Admins
CP010	Education	HBS/MS/Admins
CP011	Restaurants and hotels	HBS/MS/OIS
CP012	Miscellaneous goods and services	CFM/HBS/Admins

The comparison of independent sources allows to capture a part of non-observed economy, not reported in tax statements of companies, and also to integrate phenomena partially measurable on the basis of information collected from households. Information from HBS is examined and then integrated with other sources and used mainly to estimate spending on food, housing, health services (particularly on health outpatient services), communications and other services included in the Coicop division which refers to miscellaneous goods and services.

The balancing procedure is the last step and corrects the discrepancies between the aggregates of resources and uses according to the domestic concept.

3. Micro-macro gap

Once all possible adjustments have been made, remaining gaps have to be allocated. The EG DNA guidelines suggest four methods for the gap allocation in order to distribute the NA totals using micro data:

- *Method A (direct method)*: the distribution of the gap is made proportionally to the micro values of same indicator, i.e. applying the same adjustment coefficient (macro total/micro total) to all households (their totals match NA totals);
- *Method B (indirect method based on proxies)*: a missing or unreliable micro component is estimated by using the distribution of another consumption component as a proxy;
- *Method C (indirect method based on external data)*: a missing or unreliable micro component considered can be distributed according to exogenous data (e.g.

sociodemographic information) available at the level of the individual or of the household;

- *Method D (invariant method)*: the remaining components are distributed in proportion to the total of all the NA and the imputations are made in such a way that the inclusion or exclusion of the component does not affect the distributional results of the main indicators.

Only M1 and M3 methods were deemed suitable for consumption by Eurostat and applied in the centralised exercise. Each Coicop category requires a separate analysis to choose the most suitable method.

All the considerations made so far are summarized in table 7 which shows the coverage rate for each Coicop division, but also, in the following two columns, a qualitative assessment which depends respectively on the conceptual fit and the use of HBS as a source in NA.

Table 7 – *Assessment of "linkage" in Coicop divisions between NA and HBS.*

Division		HBS/NA (adjuested for population and conceptual differences)	Conceptual link	HBS use
CP01	Food and non-alcoholic beverages	96.5	high	high
CP02	Alcoholic beverages, tobacco and narcotics	51.7	low	medium
CP03	Clothing and footwear	55.7	high	low
CP04	Rents, fuels and maintenance of the dwelling	111.8	medium	medium
CP05	Goods and services for the dwelling	54.2	high	medium
CP06	Health	97.9	high	high
CP07	Transport	64.4	medium	medium
CP08	Communication	80.5	high	high
CP09	Goods and services for recreation and culture	71.3	medium	medium
CP010	Education	51.4	high	low
CP011	Restaurants and hotels	40.5	medium	low
CP012	Miscellaneous goods and services	50.6	low	low

The result of two last columns "shows" in which division we can assume that using HBS to obtain distributional estimates is a good approximation, i.e. food or communication, whereas miscellaneous good and services where the assessment is low-low, probably need to be investigated in depth.

Summing up where the conceptual link and HBS use are indicated as high, method A can be applied. Method M1 can be also applied to the items where only the conceptual link is high, even if the use of HBS is indicated as low or medium: micro data are in fact close in conceptual term to the adjusted totals of NA. The only two Coicop items with low conceptual link are CP02 and CP12, mainly due to illegal activities and FISIM. Where the assessment is low both in conceptual link and HBS use probably need of another method for the gap allocation.

4. Final remarks and way forward

Reconciliation of micro and macro data is a key issue to define distributional accounts. HFCE need to be harmonised with HBS in order to use the distributional information provided by the survey in the framework of National accounts. The distance between HFCE and HBS is not only related to conceptual differences and reference population but the most part derives from the sources used in HFCE estimates: this makes reconciliation challenging.

The empirical approach required the investigation of all available sources to define and better understand the micro-macro gap and then try to allocate it as properly as possible (Coli, Tartamella, 2017). All adjustments discussed above try to lead a better alignment between HFCE and HBS, not only in terms of amounts but especially in terms of definitions and concepts: the more these two domains are close in definitions and concepts, more is reasonable using the available HBS distributional information in the National account framework (Zwijnenburg *et al.*, 2021).

The incoming step is to analyse the estimated household consumption expenditure by quintiles - according to the equivalised sum of the HBS variables related to monetary net income plus imputed rent - and by socio-demographic characteristics (Chancel *et al.*,2021) .

References

BLANCHET T., CHANCEL L., GETHIN A. 2019. How Unequal Is Europe? Evidence from Distributional National Accounts, 1980-2017. *WID.world Working Paper*, No. 2019/06.

CHANCEL L., ETROPOLSKA V., FIXLER D., JAYYOUSSI H., JOHNSON D., MARTINS M., NERI A., SABELHAUS J., SENEVIRATNE A., SOLA P., TODOROVA A., TONKIN R., VALDES MARTINEZ J., WEBBER D. AND ZWIJNENBURG J. 2021. *Guidance note on Distribution of household income, consumption and wealth* - Version: July 2021.

COLI A., TARTAMELLA F. 2017. Developing Households' sub-sectors accounts: Pros and cons of the top-down and the bottom-up methods. *Statistical Journal of the IAOS*, Vol.33, pp. 525–535.

COLI A., ISTATKOV R., JAYYOUSI H., OEHLER F., TSIGKAS O. 2022. Distributional national account estimates for household income and consumption: methodological issues and experimental results, *Final report of the Expert Group on Disparities in a National Accounts Framework*, *third exercise*, Eurostat and OECD. Publication Office of the European Union, Luxembourg.

EUROSTAT 2013. European system of accounts ESA 2010, Luxembourg.

FESSEAU M. AND VAN DE VEN P. 2014. Measuring inequality in income and consumption in a national accounts framework. *Statistics Brief, OECD* No. 19.

LUSTIG N. 2018. Measuring the Distribution of Household Income, Consumption and Wealth: State of Play and Measurement Challenges. *Working Papers* No.1801, Tulane University, Department of Economics.

OECD 2020. Distributional Information on Household Income, Consumption and Saving in Line with National Accounts – Guidelines (Version December 2020).

UNITED NATIONS STATISTICS DIVISION (UNSD) 2000. Classifcation of Individual Consumption According to Purpose (COICOP) 1999. *Statistical Papers Series M* No. 84.

ZWIJNENBURG J., BOURNOT S., GRAHN D. AND GUIDETTI E. 2021. Distribution of household income, consumption and saving in line with national accounts – Methodology and results from the 2020 collection round. *OECD Statistics Working Papers*, No. 2021/01.

ZWIJNENBURG J. 2022. The Use of Distributional National Accounts in Better Capturing the Top Tail of the Distribution. *The Journal of Economic Inequality* No.20, pp. 245–254.

Sara BASSO, Istat, basso@istat.it
Incoronata DONNARUMMA, Istat, donnarum@istat.it
Stefania MASSARI, Istat, stmassar@istat.it

AN APPROACH FOR DETERMINING AREAS OF SOCIO-ECONOMIC DEPRIVATION AT SUB-MUNICIPAL LEVEL

Marco Ballin, Gianpiero Bianchi, Giancarlo Carbonetti, Paolo Lorusso

Abstract. Municipalities have a growing need for detailed and interpretable data to implement specific measures to reduce socio-economic exclusion and deprivation while improving urban quality and promoting the economic, social, and cultural development of urban areas of the resident population. Traditional administrative boundaries may not represent the true spatial distribution of socio-economic phenomena, limiting the effectiveness of such interventions.

This study proposes a parameterized procedure for aggregating small, contiguous spatial units, such as census enumeration areas, into larger, homogeneous domains. With highly detailed demographic, social and economic data, this approach provides in-depth understanding of complex urban issues. The approach is applied to a case study focusing on socio-economic deprivation at sub-municipal level, which considers multiple dimensions such as economic distress, employment instability and low educational attainment. These factors contribute in different ways to the social exclusion of individuals.

The analysis uses results from the population and housing census and data from administrative registers collected by Italian National Institute of Statistics (Istat). The different aspects are synthesized into a single metric using the composite AMPI index at the urban enumeration area level. The results highlight the potential of data-driven spatial aggregation to support effective urban planning and decision-making, revealing patterns that go beyond conventional administrative partitions.

1. Introduction

The availability of data and indicators on individuals and households referring to the minimum territorial unit, coinciding with the census Enumeration Areas (*EAs*), has in the past offered the possibility of determining critical areas within municipalities - aggregations of *EAs* - according to criteria of statistical homogeneity and spatial contiguity.

These areas were identified by the municipalities for specific local policy planning objectives or to access funding provided by the central government to contrast social exclusion, unemployment of the weakest groups (young people, women), to favour local entrepreneurship with various forms of tax exemptions.

Some examples are the "*Zone Franche Urbane*" or the "*Aree urbane degradate*" defined through specific regulatory measures[1] and defined on a predefined set of indicators based on the results of the *EA*-level census. The design of such areas was often carried out heuristically from the most critical *EAs* by defining a spatial cluster according to specific size limits in terms of resident population.

Several well-established algorithmic frameworks exist for aggregating contiguous spatial units into internally homogeneous regions. SKATER, Spatial 'K'luster Analysis by Tree Edge Removal, (Assunção *et al.*, 2006) builds a minimum spanning tree across spatial units connected by adjacency, weighted by attribute dissimilarity, and partitions the tree by pruning edges to yield contiguous, homogeneous regions. Another notable approach that can be used is the Max-p regionalization algorithm (Duque *et al.*, 2012). Max-p determines the largest possible number of regions under constraints, such as minimum size and contiguity, by optimizing internal homogeneity using heuristic or mixed integer strategies. More recently, spatially constrained spectral clustering has emerged, embedding spatial contiguity within a similarity graph, and applying spectral partitioning to balance spatial and attribute coherence. Yuan (Yuan *et al.*, 2015) propose a version using a truncated exponential kernel and recursive partitioning to produce nested coherent regions.

With this study it is proposed an intuitive and parameterized procedure for aggregating *EAs*, into larger, homogeneous domains. With highly detailed demographic, social and economic data, this approach provides in-depth understanding of complex urban issues.

The paper is organized as follows: section 2 presents the Istat project on socio-economic deprivation at the sub-municipal level; section 3 illustrates the algorithm developed for the identification of critical areas; section 4 presents some results for the municipality of Bologna.

2. The study of socio-economic deprivation at sub-municipal level

Istat, in collaboration with several municipalities, is conducting a project to study socio-economic deprivation of population at sub-municipal level (Carbonetti *et al.*, 2025). This project makes use of the data available from the permanent population census, administrative sources and thematic registers developed by Istat in recent years, as well as the ability to integrate these sources and geocode information down

---

[1] *"Zone Franche Urbane" – Legge 27 dicembre 2006, n. 296 (Finanziaria 2007), art. 1 commi 340-343). "Aree urbane degradate" – Decreto del Presidente del Consiglio dei Ministri del 15 ottobre 2015.*

to the census *EAs*. The main objective is to measure, represent and analyse the phenomenon of socio-economic deprivation of individuals and households at sub-municipal level. This is a multidimensional phenomenon defined as "*a condition in which households and individuals experience difficulties in adequately meeting their basic needs due to a lack of or insufficient economic, employment, educational, and social resources and opportunities*".

Based on the availability of data, nine statistical ratio indicators (variables) covering different components of deprivation (economic, occupational, and educational) were defined and calculated at *EA* level:

▪ *Economic deprivation*: % of individuals aged 70 and over living alone and not owning a home (ECO1); % of individuals in households in which no member is employed or receiving a pension from work (ECO2); % of individuals in low equivalized income households (ECO3).

▪ *Occupational deprivation*: Employment rate 25-64 years old (OCC1); % of individuals aged 0-64 living in households with very low work intensity (OCC2); % of employed persons aged 25-64 "not stable" during the year (OCC3).

▪ *Educational deprivation*: % of individuals aged 25-64 without upper secondary school education (EDU1); % of individuals aged 15-29 who are not employed and are not enrolled in any regular course of study (EDU2); % of students dropping out or repeating the year (EDU3).

Table 1 describes the numerators and denominators of the nine aggregate variables and their sources.

Subsequently the nine statistical ratio indicators (e.g. OCC1=Employed persons aged 25-64/Individuals aged 25-64) have been synthesized, through the Adjusted Mazziotta-Pareto Index (AMPI$^+$) methodology (Mazziotta e Pareto, 2016; Mazziotta e Pareto, 2017), in a composite index at the *EA*-level. Such index, the Socio-Economic Deprivation Index of population (*SED-I*), measures household and individual deprivation as a tangible condition of deprivation, as distinct from simple exposure to risk[2]. In summary, *SED-I* is a partially non-compensatory composite index based on a standardization of the individual indicators, at the reference time, that makes the indicators independent from the unit of measure (De Muro *et al.*, 2011). Details about AMPI$^+$ methodology and *SED-I* computation are given in paragraph 3.1.

[2] Istat already produces an indicator which measures social and material vulnerability of population. The Social and Material Vulnerability Index measures the exposure of some population groups to situations of risk, such as uncertainty of their social and economic condition (Istat, 2020).

Table 1 – *Numerators and denominators of the nine statistical ratio indicators related to socio-economic deprivation.*

Variables	Numerator	Denominator	Sources
ECO1	Individuals aged 70 and over living alone and not owning a home	Individuals aged 70 and over	PPHC Cadaster Register
ECO2	Individuals in households in which no member is employed or receiving a pension from work	Individuals in households	PPHC Pensioners' records (INPS)
ECO3	Individuals in low equivalized income households	Individuals in households	PPHC Income Register
OCC1	Employed persons aged 25-64	Individuals aged 25-64	PPHC
OCC2	Individuals aged 0-64 living in households with very low work intensity	Individuals aged 0-64 living in households	PPHC Labour Register
OCC3	Employed persons aged 25-64 "not stable" during the year	Individuals aged 25–64 living in households with a sign of employment during the year	PPHC Labour Register
EDU1	Individuals aged 25-64 without upper secondary school education	Individuals aged 25-64	PPHC
EDU2	Individuals aged 15-29 who are not working and are not registered in any regular course of study at MIM or MUR	Individuals aged 15-29	PPHC MIM – MUR
EDU3	Students dropping out or repeating the year	Students with attendance records in the academic year t/t+1	PPHC Education Register

Sources: Istat, Permanent Population and Housing Census (PPHC); Income register, Labour register and Education register; National Social Security Institute (INPS); Ministry of Education (MIM); Ministry of University and Research (MUR); Register of dwellings and buildings (Cadastre). Reference year: 2021.

All measures (indicators and *SED-I*) are calculated at *EAs*, but analyses will be conducted for two levels of *EA* "aggregations" to suggest an integrated reading of the results:

- *SMA* – administrative Sub-Municipal Areas. Areas defined by municipalities for functional, administrative, or statistical purposes.
- *ADU* – Areas of socio-economic deprivation in urban contexts. Areas never previously identified, designed by Istat according to the experimental methodology proposed in this document to identify concentrations of deprivation within municipal boundaries.

The results of the deprivation study, at *SMA* and *ADU* level, represent a powerful knowledge and decision-making tool for municipal administrations. The choice of the level of analysis and intervention depends on political choices and economic availability, factors that may vary between different municipalities and between different administrations within the same municipality.

3. The *SED-I* computation and *ADU* identification procedure

The method for identification of *ADU* is based on a spatial aggregation procedure designed to identify and group contiguous territorial units with similar characteristics. The procedure is designed to be highly flexible and adaptable to different spatial or policy needs through a set of configurable parameters:

- P_{min} , P_{max}: minimum and maximum population per *ADU*. If the total population does not exceed the minimum threshold, the aggregation is discarded; when the maximum threshold is reached, the algorithm stops the aggregation procedure.
- P_{tot}: global population threshold. It defines the amount of population to be included in the whole set of *ADU*.
- Min_{EA}: minimum number of *EAs*. Minimum number of *EAs* to be included in a *ADU*.
- ε: *SED-I* homogeneity threshold. It sets the maximum deviation allowed from the area under construction and an additional territorial unit to be included.
- τ: minimum *SED-I* threshold (*SED-I* floor). Minimum composite indicator value that an aggregation must maintain to be validated. This constraint prevents the inclusion of less deprived sections from compromising the significance of the aggregations.

Since the procedure is heuristic and data-driven, different combinations of parameters can lead to different solutions, all of which are methodologically legitimate but have different operational implications, for instance:

- smaller values on homogeneity ε favor the construction of *ADU* more coherent with respect to *SED-I*, but geographically more fragmented;
- low demographic thresholds favor defining micro-areas of intervention, high thresholds lend themselves to larger-scale planning;
- a high *SED-I* threshold identifies only the most impaired areas, while a lower one allows more extensive coverage, including situations of moderate distress.

3.1. SED-I computation

Let the *EAs* be indexed by $i=1,..,N$ where N indicates the amount of *EA* in a given municipality. Each *EA* is characterized by K ($K=9$ in the case of this paper; see Table 1) statistical ratio indicators $R_i^{(k)}$ (i=1,..,N ; $k=1,..,K$). It is assumed that for each *EA* the population, $Pop(EA_i)$, and all numerator and denominator pairs ($NUM_i^{(k)}$; $DEN_i^{(k)}$) necessary for the K indicators computation are available.

The k^{th} indicator ratio in the i^{th} *EA* is defined as:

$$R_i^{(k)} = \frac{NUM_i^{(k)}}{DEN_i^{(k)}} \text{ assuming } \frac{0}{0} = 0 .$$

In order to compute the *SED-I* composite index, the first step consist in the "normalization" of the $R_i^{(k)}$. Normalization is computed by the following:

$$Z_i^{(k)} = 70 + 60 * \left(\frac{R_i^{(k)} - a_k}{b_k - a_k}\right) \qquad \text{(if } a_k \neq b_k \text{, 70 otherwise)}$$

where: $a_k = \mu_k - \frac{1}{2} R_{max}^{(k)}$, $b_k = \mu_k + \frac{1}{2} R_{max}^{(k)}$ (a_k and b_k are called ‘‘goalposts’’), $\mu_k = \frac{\sum_{i \in S} \mathrm{NUM}_i^{(k)}}{\sum_{i \in S} \mathrm{DEN}_i^{(k)}}$ and $R_{max}^{(k)} = \max_{i \in 1,..,N} R_i^{(k)}$.

If an indicator has negative polarity, it is transformed as $Z_i^{(k)} = 200 - Z_i^{(k)}$ (as in the case of indicator OCC1).

The *SED-I* index for *i*[th] *EA* is defined as:

$$SED\text{-}I_i = \bar{Z}_i + \sigma_i \mathrm{CV}_i$$

Where $\bar{Z}_i = \frac{\sum_{k \in K} Z_i^{(k)}}{k}$, $\sigma_i = \sqrt{\frac{\sum_{k \in K} (Z_i^{(k)} - \bar{Z}_i)^2}{k}}$ and $\mathrm{CV}_i = \frac{\sigma_i}{\bar{Z}_i}$.

3.2. Algorithm description

The aggregation algorithm performs an iterative greedy region-growing process. At the very beginning the algorithm starts considering the whole set of *EAs*. The general iteration performed for the construction of an *ADU* can be described as follow:

1. Let *U* be the set of all territorial units not yet belonging to an *ADU*. Let $U^{(max)}$ be the unit with maximal *SED-I* in *U*;
2. indicate by $C\left(U^{(max)}\right)$ the set of units candidate to be aggregated to $U^{(max)}$. $C\left(U^{(max)}\right)$ is defined as the set all *EAs* adjacent to $U^{(max)}$ belonging to *U* and with $SED\text{-}I \geq (1 - \varepsilon) * SED\text{-}I\left(U^{(max)}\right)$. Let $C^{(u)}$ be an element of $C\left(U^{(max)}\right)$;
3. form aggregated areas ($AG^{(u)}$) by collapsing $U^{(max)}$ with each $C^{(u)}$;
4. for each $AG^{(u)}$ compute the total population, *Pop*($AG^{(u)}$), and all pairs $\left(NUM_u^{(k)}; DEN_u^{(k)}\right)$ by adding the correspondent elements of $U^{(max)}$ and $C^{(u)}$. Using such quantities compute $SED\text{-}I^{(AG)}$ for each $AG^{(u)}$ following the procedure described in paragraph 3.1;

5. carry out the following checks on each $AG^{(u)}$:
   a. $Pop(AG^{(u)}) \leq P_{max}$ (size of $AG^{(u)}$ can't be over a fixed threshold);
   b. $SED\text{-}I^{(AG)} \geq \tau$, (*SED-I* minimum threshold for a *ADU*);
6. among the $AG^{(u)}$ that meet the two conditions, choose the one with the maximum *SED-I*, let $AG^{(max)}$ be such elements;
7. create the set $C\left(AG^{(max)}\right)$ as described in step 2.

Repeat the steps 3-7, using $AG^{(max)}$ instead of $U^{(max)}$, until at least a new $AG^{(u)}$ meets the two conditions of step 5 and until $C\left(AG^{(max)}\right)$ is not empty.

If after last iteration of steps 3-7 the last $AG^{(max)}$ satisfies the following conditions:

- $Pop\left(AG^{(max)}\right) \geq P_{min}$;
- Number of *EAs* included in $AG^{(max)} \geq Min_{EA}$;

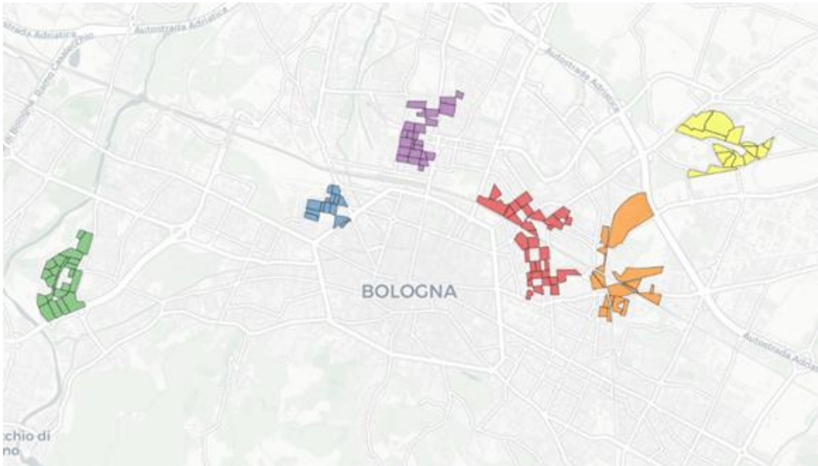
then $AG^{(max)}$ is a new *ADU*. If $AG^{(max)}$ is accepted as a new *ADU* delete all *EAs* forming such new area from *U* and repeat the algorithm from step 1.

If at least one of the last two conditions is not met, $AG^{(max)}$ is not accepted as a new *ADU*. In such case the procedure re-starts from the following *EA* with the greatest *SED-I* in *U*. If adding the population of previous *ADU* and $AG^{(max)}$ is greater than P_{tot} the algorithm, stops. Furthermore, the algorithm stops when all starting point have been tested and no additional *ADU* can be created.

4. Results

To demonstrate the effectiveness of the proposed procedure and to highlight the critical role of parameter selection, we conducted a series of simulations aimed at identifying aggregations characterized by significant socio-economic disease patterns in the city of Bologna. In the first simulation, the parameters were defined as follows: the homogeneity tolerance range (ε) was set to 0.1; the global population P_{tot} threshold was established at 38,000, corresponding to approximately 10% of the total population in the census *EAs* of the inhabited center; each aggregation was required to have a population between P_{min}=3,000 and P_{max}=7,500; the *SED-I* floor was set to 100 plus the mean squared error of the *SED-I* distribution across *EAs*, resulting in a threshold value of 103.69 (Figure 1); Min_{EA} has been set to 5.

Figure 1 – *Bologna: SED-I floor=103.69; Population 3,000 – 7,500 (Istat, 2021).*



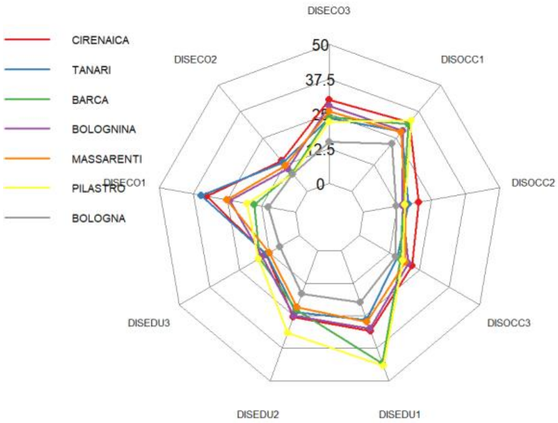
Six critical aggregations were identified, roughly corresponding to the neighborhoods of the following disctricts of Bologna: "Cirenaica", "Tanari", "Barca", "Bolognina", "Massarenti", and "Pilastro". A summary (Table 2) of these aggregations (*ADUs*) is provided: population size (POP), final *SED-I* value (*SED-I*), starting section *SED-I* (START *SED-I*), total number of sections (N.of *EAs*) included, and the number of sections with an *SED-I* value above 110 and 105, respectively.

Table 2 – *Bologna: SED-I floor=103.69; Population 3,000 – 7,500. ADU's summary (Istat, 2021).*

ADU	DISTRICT	POP	SED-I	START SED-I	N. of EAs	N. of EAs SED-I>110	N. of EAs SED-I>105
1	Cirenaica	7,323	106.40	116.09	31	8	17
2	Tanari	3,028	104.70	113.77	15	2	3
3	Barca	4,438	104.55	112.51	21	1	8
4	Bolognina	7,413	104.67	112.42	24	3	10
5	Massarenti	3,121	104.05	112.30	18	1	4
6	Pilastro	5,484	105.52	108.79	14	0	9

In the "radar chart" (Figure 2), we compare how the nine indicators of the six *ADUs* behave relative to the overall values for the municipality of Bologna.

Figure 2 – *Bologna: SED-I floor=103.69; Population 3,000 – 7,500. ADU's indicators (Istat, 2021).*



As expected, all the aggregations show higher values than the city average. The most critical issues emerge in terms of EDU1 and EDU2, where the "Pilastro" area stands out as the most critical, also recording the highest value for EDU3. The "Barca" area follows closely in terms of EDU1, while for ECO1, both "Cirenaica" and "Tanari" show values nearly four times higher than the city average.

To assess the flexibility of the procedure in terms of its applicability to different contexts, the analysis has been repeated using as population thresholds P_{min}=1,000 and P_{max}=5,000 residents (results are shown in Figure 3 and Table 3).

Figure 3 – *Bologna: SED-I floor=103.69; Population 1,000 – 5,000 (Istat, 2021).*

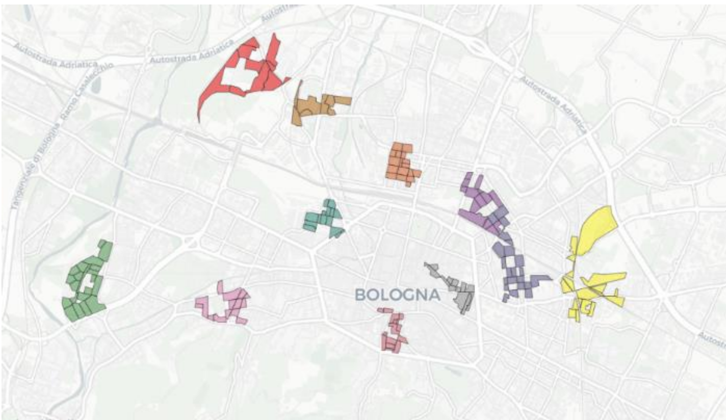
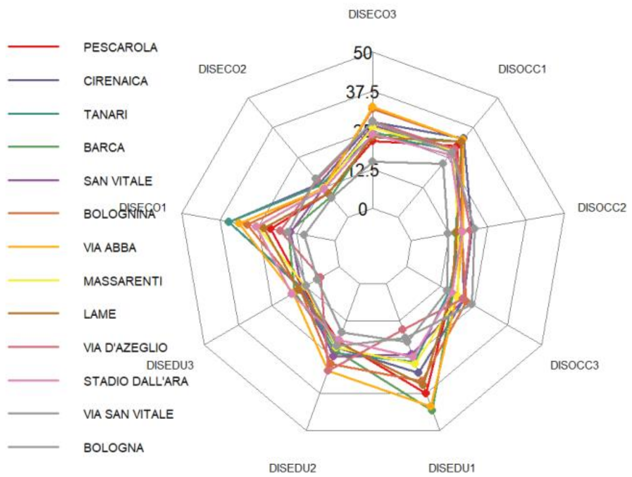


Table 3 – *Bologna: SED-I floor=103.69; Population 1,000 – 5,000. ADU’s summary (Istat, 2021).*

ADU	DISTRICT	POP	SED-I	START SED-I	N. of EAs	N. of EAs SED-I>110	N. of EAs SED-I>105
1	Pescarola	2,671	103.89	118.55	10	1	3
2	Cirenaica	4,637	106.06	116.09	22	7	10
3	Tanari	3,028	104.70	113.77	15	2	3
4	Barca	4,438	104.55	112.51	21	1	8
5	San Vitale	3,168	103.70	112.46	14	2	5
6	Bolognina	4,772	106.19	112.42	17	3	10
7	Via Abba	1,690	107.17	112.34	10	2	7
8	Massarenti	3,121	104.05	112.30	18	1	4
9	Lame	1,618	104.18	111.74	7	1	2
10	Via d'Azeglio	1,558	103.75	110.65	15	1	4
11	Stadio Dall’Ara	2,722	103.75	110.17	14	2	6
12	Via San Vitale	2,344	103.75	110.05	13	2	4

In the “radar chart” (Figure 4), the values of the nine indicators for each of the twelve *ADUs* are compared to the overall values for Bologna. Interestingly, the aggregation "Via d’Azeglio" shows lower values than the city average for indicators EDU1 and EDU3, despite exhibiting significantly higher-than-average values for the other seven indicators. It records the highest value for indicator EDU2, highlighting a specific critical issue among the 15–29 age group.

Figure 4 – *Bologna: SED-I floor=103.69; Population 1,000 – 5,000. ADU’s indicators (Istat, 2021).*



5. Conclusions

The paper describes a methodology and proposes a procedure for its implementation to identify sub-municipal areas that complement the minimum territorial units coinciding with the census enumeration areas (*EAs*). These new areas are created by aggregating contiguous and homogeneous *EAs*.

In the case analysed in this study, homogeneity is assessed with respect to a composite index of socio-economic deprivation (*SED-I*) developed by Istat, based on nine elementary indicators, and reflecting the level of economic, employment, educational, and social deprivation of the population residing in a specific area.

The area construction methodology is sequential. The procedure begins by selecting the *EA* with the highest observed socio-economic deprivation and continues by aggregating neighbouring homogeneous *EAs* around it. The area construction proceeds if there are contiguous *EAs* sufficiently homogeneous with the area already formed, or if it does not exceed predetermined size thresholds. Once the first area is completed, the process resumes by excluding already assigned *EAs* from further analysis. The process stops when there are no more *EAs* that meet the eligibility criteria for aggregation, or when the aggregate of constructed areas exceeds a predetermined fraction of the municipal population.

The objective is to identify the areas of the municipal territory where the highest concentrations of socio-economic deprivation of the population are observed.

One of the strengths of the procedure lies in the intuitiveness of the required parameterisations: the minimum degree of homogeneity; the minimum level of overall deprivation of each area; the minimum and maximum size of each area; the maximum fraction of municipal population that can be included in the set of critical areas. In addition to describing the procedure in detail, the document presents some examples for the city of Bologna.

Future developments will focus on the identification of a set of parameterisations applicable to groups of municipalities homogeneous with respect to their demographic size, as well as on the exploration of potential alternative forms of parameterisation of the procedure.

An important aspect will be the definition of evaluation criteria for the resulting areas. Currently, the validation of the results is conducted together with the municipalities, which are the final users of the results. The identification of these areas will in fact offer the possibility to define targeted policy actions on these areas, also considering the socio-demographic characteristics of the resident individuals and households.

References

ASSUNCÃO R. M., NEVES M. C., CÂMARA G., DA COSTA FREITAS C. 2006. Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. I*nternational Journal of Geographical Information Science*, Vol. 20, No. 7, pp. 797-811.

CARBONETTI G., BIASCIUCCI F., CUTILLO A., MAZZIOTTA M., QUONDAMSTEFANO V., TAMBURRANO M. T., TRONU. D. 2025. Measuring socio-economic deprivation at sub-municipal level through the integration of census and administrative data. *Italian Journal of Economic, Demographic and Statistical Studies – RIEDS*, Vol. LXXIX No.1 Gennaio-Marzo 2025.

DE MURO P., MAZZIOTTA M., PARETO A. 2011. Composite Indices of Development and Poverty: An Application to MDGs. *Social Indicators Research*, Vol. 104, No. 1, pp.1-18.

DUQUE J. C., ANSELIN L., REY S. J. 2012. The max-p-regions problem. *Journal of Regional Science*, Vol.52, No. 3, pp. 397-419.

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*, Springer, pp.159-191.

MAZZIOTTA M., PARETO A. 2016. On a Generalized Non-compensatory Composite Index for Measuring Socio-economic Phenomena. *Social Indicators Research*, Vol. 127, No. 3, pp. 983-1003.

YUAN S., TAN P. N., CHERUVELIL K. S., COLLINS S. M., SORANNO P. A. 2015. Constrained spectral clustering for regionalization: Exploring the trade-off between spatial contiguity and landscape homogeneity. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 1-10). IEEE.

Marco BALLIN, Italian National Institute of Statistics, ballin@istat.it
Gianpiero BIANCHI, Italian National Institute of Statistics, gianbia@istat.it
Giancarlo CARBONETTI, Italian National Institute of Statistics, carbonet@istat.it
Paolo LORUSSO, Italian National Institute of Statistics, paolo.lorusso@istat.it

STATISTICS VS MACHINE LEARNING: AN APPLICATION ON TIME-TO-EVENT DATA IN TERRORIST KIDNAPPING EVENTS

Luciano Nieddu, Cecilia Vitiello

Abstract. Hostage-taking durations in terrorist attacks using the Global Terrorism Database have been considered in this paper using data from 1970 to 2021. Employing conditional inference trees and Cox proportional hazards models, we managed to determine factors influencing hostage release time. Attack type and ransom demands significantly prolong incident duration, with effects varying over time. Regional variations have also been detected, with Middle East & North Africa and Southeast Asia showing longer median durations. The complementary insights from machine learning providing a clustering of individuals based on survival time and statistical methodologies which provide a clustering of the effects of the covariates on the instantaneous risk of surviving yield a robust framework for understanding complex events such as kidnapping.

1. Introduction

The Global Terrorism Database (GTD) (LaFree and Dugan, 2007; START, 2022) is one of the most extensive and systematically compiled longitudinal dataset on terrorist events. GTD covers events from 1970 to 2021, except for 1993 due to data loss, and includes over 214000 recorded attacks. The database is constructed from open-source materials such as media reports and official documents and provides information on various dimensions of each incident, covering event-level information such as the date, location, weapons used, target type, and number of fatalities and injuries. The GTD has been widely employed in empirical research on terrorism, including statistical modeling of temporal and spatial trends (Dugan et al, 2005; Phillips, 2011), and assessments of counter-terrorism policy effectiveness (Enders and Sandler, 2012). Several studies have used GTD data to investigate tactical selection and target profiles (Deloughery et al. 2012), the impact of terrorist negotiations and concessions (Piazza, 2017), and the effectiveness of rescue operations and government responses (White et al 2014, Wilson, 2000).

The longitudinal nature of GTD makes it particularly well-suited for time series modeling and survival analysis. In this study, we consider the hostage subset of GTD records to quantify risk factors and detect patterns in kidnapping events over time and across geographic regions. Several studies have examined the duration of kidnappings: Brandt and Sandler (2009) analyzed the dynamic properties of hostage-taking series using time series methods, finding that past concessions influence future kidnapping events. Kim and Sandler (2021) applied survival analysis to a dataset of terrorist hostage-taking incidents from 1978 to 2018 identifying key determinants that prolong the duration of hostage situations (i.e. demanding release of imprisoned comrades). Gaibulloev and Sandler (2009) investigates the determinants of logistical and negotiation success in terrorist hostage-taking missions finding that incident durations are positively associated with negotiation success. Phillips (2014) used survival analysis to investigate the factors contributing to the varying longevity of terrorist groups.

This paper uses machine learning and semi parametric survival analysis techniques to assess the instantaneous risk of a hostage release as a function of several structural characteristics of the attack.

2. Data and Methods

Out of the total number of events only those classified as “hostage kidnapping” (i.e. held against their will) or “kidnapped” (i.e. held against their will and taken to another location) have been selected resulting in a dataset of 17161 cases. Duration of the kidnapping has been measured in hours since the incident. Cases where the duration was not available have been removed from the dataset resulting in a dataset of 5534 cases. Events have been right-truncated to the 95th percentile of all the times resulting in a dataset of 5258 observations. Events happening at time zero (240) or with a negative duration (3) were removed resulting in a dataset of 5015 records.

The event has been defined as the release of the hostage within the observational period. Hostages not released or killed have been considered as censored. In each attack the number of persons taken hostages and the number of persons released have been used as weights for the analysis. Those cases where the weight was not available were removed (162) resulting in a final dataset of 4853 cases.

The survival function $S(t)$, modelling the time-to-event of the hostage, has been estimated via the product limit Kaplan Maier method (Kaplan and Meier, 1958) and the semi-parametric Cox proportional hazards model (Cox, 1972). The results of these statistical techniques have been compared with the machine learning approach based on survival trees (Hothorn et al. 2006; Hothorn and Zeileis 2015). The variables that have been considered int the models are:

- Attack Type (attacktype): primary tactical method of the incident (e.g., armed assault, bombing, hostage-taking).
- Region (region): geographical region where the incident occurred.

- Ransom (ransom): binary variable indicating whether a monetary ransom was demanded during the incident.
- Suicide (suicide): binary variable indicating if the attack was a suicide operation.
- Success (success): binary indicator of the attack's outcome, denoting whether the violent act was executed as intended (e.g., a bombing is successful if the device detonates; an assassination is successful only if the specific target is killed).

The proportionality assumption for the hazard related to the Cox model has been assessed via the Shoenfeld residuals. For those variables whose effect does not meet the proportionality assumption, time-dependent effects were included in the model by introducing the interactive effect between the variables themselves and the log(t).

2.1. Kaplan Meier Product limit estimator

The Kaplan-Meier (KM) product limit estimator is a nonparametric method for estimating the survival function from time-to-event data in the presence of censored observations. Developed by Kaplan and Meier (1958), the estimator computes the probability of survival beyond specific time points by multiplying conditional survival probabilities observed at each distinct event time. It does not require any assumptions about the underlying hazard function or the effect of covariates, making it especially useful for descriptive analysis. Censored observations are accounted for by reducing the number at risk when the censoring occurs without contributing to the event count.

$$\hat{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right)$$

Where:

- t_i represents the ordered distinct event times $(t_1 < t_2 < \cdots < t_k)$.
- d_i is the number of events (deaths or failures) observed at time t_i.
- n_i is the number of individuals at risk (alive and uncensored) just prior to time t_i.

2.2. Cox Model

The Cox proportional hazards model is a semi-parametric method used to analyze time-to-event data, originally introduced by Cox (1972). It estimates the hazard function, i.e., the instantaneous risk of the event, conditional on a set of covariates. The model assumes that covariates have a multiplicative effect on the baseline hazard and that these effects are constant over time (proportional hazards assumption).

$$h(t) = h_0(t) \cdot e^{X\beta}$$

Estimation is performed via partial likelihood, which allows the baseline hazard to remain unspecified while estimating the regression coefficients. The estimates of the β coefficients are the conditional effect of each single covariate on the hazard on the log scale.

2.3. Survival Trees

Survival trees used in this paper are based on conditional inference trees. The method is based on the conditional inference framework developed by Hothorn et al (2006), which uses statistical hypothesis testing to grow a binary tree. The global null hypothesis of independence between each covariate and the survival response is evaluated using the log-rank test. If a significant association is detected, the variable with the strongest evidence against the null hypothesis is selected, and the best binary split is identified to maximize the separation in survival experiences between resulting groups. The recursive process continues until no covariate exhibits a significant relationship with the survival outcome. At each terminal node, the model yields an estimated survival function (Kaplan-Meier) for the statistical units assigned to that terminal node. Conditional inference trees for survival data offer a robust, nonparametric method to identify covariates influencing time-to-event outcomes, especially in the presence of complex interactions and non-linear effects.

3. Results

From the KM estimator, the median event durations provide a robust measure of central tendency (see Table 1), less susceptible to the influence of extreme values. Central Asia, Middle East & North Africa (MENA), and Sub-Saharan Africa all report a median duration of 120 hours (Table 2). This group primarily consists of regions that have historically experienced complex geopolitical conditions, internal conflicts, or varying degrees of institutional instability. Numerous regions, such as Australasia & Oceania, Central America & Caribbean, and East Asia, present a median duration of 24 hours, suggesting a prevalence of relatively short events.

Table 1 – *KM statistics for survival time.*

N	n.start	events	rmean	se(rmean)	median	LCL	UCL
38476	38476	32685	335.46	4.58	72	72	72

The median survival time for the entire dataset is 72 hours from the kidnapping (see Table 1). In Figure 1 the survival tree output has been displayed. The four terminal nodes identify 4 clusters of regions with homogenous behavior according to the time-to-event data. Node 6 (MENA and Southeast Asia) is the one presenting the cases more at risk of presenting a long duration of the kidnapping compared to the other 3 clusters. The cluster showing the fastest resolution time is represented by node 3 (Australasia and Oceania, Central America and Caribbean, Central Asia, East Asia, Western Europe).

Table 2 – *KM statistics using region as covariate.*

Region	N	n.start	events	rmean	se(rmean)	median	0.95LCL	0.95UCL
A/O	47	47	44	73.02	15.20	24	16	24
CA/Car	488	488	459	98.45	14.39	24	24	24
C.Asia	243	243	218	197.43	32.91	120	72	120
E.Asia	275	275	274	24.18	0.27	24	24	24
E.E.	2847	2847	2330	152.76	10.66	72	72	72
MENA	8151	8151	6002	698.89	16.92	96	96	96
N.Am.	647	647	376	93.44	19.54	30	30	48
S.Am.	3082	3082	2816	315.49	14.84	48	48	48
S.Asia	9346	9346	7872	215.50	6.69	48	48	48
SE.Asia	2866	2866	2673	506.01	19.22	48	48	72
SSA	8745	8745	7995	270.16	6.15	120	120	120
W.E.	1739	1739	1626	179.64	9.65	37	37	37

Note: Australasia & Oceania (A/O), Central America & Caribbean (CA/Car), Central Asia (C.Asia), East Asia (E.Asia), Eastern Europe (E.E.), Middle East & North Africa (MENA), North America (N.Am.), South America (S.Am.), South Asia (S.Asia), Southeast Asia (SE.Asia), Sub-Saharan Africa (SSA), Western Europe (W.E.)

Figure 1 – *output of survival tree using region as covariate.*

```
[1] root
|   [2] region in A/O, CA/Car, C.Asia, E.Asia, E.E., N.Am., SSA, W.E.
|   |   [3] region in A/O, CA/Car, C.Asia, E.Asia, W.E.: 72.000 (n = 279)
|   |   [4] region in E.E., N.Am., SSA: 120.000 (n = 987)
|   [5] region in MENA, S.Am., S.Asia, SE.Asia
|   |   [6] region in MENA, SE.Asia: 288.000 (n = 1434)
|   |   [7] region in S.Am., S.Asia: 192.000 (n = 2153)

Number of inner nodes:    3
Number of terminal nodes: 4
```

Note: Australasia & Oceania (A/O), Central America & Caribbean (CA/Car), Central Asia (C.Asia), East Asia (E.Asia), Eastern Europe (E.E.), Middle East & North Africa (MENA), North America (N.Am.), South America (S.Am.), South Asia (S.Asia), Southeast Asia (SE.Asia), Sub-Saharan Africa (SSA), Western Europe (W.E.)

The Cox proportional hazards model provides a quantification of the effect on the log-hazard of each value of the covariate while the survival tree provides a clustering of the regions based on the survival probability. Estimates for the Cox model using region as covariate have been reported in Table 3.

Comparing the groups of regions obtained by the survival tree approach (Figure 1 and 2) with the effect on the log hazard (Table 3) we observe:

- Node [3]: fastest release group (Median 72 hours): Australasia & Oceania (Reference), Central America & Caribbean, Central Asia, East Asia, Western Europe. *The Cox model shows that regions in this group show non significant effect on the log(HR)*
- Node [4]: moderately fast release Group (Median 120 hours): Eastern Europe, North America, Sub-Saharan Africa. *The Cox model's results for these regions present moderately negative log(HR) values that are either statistically significant or borderline.*
- Node [7]: fairly slow-release group (Median 192 hours): South America, South Asia. *The Cox model presents significantly negative log(HR) values for these regions although not the highest observed effects.*
- Node [6]: slow-release group (Median 288 hours): MENA, Southeast Asia. *The regions in this node show the highest and statistically significant negative log(HR) values in the Cox model.*

Table 3 – *Cox proportional hazard model estimates.*

Characteristic	*log(HR)*	*95% CI*	*p-value*
region			
Australasia & Oceania	—	—	
Central America & Caribbean	***-0.57***	***-1.3, 0.20***	***0.150***
Central Asia	***-0.53***	***-1.4, 0.37***	***0.300***
East Asia	***0.58***	***-0.51, 1.7***	***0.300***
Western Europe	***-0.71***	***-1.5, 0.05***	***0.069***
MENA	*-1.3*	*-2.0, -0.55*	*<0.001*
Southeast Asia	*-1.3*	*-2.1, -0.56*	*<0.001*
South America	*-1.2*	*-1.9, -0.43*	*0.002*
South Asia	*-1.1*	*-1.8, -0.36*	*0.004*
Eastern Europe	*-0.87*	*-1.6, -0.10*	*0.026*
North America	*-0.80*	*-1.6, 0.00*	*0.051*
Sub-Saharan Africa	*-0.91*	*-1.7, -0.17*	*0.017*

Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Using the clustering of regions suggested by the survival tree we observe that regions belonging to Node 3 are those regions in the Table 3 that show not significant or marginally significant effects on the risk of the event. Regions in node 4 show significant and average negative effect on the risk to experience the event. Regions from Node 6 show very strong negative significant effect on the risk of experiencing the event and regions in Node 7 show strong middle high negative significant effect

3.1. Extended Model

The goal of this section is to compare the performances of the machine learning approach and the semi-parametric approach on a model where the hazard function depends on more than one variable. Namely the following model has been considered:

$$S(t) = f(Region, Attacktype, Ransom, Suicide, Success)$$

where:

- *Region* is the geographical region where the event took place
- *Attacktype* is a variable detailing the general method of attack. It takes value as “Assassination”, “Hijacking”, “Kidnapping”, “Barricade Incident”, “Bombing/Explosion”, “Armed Assault”, “Unarmed Assault”, “Facility/Infrastructure Attack”, “Unknown”;
- *Ransom* is a binary variable which is 1 iff there is evidence that a ransom was requested
- *Suicide* is a binary variable which is 1 iff there is evidence that the terrorists were planning to commit suicide
- *Success* is a binary variable which is 1 if the attack was successful

The survival tree yields a final partition of the statistical units into nine terminal nodes (Figure 2). The first splitting has been performed using type of attack. while the following split involves different choices of variables implying a different role of the variables within subsets of the upper level variables. This suggests a possible presence of interaction and nonlinear effect of covariates on the outcome.

The survival tree suggests once again a partition of the regions of interest into two major subsets, namely:

- Australasia & Oceania, Central America & Caribbean, Central Asia, East Asia, Eastern Europe, North America, Sub-Saharan Africa, Western Europe
- MENA, South Asia, Southeast Asia

In Figure 2 the median survival times for each terminal node have been reported together with the size of the terminal node; “Inf” values for the median indicate that the cumulative incidence in that cohort of cases remained below 50% making impossible the estimation of the median time. This identifies a small subgroup with a relatively low hazard rate, and the follow-up time is insufficient to observe the point at which half of the cohort has experienced the event.

Figure 2 – *Output of survival trees using all covariates.*

```
[1] root
|   [2] attacktype in Assau, Assas, Kidnap, Other
|   |   [3] region in A/O, CA/Car, C.Asia, E.Asia, E.E., N.Am., SSA, W.E.
|   |   |   [4] attacktype in Assau, Kidnap: 120.000 (n = 1096)
|   |   |   [5] attacktype in Assas: 576.000 (n = 45)
|   |   [6] region in MENA, S.Am., S.Asia, SE.Asia
|   |   |   [7] region in MENA, SE.Asia
|   |   |   |   [8] attacktype in Assau, Assas
|   |   |   |   |   [9] ransom <= 0
|   |   |   |   |   |   [10] region in MENA: Inf (n = 32)
|   |   |   |   |   |   [11] region in SE.Asia: 432.000 (n = 20)
|   |   |   |   |   [12] ransom > 0: 4320.000 (n = 10)
|   |   |   |   [13] attacktype in Kidnap, Other: 288.000 (n = 1302)
|   |   |   [14] region in S.Am., S.Asia: 192.000 (n = 2076)
|   [15] attacktype in Expl, Inf.Att, Hijack, Barr
|   |   [16] attacktype in Expl, Hijack: 48.000 (n = 124)
|   |   [17] attacktype in Inf.Att, Barr: 10.000 (n = 148)

Number of inner nodes:    8
Number of terminal nodes: 9
```

Note: when a median survival time is reported as (Inf) (infinity) for a terminal node, it indicates that the median survival time for subjects in that specific subgroup was not reached during the study's follow-up period.
Regions: A/O: Australasia & Oceania, CA/Car: Central America & Caribbean, C.Asia: Central Asia, E.Asia: East Asia, E.E.: Eastern Europe, MENA: Middle East & North Africa, N.Am.: North America, S.Am.: South America, S.Asia: South Asia, SE.Asia: Southeast Asia, SSA: Sub-Saharan Africa, W.E.: Western Europe.
Attack Types: Assau: Armed Assault, Assas: Assassination, Kidnap: Hostage Taking (Kidnapping), Other: Unknown, Expl: Bombing/Explosion, Inf.Att: Facility/Infrastructure Attack, Hijack: Hijacking, Barr: Hostage Taking (Barricade Incident)

The large medians in some subgroups (4320 hours in Node 12 for instance) can be attributed to the recursive partitioning of the covariate space: the tree algorithm has isolated a very specific and small subset of the data (n=10 for Node 12) defined by the pathway leading to that node. This partition identifies a cohort with a relatively low hazard rate.

The Cox model has been fitted starting from the null model with just the intercept and using a stepwise forward selection strategy to improve the model fitting using penalized likelihood (AIC).

$$h(t|X) = h_0(t)e^{(\beta_1 \cdot \text{attacktype} + \beta_2 \cdot ransom + \beta_3 \cdot \text{region} + \beta_4 \cdot success + \beta_5 \cdot suicide)}$$

The covariates in the final model are shown in Table 4 according to the order of inclusion.

Table 4 – *Cox model forward selection.*

Step	d.o.f.	Deviance	Resid. Df	Resid. Dev	AIC
	--	--	3158	46390.86	46390.86
+Attacktype	-7	326.45654	3151	46064.41	46078.41
+Region	-11	100.99569	3140	45963.41	45999.41
+Ransom	-1	30.68959	3139	45932.72	45970.72

The final Cox proportional hazards model selected by the stepwise procedure is:

$$h(t|X) = h_0(t)e^{\beta_1 \cdot \text{attacktype} + \beta_2 \cdot \text{region} + \beta_3 \cdot \text{ransom}}$$

The proportionality assumption does not hold for *Ransom* and *Attacktype*, therefore Attacktype has been recoded as $Attackidnap$ which is equal to one *iff* Attacktype="Hostage Taking (Kidnapping)", i.e. those attacks that originally are conceived with the idea of taking hostages. A time dependent effect has been used for both Attackidnap and Ransom allowing therefore the effects of these two variables to change over time. The new model is therefore:

$$h(t|X) = h_0(t)e^{(\beta_1 \cdot \text{attackidnap} + \beta_2 \cdot \log(t) \cdot \text{attackidnap} + \beta_3 \cdot \text{region} + \beta_4 \cdot \text{ran} + \beta_5 \cdot \log(t) \cdot \text{ran})}$$

All other things being equal, Attackkidnap (Table 5) shows a negative significant influence on the instantaneous hazard of incident resolution (log(HR) = -1.7; p<0.001), i.e. a significantly protracted duration for these specific events. The time-dependent effect for Attackkidnap is also highly significant and positive (log(HR) = 0.36; p<0.001) therefore while pure hostage-taking attacks generally prolong incident duration, this effect is dampened as time progresses, all other things being equal.

The presence of a ransom note or a ransom demand shows a similar behavior: highly significant negative association with the hazard of resolution (log(HR) = -0.75; p<0.001). This finding is coherent with the expectation that ransom negotiations introduce initial delays. As for Attackidnap the time-dependent effect for ransom is statistically significant and positive (log(HR) = 0.11; p<0.001).
Considering the regional effect on the instantaneous hazard of incident resolution the model suggests different patterns across different geographical regions, (reference category: Australasia & Oceania). Namely: a significant negative effect on the hazard for a group of regions where the kidnapping attacks tend to be of longer duration. This group includes:

- MENA (log(HR) = -1.1, p-value = 0.003)
- Southeast Asia (log(HR) = -1.1, p-value = 0.003)
- South America (log(HR) = -1.0, p-value = 0.008)

- South Asia (log(HR) = -0.97, p-value = 0.010)

In these regions, the instantaneous hazard of resolution is reduced compared to Australasia & Oceania suggesting a significantly longer time-to-event. Specifically, the hazard is about 67% lower in the MENA and Southeast Asia regions, 63% lower in South America, and 62% lower in South Asia compared to Australasia & Oceania. This pattern suggests potentially shared regional dynamics, varying levels of state response capacity, or differing cultural/political approaches to hostage situations that can contribute to more prolonged kidnapping time.

Table 5 – *Cox proportional hazard model estimates (forward selection).*

Characteristic	*log(HR)*	*95% CI*	*p-value*
attackidnap	*-1.7*	*-1.9, -1.4*	*<0.001*
tt(attackidnap)	*0.36*	*0.29, 0.43*	*<0.001*
factor(region)			
Australasia & Oceania	—	—	
Central America & Caribbean	*-0.38*	*-1.2, 0.39*	*0.300*
Central Asia	*-0.48*	*-1.4, 0.42*	*0.300*
East Asia	*0.44*	*-0.66, 1.5*	*0.400*
North America	*-0.70*	*-1.5, 0.11*	*0.089*
Western Europe	*-0.53*	*-1.3, 0.23*	*0.200*
Eastern Europe	*-0.78*	*-1.6, -0.01*	*0.046*
Sub-Saharan Africa	*-0.76*	*-1.5, -0.01*	*0.046*
MENA	*-1.1*	*-1.9, -0.40*	*0.003*
South America	*-1.0*	*-1.8, -0.26*	*0.008*
South Asia	*-0.97*	*-1.7, -0.23*	*0.010*
Southeast Asia	*-1.1*	*-1.9, -0.38*	*0.003*
ran	*-0.75*	*-1.0, -0.49*	*<0.001*
tt(ran)	*0.11*	*0.06, 0.16*	*<0.001*

Another group are those regions that show a marginally significant effect on the log hazard

- Eastern Europe (log(HR) = -0.78, p-value = 0.046)
- Sub-Saharan Africa (log(HR) = -0.76, p-value = 0.046)

These regions show a reduction in hazard of approximately 53% to 54% relative to Australasia & Oceania. These effects, while significant, are of a smaller magnitude than those in the previous group.

The last group is those of regions that do not show any statistically significant effect compared to the reference category (Australasia & Oceania):

- Central America & Caribbean (p-value = 0.300),
- Central Asia (p-value = 0.300),
- East Asia (p-value = 0.400),

- Western Europe (p-value = 0.200),
- North America (p-value = 0.089)

Kidnapping incident durations in these regions are not statistically distinguishable from those in Australasia & Oceania all other things being equal.

4. Conclusions

This study applied both conditional inference trees and the Cox proportional hazards model to analyze survival times in terrorist events involving hostages. Conditional inference trees offered a non-parametric framework to simultaneously perform feature selection, identify clusters, and handle both qualitative and quantitative variables without requiring distributional assumptions. They overcome the difficulty of Kapan-Meier product limit estimator to efficiently handle quantitative covariates or qualitative covariates with a large number of categories. The Cox model, as a semi-parametric method, provided an inferential framework that allowed for quantitative estimates of hazard ratios, testing of covariate effects and varying covariate effects over time. It supported feature selection and clustering of events based on estimated coefficients and their statistical significance. The Cox model findings quantified the impact of various factors on event duration, highlighting that ransom demands are significantly associated with longer survival times.

The results obtained from the tree-based and Cox model approaches were coherent and complementary. While the Cox model offered effect estimates and statistical inference, machine learning trees suggested complex interactions and clusters in terrorist behavior. These methods can be used in cascade (either tree and KM or tree and cox model), where exploratory insights from the tree model can inform the specification of the Cox model or suggest categorization of quantitative variables for the Kaplan-Meier. The congruence between the Cox model's parameter estimates (magnitude and significance of log(HR)) and the regional partitions obtained by the survival tree provides evidence that geographical region is a fundamental determinant of hostage release duration.

References

BRANDT P.T., SANDLER T. 2009. Hostage taking: Understanding terrorism event dynamics, *Journal of Policy Modeling*, Vol. 31, No. 5, pp. 758–778.

COX D.R. 1972. Regression models and life-tables (with discussion), *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 34, No. 2, pp. 187–220.

DELOUGHERY K., KING R.D., ASAL V. 2012. Close cousins or distant relatives? The relationship between terrorism and hate crime, *Crime & Delinquency*, Vol. 58, No. 5, pp. 663–688.

DUGAN L., LAFREE G., PIQUERO A.R. 2005. Testing a rational choice model of airline hijackings, *Criminology*, Vol. 43, No. 4, pp. 1031–1065.

ENDERS W., SANDLER T. 2012. The Political Economy of Terrorism. Cambridge: Cambridge University Press.

GAIBULLOEV K., SANDLER T. 2009. Hostage taking: Determinants of terrorist logistical and negotiation success, *Journal of Conflict Resolution*, Vol. 46, No. 6, pp. 882–901.

HOTHORN T., HORNIK K., ZEILEIS A. 2006. Unbiased recursive partitioning: A conditional inference framework, *Journal of Computational and Graphical Statistics*, Vol. 15, No. 3, pp. 651–674.

HOTHORN T., ZEILEIS A. 2015. partykit: A Modular Toolkit for Recursive Partytioning in R, *Journal of Machine Learning Research*, Vol. 16, pp. 3905–3909.

KAPLAN E.L., MEIER P. 1958. Nonparametric estimation from incomplete observations, JASA, Vol. 53, No. 282, pp. 457–481.

KIM W., SANDLER T. 2021. Non-UN Peacekeeping Effectiveness: Further Analysis, *Defense and Peace Economics*, Vol. 33, No. 5, pp. 511–533.

LAFREE G., DUGAN L. 2007. Introducing the Global Terrorism Database, *Terrorism and Political Violence*, Vol. 19, No. 2, pp. 181–204.

PHILLIPS B.J. 2014. Terrorist group cooperation and longevity, *International Studies Quarterly*, Vol. 58, No. 2, pp. 336–347.

PHILLIPS P.J. 2011. The Diseconomies of Terrorism, *World Economics*, Vol. 12, No. 4, pp. 171–192.

PIAZZA J.A. 2017. Repression and terrorism: A cross-national empirical analysis of types of repression and domestic terrorism, *Terrorism and Political Violence*, Vol. 29, No. 1, pp. 102–118.

START - National Consortium for the Study of Terrorism and Responses to Terrorism. 2022. Global Terrorism Database (GTD). Retrieved from http://www.start.umd.edu/gtd

WHITE G., MAZEROLLE L., PORTER M., CHALK P. 2014. Modelling the effectiveness of counter-terrorism interventions, *Trends & Issues in Crime and Criminal Justice*, No. 475. Canberra: Australian Institute of Criminology.

WILSON M.A. 2000. Toward a model of terrorist behavior in hostage-taking incidents, *Journal of Conflict Resolution*, Vol. 44, No. 4, pp. 403-424.

Luciano NIEDDU, UNINT University, l.nieddu@unint.eu
Cecilia VITIELLO, Sapienza University of Rome, cecilia.vitiello@uniroma1.it

WEB DATA SOURCES AND OFFICIAL STATISTICAL STANDARDS: THE WIN EXPERIENCE[1]

Giuseppina Ruocco, Renato Magistro, Giulio Massacci

Abstract. In order to promote the integration of web data sources for official statistics, Eurostat has launched a four-year initiative, the Web Intelligence Network project. One of the main objectives of the project is to create a network within the European Statistical System (ESS) and to develop a common infrastructure, the Web Intelligence Hub (WIH), providing services and tools for web data collection and management. The WIH is designed to support National Statistical Institutes (NSIs) in all stages of web data collection and processing. In the long term, the WIH can be further developed to explore the potential of additional innovative data sources for official statistics. During the WIN project, the architectural task has dealt with the Enhancement and Enrichment (E&E) of the architectural standard Big Data REference Architecture and Layers (BREAL). BREAL is an architectural framework developed in the Big Data II project to support NSIs in planning their investments in big data. It provides a set of tools for defining the business objectives, application components and data models needed to develop statistical processes based on big data. One of the main outcomes achieved by the architectural task is the specialization of BREAL for web data. The adoption and E&E of the BREAL framework has enabled the harmonization between the requirements of the project use cases and the services provided by the WIH. BREAL specialization results from the combination of the project experience and it is intended to promote the deployment of web data workflows in the production environment. This enhancement of BREAL enables a NSI to assess the maturity level of a use case. In addition, the adoption of this approach increases process standardization and the development of shareable tools that can be reused by other statistical organizations, and/or adapted to other use cases and/or statistical domains.

1. Introduction

The availability of new data sources resulting from technological developments represents a significant opportunity for official statistics. The new data sources increase the efficiency of statistical processes, reduce the burden on respondents, and improve the quality of the output produced. This improvement is evident in terms of both timeliness and the enrichment of statistical products, resulting in an enhancement of disseminated results.

[1] The individual contributions are as follows: Giuseppina Ruocco: Paragraphs 3,6; Renato Magistro: Abstract, paragraphs 4,5; Giulio Massacci: Paragraphs 1,2.

Although traditional data sources have been thoroughly investigated and standardised, new data sources are characterised by heterogeneity in terms of both their origin and data format. Examples include satellite images and information gathered using sensors in various contexts.

In all these cases, the issue of heterogeneity affects the entire process, not just the input and output phases. Particular focus is given to quality assessment. This assessment should consider potential new sources of error and factors that have not yet been fully investigated by the statistical frameworks adopted for traditional data sources. In order to collect, process and validate these new types of data, the methodology, tools and technological infrastructures must first be designed and tested in relation to this heterogeneity. The main challenge will then be to adapt or create new reference standards.

Starting from the experience of the Web Intelligence Network project, this article provides an overview of the Enhancement and Enrichment (E&E) of the architectural standard Big Data REference Architecture and Layers (BREAL) for web data. This enhancement of BREAL enables a National Statistical Institute (NSI) to assess the maturity level of a use case. In addition, the adoption of this approach increases process standardization and the development of shareable tools that can be reused by other statistical organizations, and/or adapted for other use cases and/or statistical domains.

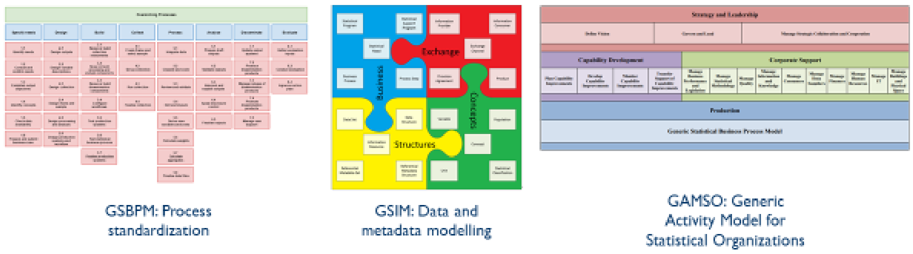
2. Official standards and new data sources

The main official statistical standards reported in Figure 1 have so far played a significant role in improving the efficiency and quality of statistical processes based on traditional data sources. The following standards have improved the efficiency of NSIs with regard to the management of processes, data and metadata. In particular:

- The GSBPM (Generic Statistical Business Process Model) provides a standardised model to describe all phases of the statistical process with harmonised and internationally shared terminology, from the definition of information needs to the dissemination of data, including the final evaluation. It is useful within the NSIs for mapping, analysing and improving statistical processes, but it is not a “rigid” framework in which all steps must be followed in a strict order.
- The GSIM (Generic Statistical Information Model) is a reference framework that describes different types of information (information objects), their attributes, and their relationships. Its purpose is to model the information used to standardise the management and use of data and metadata throughout the entire statistical production process.

- The GAMSO (Generic Activity Model for Statistical Organisations) is a standard that focuses on organisational management, aiming to describe all the activities of a statistical organisation, not just those related to production. Indeed, GAMSO builds upon the GSBPM by incorporating strategic, institutional, and support activities such as resource management, governance, planning, innovation, and external relations.

Figure 1 – *GSBPM, GSIM and GAMSO models.*



Source: Generic Statistical Business Process Model – GSBPM, Generic Statistical Information Model (GSIM) – User guide, Generic Activity Model for Statistical Organisations – GAMSO

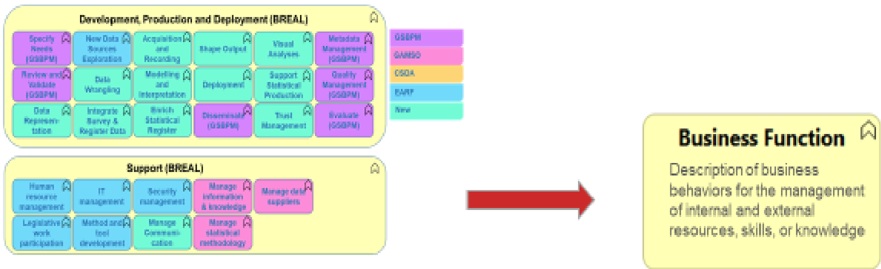
3. BREAL: an architectural framework for big data sources

At the European level, several projects integrating new data sources into statistical processes have been launched in order to address these new challenges and respond to increasing complexity. In this context, one of the main outputs of Big Data II was the BREAL framework. The goal of BREAL is to provide guidance for investments in big data and facilitate the transition from using standards in well-known environments (traditional sources) to adapting and using them in innovative contexts.

Specifically, the BREAL framework consists of several elements, each of which focuses on a particular stage of the data processing lifecycle. Together, these elements provide a roadmap that considers all aspects when dealing with new data sources, known as big data, at the beginning of their use.

During the project, the BREAL framework was applied to several use cases, the most important of which were OBEC (Online Based Enterprise Characteristics) and OJA (Online Job Advertisements). OBEC relates to the use of information on company websites to enrich the information contained in the Business Register, while OJA relates to the use of job advertisements published on job portals to develop new labour demand indicators. The framework has been designed to help statistical organisations deal with the complexities of using big data, while ensuring consistency with existing standards. Among the various components of the BREAL framework, the Business Layer is particularly relevant for modelling and standardising statistical processes based on new sources. It shows the main Business Functions aimed at acquiring and subsequent processing. From an architectural point of view, the business layer in the BREAL framework defines "WHAT" is to be realised in terms of behaviour for managing internal and external resources, competencies or knowledge. As shown in Figure 2, the BREAL business layer includes functions from various official statistics standards, including those relating to the statistical process, such as GSBPM and GAMSO, as well as architectural elements derived from CSDA and EARF, and new Business Functions related to new data sources. The Business Functions are generally divided into two main categories: the first is aimed at implementing and putting big data processes into production (Development, Production and Deployment subset), while the second relates to all the functions that support the statistical process at various levels (Support subset).

Figure 2 – *The BREAL business layer and Business function definition*.



Source: For BREAL business layer - Scannapieco M. et al. (2019): BREAL. Big Data Reference Architecture and Layers. Version 2019-12-09. Edited by EUROSTAT

In order to identify the methodological implications of implementing such architecture across the various countries, an inventory of the different use cases should be made to achieve:

- A classification of each kind of implementation by type of source (input)
- A list of the methods adopted for the pre-processing and data processing steps (throughput activities)
- The outputs produced and their integration within the related statistical domains.

Methods and, above all, quality indicators must be associated with each business function for each source type, so that the tools and procedures that implement these methods can be shared among different NSIs.

The WIN project's experience demonstrates how the BREAL framework achieves a balance between generalisation at the European level and the necessary customisation at the national level with regard to web data.

Although one of the goals of the WIN is to produce a standardised architecture and reusable tools, its effectiveness ultimately hinges on NSIs' ability to adapt it to their institutional, regulatory, and operational contexts.

BREAL defines a set of Business Functions that are universally applicable to NSIs at a 'high' level. The WIN facilitates the sharing of open source tools for data acquisition and pre-processing (e.g. deduplication and structuring), as well as data processing. In addition, it explored some quality indicators to evaluate the web data sources. This approach is consistent with the GSBPM/GSIM/GAMSO models and ensures comparability and interoperability between countries. However, while based on a shared core, each NSI should adapt the framework to its own specific requirements. The main areas of customisation include:

- Legal and ethical compliance. Development of data acquisition strategies to national context with respect to data protection and use legislation (e.g. specific GDPR interpretations, local data provider agreements).
- Integration with national IT infrastructure. It involves, for example, the inclusion of WIN tools within existing IT architectures, focusing on compatibility with cloud infrastructures.
- Statistical domains and national priorities, focusing on use cases that are most relevant to each country's needs (e.g. tourism indicators for some NSIs and labour market indicators for others).

The proposed framework addresses these challenges by offering sufficient flexibility for national adaptations, all the while maintaining a robust shared methodological framework. In this sense, the framework developed by the WIN project should not be understood as a fixed solution, but rather as an adaptable architecture designed to promote both harmonisation and innovation. By combining shared services with opportunities for national customisation, the framework provides a practical approach to integrating web data into official statistics within the ESS.

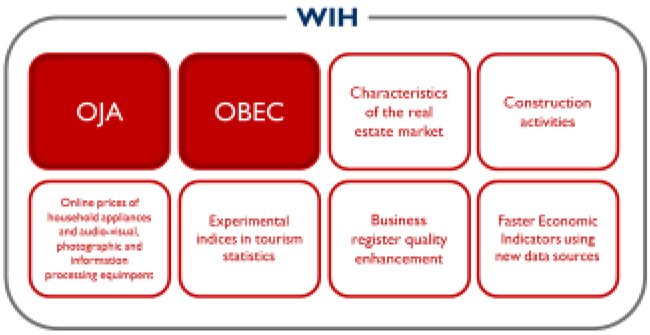
4. The WIN project and the BREAL specialization for web data

The Web Intelligence Network (WIN) project was launched in 2021 with the aim to bring the most mature use cases into production and develop the Web Intelligence Hub (WIH). The WIH is a shared European infrastructure for the efficient management of the acquisition and processing of web data for statistical purposes related to specific statistical domains. The WIH is designed to provide services for supporting NSIs at every stage of web data collection and processing.

During the WIN project, several use cases examined how web data could be utilised in various statistical fields, including the real estate market, tourism, electronics retail, labour demand trends, and business characteristics. Regardless of the statistical domain, the following characteristics were common to all of these use cases:

- The selection of URLs relating to portals and business sites
- Acquisition of information, mainly textual, from the selected URLs
- Pre-processing the acquired data, to select the useful information for statistical purposes and preparing it for further processing, after de-duplication and data structuring.

Figure 3 – *The WIH use cases.*



During the project, the architectural task specialized BREAL's business layer, building on experience with more mature use cases, mainly OBEC and OJA, to ease their transition into production environments.

The objective was to standardize the main process steps and develop shareable, reusable tools, with a focus on methodological aspects and quality assessment.

BREAL's specialization mainly concerned the subset 'Development, Production and Deployment'. It is important to note that the approach adopted was bottom-up, based on experience in the field rather than theoretical standardization. This approach not only enabled existing functions to be refined, but also allowed the BREAL model to be extended in a coherent and practical manner while maintaining alignment with other existing standards. In summary, the specialization is an evolution of official standards driven by experience, not a replacement. As an example, Figure 4 shows the results obtained for the main exploration and acquisition phases of new data sources.

More in detail:

- New Data Sources Exploration: This function involves identifying and evaluating new web data sources that could be useful for statistical purposes. A set of criteria has to be defined to evaluate and rank potential sources according to their relevance of content. Web data sources are explored to evaluate the key statistical concepts that can be derived from them and the information they provide on a specific domain of interest.

- Acquisition and Recording: This function concerns the automatic collection of data (via scraping or API) and the consequent transformation into a usable format. Adaptive solutions are required because web data sources are liable to change their structure. In addition, it is particularly important to monitor the URL stability and relevance over time, as well as accessibility issues. It is also necessary to pre-process the scraped data to avoid duplicate information and select the information that is useful for statistical purposes.

Figure 4 – *Examples of enhanced Business Functions from the 'Development, Production and Deployment' subset.*

	BREAL Original Description	Enhancement for web scraped data
New Data Sources Exploration	*Big Data specific* Besides the exploration of the new data sources the ability to find Big Data sources and to make these sources available for statistical research and development becomes important. The latter is part of the business function "Manage data suppliers"	• Define and share a **set of criteria to assess and select potential web data sources** based on a ranking approach • Explore web data sources to **compare key statistical concepts and the information** provided by web data for a specific domain of interest • Adopt an iterative approach to identify and compare web data portals and variables, population frames, lists of reference/target units
Acquisition and Recording	The ability to collect data from a given Big Data source, e.g. through API access, web scraping, etc. In addition, this function includes the ability to store and make data accessible within the NSI	The ability to: **identify and list relevant URLs**; collect and store data from the web e.g. through API access, web scraping or crawling. • After an initial phase of URL selection and landscaping, also through a list of keywords, **monitoring of stability and relevance of URLs over time**, as well as URLs accessibility issues. • Identifying and defining the reference/target units to enable the creation of population frames. **Early validation of scraped data** to prevent storing inconsistent information

Source: For the descriptions - G. Ruocco et al. (2025): Deliverable D4.7 BREAL - Big Data REference Architecture and Layers for web scraped data. Final version 2025-03-31. Edited by EUROSTAT

As an example, the specialisation of overarching Business Functions relating to the evaluation phase and quality management has emphasised the importance of using additional data sources as benchmarks for evaluating, monitoring and documenting the quality of web data. This includes a set of quality indicators relating to:

- The selection criteria and the degree of accessibility of websites
- The methods applied in the different phases of deriving statistical information, and their representativeness with respect to reference populations in the different statistical domains. The aim is to ensure compliance with minimum standards of accuracy, consistency, completeness and timeliness.

In order to evaluate the implemented steps, it is important to be able to assess and monitor the main issues (methodological, technical, operational and organizational), for example unforeseen changes to websites or the need to enter into special agreements with website owners to guarantee long-term data accessibility. The evaluation phase also identifies possible areas for improvement and contributes to progressive standardization, implementing the iterative vision of the process, in full consistency with the GSBPM principles, while adapting to the flexibility required by web data. The figure below shows the original BREAL description and the enhancement version of these Business Functions for web data.

Figure 5 – *Examples of enhanced Business Functions from the 'Support' subset.*

	BREAL Original Description	Enhancement for web scraped data
Quality management (GSBPM)	For the original description, see GSBPM, paragraph 106 through 114: The main goal of quality management within the statistical business process is to understand and manage the quality of the statistical products. In order to improve the product quality, quality management should be present throughout the statistical business process model. All evaluations result in feedback, which should be used to improve the relevant process, phase or sub- process, creating a quality loop	Considering the specific features of web data sources, a set of auxiliary information is required to assess, monitor and document the quality of web data, the process and the results. In particular, a set of quality indicators is required to evaluate a web data source: • **Websites selection criteria and accessibility** • **Accuracy of collected information and applied methods** • **Representativeness with respect to population frames** • **Comparability with other statistical sources** • **Timeliness and relevance of the final results**
Evaluate (GSBPM)	*Big Data specific* When Big Data sources are used, evaluation plays an important role. Most of the specificities of Big Data are related to its quick pace of change, both in terms of the population covered and of their behaviour. Thus issues like coverage, accuracy and fitness of the model must be constantly assessed and monitored	Starting from the auxiliary information related to the previous BBFs (Metadata Management and Quality Management), the **ability to assess and monitor the main issues** (methodological, technical, operational, organizational) affecting the workflow developed for a specific use case. As an example: • **Unexpected websites changes** • **Agreements with websites owners to ensure data accessibility over time** • **Models decay due to data or concept drifts** • **Quality improvements through manual revisions of a sample of units** • **Staff trainings**

Source: For the descriptions - G. Ruocco et al. (2025): Deliverable D4.7 BREAL - Big Data REference Architecture and Layers for web scraped data. Final version 2025-03-31. Edited by EUROSTAT

In addition to methodological and architectural guidelines, the WIN project has produced a series of empirical results that demonstrate the validity and feasibility of the proposed approach in supporting web data acquisition and management processes in real-world contexts.

Applying the standard has helped to highlight which stages of the process need to be monitored over time to improve input data quality. For instance, in the OJA use case, the assessment of the web data sources (in the landscaping phase of New Data Sources Exploration) resulted in the definition of indicators to evaluate and monitor the relevance and stability of sources over time[2]. Overall, this evidence shows that the project has provided not only a conceptual model, but also an operational platform capable of producing concrete results and driving innovation in statistical processes across various domains.

The main lessons learnt from the initial applications are as follows:

- As previously mentioned, the processes that perform specific Business Functions are influenced by the particular characteristics of the national contexts in which they are implemented. This is particularly evident in ethical and legal aspects, as well as those relating to environmental and infrastructure

[2] These results are documented in Deliverable 4.8, *Quality Assessment for the Statistical Use of Web-Scraped Data*, which provides further details on performance metrics, critical issues encountered, and solutions adopted.

management. This limits the applicability of a framework and highlights areas for improvement

- Sharing experiences, methods, and tools creates synergies that benefit the entire statistical community. Therefore, it is crucial to monitor various use cases over time and support the development of shared infrastructures to facilitate the integration of new data sources into official statistics

• From an architectural perspective, combining a top-down approach — from Business Functions to the implementation of processes to achieve them — with a bottom-up approach — from implemented use cases to Business Functions — fosters process standardization based on the type of source and data. This promotes the sharing of tools and methods between NSIs, starting from a common framework.

5. Specialization of official standards based on use cases experience: a SWOT analysis

In order to assess the pros and cons of specializing official statistical standards and frameworks based on use cases experience, the SWOT analysis reported in Figure 6 provides an overview of the benefits and challenges.

Figure 6 – *Specializing official standards through use cases experience: SWOT Analysis***.**

Specifically, the main Strengths (S), Weaknesses (W), Opportunities (O) and Threats (T):

S

- Testing the applicability of a standard and highlighting areas for improvement through a 'learning by doing' approach
- Reusing available tools and increasing interoperable solutions

- Competence sharing: developing a collaborative ESS community that is active in exchanging best practices
- Open-source integration: improving transparency and collaboration
- Fostering the development of a shared, centralized infrastructure to facilitate cooperation and continuous improvement.

W

- The maturity level of use cases influences the enrichment of a standard
- Problems that need to be addressed at a national level, e.g. national regulations, may prevent a high degree of standardization
- Ethical and legal constraints: uncertainties about the conditions of use of new data sources
- Complexity in stakeholders' management, necessitating clarity in operational roles and responsibilities.

O

- Enriching official statistics by integrating traditional and innovative data sources
- Improvement of traceability and documentation of processes
- Harmonization of bottom-up (use cases) and top-down (standards) approaches
- Improvement and updating statistical standards based on experience
- Increase of efficiency through the centralization of complex processes to save resources within NSIs.

T

- Determining the feasibility of adapting the standard to avoid creating a new one
- Peculiarities of the domain that require specific solutions in terms of methods and tools
- Risks and sources of errors that arise from using new data sources for further investigation (e.g. issues related to content ownership, accessibility and privacy)
- Large investments that are required to design, develop, create and improve shared infrastructures for specific types of new data sources.

6. Conclusions

The WIN project demonstrated the benefits of applying and enhancing the BREAL architectural framework, in order to foster the implementation of solutions in the production environments. This approach could be extended to other types of new data sources with the aim of identifying common aspects that could be standardized. This strategy is useful and feasible across different statistical domains, facilitating the implementation of shared infrastructures at a European level to integrate new data sources into the statistical process and enrich official statistics. Adopting this approach increases process standardization and the development of shareable tools that can be reused by other statistical organizations, and/or adapted to other use cases and/or statistical domains.

In addition to the operational benefits demonstrated in the use cases, the WIH is not only a common infrastructure, but also a strategic element if considered within the data ecosystem strategy conceived for the process of modernising and innovating the ESS.

Firstly, the WIH is fully in line with European strategic guidelines, pursuing the objectives of efficiency, cooperation and reduction of production costs. The centralisation of complex functions – such as the acquisition and pre-processing of web data – allows NSIs to free up internal resources, focusing on analyses and adaptations specific to national priorities.

Secondly, the WIH provides resources for addressing emerging legal and ethical frameworks.

In a context characterised by complex regulations (e.g., the GDPR and the recent EU AI Act on artificial intelligence), the hub provides a space in which common guidelines can be developed and shared for ethical scraping, data anonymization and transparent governance. In this sense, the WIH reduces regulatory uncertainty and strengthens user and citizen trust in official statistics.

Another aspect to consider is the potential for future developments. Although the WIH is currently designed for web data, its operating model can be extended to other innovative sources (e.g., data from sensors, satellite images) as well as new-generation administrative sources. Thus, the WIH can evolve from a thematic platform into a reference infrastructure for integrating heterogeneous sources, thereby ensuring methodological continuity and consistency with official statistical standards.

Finally, in terms of the relationship between national priorities and European cooperation, the WIH acts as an enabling tool. On the one hand, it enables NSIs to respond quickly to specific national requirements (e.g. labour market indicators), and on the other, it fosters the exchange of best practices and the reuse of tools, thereby promoting innovation at the ESS level.

In summary, the WIH is not only the outcome of a project, but also an innovation platform for official statistics. It strengthens the ability of the European system to respond to the challenges posed by new data sources by integrating methodological rigour, institutional cooperation, ethical responsibility, and regulatory compliance.

References

AA.VV. 2025. Chapter 5 National data ecosystems and governance. In *The Handbook on Management and Organization of National Statistical Systems.*

Edited by UN, pp. 97-111. https://projects.officialstatistics.org/hb-mgnt-org-nss/handbook/intro.html .

ASCHERI A., MUSEUX J. M., WIRTHMANN A., GIANNAKOURIS K., KARLBERG M., BALDACCI E.2022 Innovation in the European Statistical System: Recent achievements and challenges ahead, *Statistical Journal of the IAOS*, Vol. 38, No. 3, pp. 805-813. https://doi.org/10.3233/SJI-220053.

AUNO V. *et al*. 2024. *Deliverable 4.8: Quality Assessment for the Statistical Use of Web Scraped Data. Final version, 2024-11-20*. Edited by Eurostat.

DAAS P. J. H. 2015. Big data as a source for official statistics, *Journal of Official Statistics*, Vol. 31, No.1, pp. 249–262. https://doi.org/10.1515/jos-2015-0016.

Generic Statistical Business Process Model – GSBPM. https://statswiki.unece.org/display/GSBPM

Generic Statistical Information Model – GSIM. https://statswiki.unece.org/display/gsim

Generic Activity Model for Statistical Organizations – GAMSO. https://statswiki.unece.org/spaces/GAMSO/pages/105580149/Generic+Activity+Model+for+Statistical+Organizations

KOWARIK A. *et al.* 2021. *Deliverable 4.1: Minimal guidelines and recommendations for implementation. Version 2021-09-23.* Edited by Eurostat.

RUOCCO G. *et al.* 2025*. Deliverable D4.7 BREAL - Big Data REference Architecture and Layers for web scraped data. Final version, 2025-03-31.* Edited by Eurostat.

SCANNAPIECO M. *et al*. 2021. *(Deliverable F2) BREAL. Big Data Reference Architecture and Layers. Application layer and Information layer. Version 2021-03-31.* Edited by Eurostat.

SCANNAPIECO M. *et al*. 2019. *(Deliverable F1) BREAL. Big Data REference Architecture and Layers. Business Layer. Version 2019-12-09.* Edited by Eurostat.

SIX, M. *et al*. 2025. *Deliverable 4.6: WP4 Methodology report on using web scraped data. Final version, 2025-03-25.* Edited by Eurostat.

SIX M. *et al*. 2025. *Deliverable 4.5: Quality Guidelines for acquiring and using web scraped data. Revised final version, 2025-02-20.* Edited by Eurostat

STRUIJS P, BRAAKSMA B., DAAS P. J.. 2014. Official statistics and Big Data, *Big Data & Society*, Vol. 1, No. 1 https://doi.org/10.1177/2053951714538417.

Giuseppina RUOCCO, Istat, giruocco@istat.it
Renato MAGISTRO, Istat, renato.magistro@istat.it
Giulio MASSACCI, Istat, giulio.massacci@istat.it

BIG DATA AND ROAD SAFETY: OPEN STREETMAP AND THE TERRITORIAL ANALYSIS OF ACCIDENTS[1]

Marco Broccoli, Silvia Bruzzone

Abstract. Road safety is one of the main challenges for sustainable mobility policies, aiming to reduce the number of accidents and their consequences. Analysing road accidents is essential to identify risk factors and to develop prevention strategies, based on data. The aim of this study is threefold. First, to utilize OpenStreetMap (OSM) data to calculate road accident, mortality, and injury indices by correlating them with the length of road lanes (in meters). Second, to conduct a territorial analysis to identify high-risk areas, thereby supporting road safety planning and third, to enhance national statistical information by estimating accident involvement probabilities, with the ultimate goal of determining real traffic flows (vehicles/km) and actual risk exposure rates. The approach uses an integrating geographic and statistical data using GIS techniques. The researchers implement a spatial join algorithm to overlay information layers derived from OSM and traffic points (PoT).

The analysis includes a new classification of road segments, updated to 2021, and the application of the "Ranker" tool to generate synthetic indicators. Accident data, provided by Istat and other administrative sources, are georeferenced and analysed to highlight territorial variations in risk distribution.

The main innovation of this study lies in the use of Big Data from OSM for statistical purposes, aligning with Trusted Smart Statistics (TSS) initiatives. The integration of geographic and statistical data overcomes the limitations of traditional risk measures based on resident population or vehicle ownership. Furthermore, the introduction of traffic points refines risk indicators, providing a more detailed framework for accident prevention on a territorial scale.

1. Introduction

Road safety remains a paramount challenge within sustainable mobility policies, with concerted efforts aimed at reducing the number of accidents and their severe consequences. Road Safety Performance Indicators (RSPI), as recommended by European Commission programs for EU countries, offer a multidimensional approach to investigating accidents, considering roads, vehicles, and individuals involved. Preventing road trauma on public roads is a core responsibility for governments, their agencies, and stakeholders, necessitating a common and shared commitment. The scale

[1] Marco Broccoli edited paragraphs 2, 2.1, 2.2, 2.3, 2.4, 2.5, 3, 3.1, 3.2, 3.3, 3.4 and 3.6; Silvia Bruzzone edited paragraphs 1, 3.5 and 4.

of the road safety challenge and the diversity of its impacts underscore the importance of exploring synergies among decision-makers within the road network.

In Italy, road accidents resulting in death or injury continue to pose a significant public health and social burden. Data for the years 2010-2023 illustrate ongoing trends, with a target for 2030 aiming at halving the number of deaths[2]. While there has been a general downward trend in accidents and injuries over the past decade, the number of fatalities remains a critical concern, with an estimated 3,030 deaths in 2024. This perspective highlights the urgent need for effective, data-driven prevention strategies.

Traditionally, analysts calculate road fatality and accident rates based on denominators such as the resident population or the vehicle fleet in each province of registration. However, a clear information bias exists regarding the appropriate reference denominators for constructing robust statistical indicators linked to road accidents (Broccoli and Bruzzone, 2021). Resident population, often used as a common proxy for the population exposed to risk in a specific geographical area, is not always an appropriate solution. This is especially true given the seasonal nature of road accidents and their concentration in specific locations or periods, which means resident population figures, do not accurately reflect the actual population present at the time and place of an event. Similarly, while the vehicle fleet provides more targeted information, it still suffers from a deductible bias due to the mobility of road users, failing to account for vehicles transiting through an area.

These traditional indicators, though more easily accessible, are therefore inherently limited. The high mobility of individuals for work, leisure, family needs, and commuting creates a substantial distortion when attempting to assess accident risk in specific geographical areas. In contrast, this study leverages the length of the road network (carriageway length in meters) from Open Street Map (OSM) as a more stable and geographically anchored denominator. The road infrastructure, unlike population or vehicle fleet, remains a constant physical presence within a territory, offering a more consistent basis for risk assessment. The ultimate goal is to enhance national statistical information by estimating the probability of accident involvement, accounting for different risk exposures, and eventually estimating effective traffic flows (vehicles/km) on the national road network.

This research aligns with Istat's initiatives in Experimental Statistics and the pursuit of "Trusted Smart Statistics" (TSS). Since 2019, Istat has been developing experimental statistics that include road accident, mortality, and harmfulness rates, comparing traditional measures with those based on roadway segment length from OSM. The final aim for these experimental statistics is to publish the documents as official statistics, with a continuous and planned timetable by 2026. Beyond academic advancement, a

[2] European Commission. EU Road Safety Policy Framework 2021-2030 - Next steps towards "Vision Zero". Brussels 19.6.2019, SWD (2019) 283 final. Link: https://transport.ec.europa.eu/system/files/2021-10/SWD2190283.pdf

primary driver of this research is the development of robust synthetic indicators that can serve as actionable tools for policymakers. Such indicators are crucial for enabling more informed decision-making in the programming of preventive actions, the strategic modernization of road infrastructures, and, where appropriate, the targeted enforcement of driving behaviours to enhance overall road safety. This paper presents the methodology and main results of using OSM data for a territorial analysis of road accidents, highlighting the advantages of innovative denominators and synthetic indicators.

2. Data and Methods

The approach integrates geographic and statistical data using Geographic Information Systems (GIS) techniques and OSM as a key Big Data source.

2.1 Open Street Map Data

OSM is a collaborative project aimed at creating free, editable content maps of the world. It provides a vast collection of geographical data, including detailed road networks, with the primary purpose of creating maps and cartography. A key feature of OSM data is its free license (Open Database License), allowing for its use for any purpose with the only constraint of mentioning the source. A global community contributes data using GPS devices, aerial photography, and other free sources. Most Android and iOS GPS navigation software (e.g., WisePilot, Maps.me, NavFree) are powered by OSM.

For this study, key OSM vector layers, which are daily updated and freely downloadable, include:

- Road graph (detailing road segments and their characteristics);
- Traffic Points (PoT), indicating locations on road segments where traffic intensity.

2.2 GIS Techniques and Data Integration

A spatial join algorithm integrates the data by linking Istat’s census geography with the OSM road network. This process overlay the two vector layers, enriching Istat administrative features (e.g., localities) with OSM road segment attributes based on spatial location. The procedure generates a unified dataset that incorporates both administrative boundaries and detailed infrastructure characteristics through the key-reference-by-position algorithm.

2.3 Classification of Road Segments and Localities

To build road accident indicators with denominators represented by road segment length from OSM, Istat uses a "bridge coding table". This table systematically classifies OSM road segments according to Istat survey on road accidents classification and Istat Population Census Localities.

The classification considers:

- OSM road segment types (e.g., motorway, trunk, primary, secondary, tertiary, residential, service) (Table 1).
- Istat Population Census Localities types: Urban areas, Small inhabited areas, Productive areas, widespread houses;
- Road accidents survey road types (e.g., Motorway, Urban Road and Rural Road).

Table 1 – *OSM road segments classification (a).*

Road type	Road type description
Secondary link	The link roads (slip roads/ramps) leading to/from or from/to a secondary road or lower class highway.
Tertiary	Roads of local rank. They connect smaller municipalities together. In urban areas, they are side roads to primary and secondary roads with a medium flow of traffic.
Tertiary link	The link roads (slip roads/ramps) leading to/from or from/to a tertiary road or lower class highway.
Unclassified	Classification for some extra-urban road.
Residential	Roads in a residential area, which serve as an access to housing, without function of connecting settlements.
Living Street	Residential road where pedestrians have legal priority over cars, speeds are very low.
Pedestrian	Pedestrian areas (roads or squares in urban areas), accessible mainly or to pedestrians.
Service	Access roads or internal service areas, beaches, camping, industrial areas, shopping centers, parking places etc.
Track	Roads for mostly agricultural or forestry uses
Cycle way	Cycle paths on dedicated carriageway, mainly or exclusively for cycling tourism.
Footway	Paths mainly/exclusively for pedestrians. This includes walking urban tracks, paths in a public park and footpaths
Path	Paths not structured for a public use
Steps	Stairs in steps, exclusively accessible by pedestrians
Unknown	Not classified

(a) Road Segments by Open Street Map – update 1/1/2024.

The study adopts a new analytical classification with respect to the first release (Broccoli and Bruzzone, 2021), using a more refined technique for attributing individual road segments, approximately 3.5 million in total from OSM to the Istat classification groups (Table 2). The operational criterion applied provides the roads classification, through the textual analysis of the Name and Reference attributes, according to the different classes of road segment and spatial attribution of the location type.

Table 2 – *Bridge coding table between roads segments classification by OSM, localities and road type (a).*

Road Segments classification by OSM	Localities at Census 2011			
	Urban areas + Small		Productive areas + Wide	
	Road Localisation by Road accidents survey			
	Motorways	Urban Roads	Motorway	Rural Roads
Motorway	x		x	
Trunk	x		x	
Primary		x		x
Secondary		x		x
Tertiary		x		x
Unclassified		x		x
Residential		x		x
Living Street		x		x
Motorway Link	x		x	
Trunk Link	x		x	
Primary Link		x		x
Secondary Link		x		x
Tertiary Link		x		x
Service		x		x
Unknown		x		x

(a) *Istat computing*

2.4 *Road Accident Indicators and Traffic Point (PoT) Weighting*

The study calculates road accident, mortality, and injury indices. The innovation lies in correlating these with the length of road lanes (in meters) by driving direction from OSM. As a further refinement, the study used additional information on Traffic Points (PoT) on road segments, downloaded from the OSM detection system. Since the 2021 edition of experimental statistics, Istat researchers propose road accident indicators "weighted" with information on traffic intensity. This information considers the kilometres of roadway with the presence of a traffic point as a discriminating element, aiming to better approximate actual traffic exposure. The ultimate objective is to estimate real traffic flows (vehicles/km).

2.5 *Synthetic Indicators: The Ranker Tool and MZ Method*

For the analysis and comparison of the synthetic indicators, Istat uses two tools developed in-house:

- Ranker Tool desktop software: (http://www.istat.it/en/tools/methods-and-it-tools/analysis-tools/ranker)
- i.Ranker web application: (https://i.ranker.istat.it)

Three synthesis methods are evaluated: the arithmetic mean of z-scores (MZ), the Relative Index method (MR), and the Mazziotta-Pareto Index method (MPI; De Muro *et al.,* 2010). For the analysis and comparison of the synthetic indicators, Istat uses tools developed in-house, specifically the "Ranker" and "COMIC" (COMposite Indices Creator) software. Three synthesis methods are evaluated: the arithmetic mean of z-scores (MZ), the Relative Index method (MR), and the Mazziotta-Pareto Index method (MPI; De Muro *et al.,* 2010).

The robustness analysis entailed a comparative assessment of these methods using the COMIC software to verify their internal consistency. Specifically, the stability of the results was measured by analyzing the mean and standard deviation of the rank "shifts" produced by each method. The results demonstrated that both MZ and MPI offered the highest robustness (i.e., minimal rank volatility). Ultimately, the MZ method was selected as the primary criterion due to its computational transparency and ease of interpretation for non-statistical stakeholders.

Regarding the construction of the indicator, the MZ method operates by first standardizing the elementary indicators (accident, mortality, and injury rates) to z-scores (with zero mean and unit variance) to ensure comparability across different scales. Subsequently, the arithmetic mean of these standardized values is calculated to generate the final synthetic index, allowing for consistent territorial comparisons.

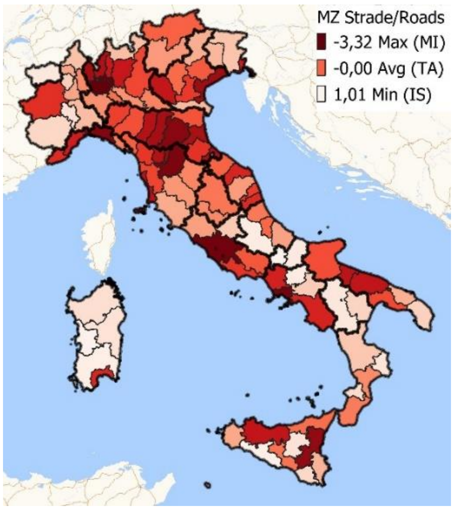
3. Main results

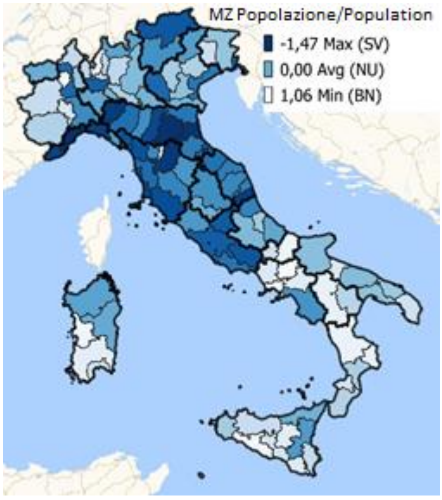
3.1. Comparing Rates: Road Length vs Population Denominators

Focusing on 2023 data, road accident indicators by road length (number of accidents, vehicles involved, deaths, and injuries per 1000 km of carriageway in the province) reveal the maximum exposure to risk for motorways and urban roads, primarily stated in main cities. Comparing the road mortality rates with those calculated using resident population as a denominator shows that the ranking of provinces is completely different. For example, Milan and Rome, which rank high for road mortality risk (1st and 4th, respectively) based on road length, drop to 88th and 60th positions out of the resident population. For motorways specifically, their positions shift from 3rd and 11th (by road length) to 17th and 33rd (by population). This clearly demonstrates the distortion introduced by population-based denominators, which do not account for the actual use of the road network. The same results arise even more clearly analyzing the synthetic index provinces ranking (Figure 1, 2).

Figure 1 -*Synthetic Indices: MZ Rank by road resident segments length*

Figure 2 - *Synthetic Indices: MZ Rank by population*





Source: Istat computing on Istat Road Accidents survey (2023), OSM data (1/1/2024) and Istat resident population (1/1/2023).

3.2. Detailed Provincial Rankings and the Impact of Denominators

The analysis of provinces exhibiting the highest and lowest road safety starkly reveals the profound impact of the chosen denominator. The list of the five "worst-performing" provinces, for instance, undergoes significant alterations when switching from road length-based indicators to those reliant on resident population or vehicle fleet. This volatility underscores the critical need to select denominators that accurately reflect risk exposure, as traditional metrics can lead to divergent and potentially misleading policy priorities (Table 3).

Table 3 - *Best and Worst Z Road graph performance. Year 2023 (a).*

Best Z Road graph performance. Year 2023						
	Ranking indicators by road length		Ranking indicators by vehicles fleet		Ranking indicators by resident population	
Best	Benevento	0.883	Biella	0.780	Biella	0.633
	Sud Sardegna	0.919	Agrigento	1.030	Oristano	0.745
	Agrigento	0.929	Trento	1.045	Napoli	0.767
	Oristano	0.999	Benevento	1.046	Agrigento	1.011
	Isernia	1.017	Aosta	1.330	Benevento	1.062

Table 3 (cont.) - *Best and Worst Z Road graph performance. Year 2023 (a).*

Worst Z Road graph performance. Year 2023						
	Ranking indicators by road length		Ranking indicators by vehicles fleet		Ranking indicators by resident population	
Worst	Milano	-3.330	Bologna	-1.602	Savona	-1.475
	Monza e Brianza	-2.536	Genova	-1.114	Bologna	-1.252
	Roma	-2.085	Savona	-1.108	Ravenna	-1.088
	Napoli	-1.768	Piacenza	-1.072	Firenze	-1.067
	Firenze	-1.720	Ravenna	-1.026	Piacenza	-0.991

(a) A Z-score measures the distance between a data point and the mean using standard deviations. Z-scores can be positive or negative. The sign tells you whether the observation is above or below the mean. The research use the negative polarity to represent the higher risk.

3.3. Correlation Analysis: Highlighting Methodological Differences

The correlation matrix results illustrate the distinctiveness and innovative strength of using road length as a denominator (Table 4).

Table 4 - *Correlation Matrix between indicators (analytical classification of road segments).*

Ranks	Road segment (Analytical classification)	Population	Vehicle Fleet
Road segment	1.0000	0.3794	0.4830
Population	0.3794	1.0000	0.8843
Vehicle Fleet	0.4830	0.8843	1.0000

The low correlation coefficients between indicators based on road segment length and those based on resident population (0.3794) or vehicle fleet (0.4830) empirically confirm the significant differences in risk assessment these approaches yield. This divergence underscores how denominators tied to human mobility (population, vehicle fleet) – influenced by work-related travel, leisure trips, family obligations, and daily commuting – introduce a considerable bias in risk perception. The road network's physical extension, on the other hand, provides a more objective and territorially consistent reference for evaluating accident phenomena. The application of different weighting criteria leads, in fact, to very divergent results. The road accident indicators referred to road length by province, therefore, seem to lead to a better result for the risk measure of the road accidents and are closer to the values by traffic flows data.

3.4. Synthetic Indicators (MZ Method) and Cartographic Representation

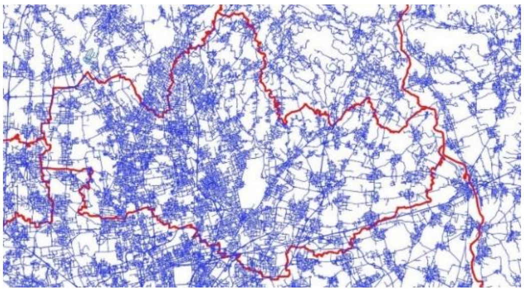
Synthetic indicators using the MZ method provide a composite view of road safety risk. Cartographic representations, applying a quantile method in color classes (higher values highlighted with a more intense tone), visualize these territorial differences

These maps clearly show different risk patterns depending on the normalization basis, reinforcing the utility of road-length based synthetic indicators for policymakers. This refined understanding moves beyond simple incident counts, offering a strategic basis for prioritizing infrastructure upgrades, tailoring public awareness campaigns, and guiding law enforcement.

3.5. Case Studies: Illustrating Denominator Impact

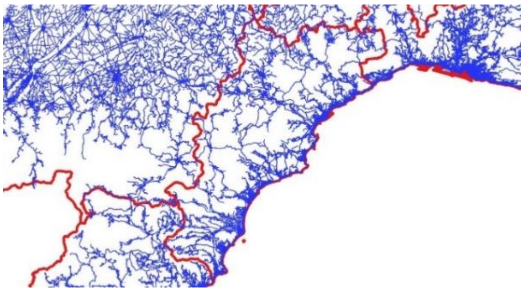
The province of Savona serves as a compelling illustration of the distortions caused by traditional denominators. A significant network of connecting infrastructures and many tourist destinations characterize Savona, leading to substantial traffic volumes from transit and tourism. Consequently, its resident population is relatively small when compared to the actual traffic volumes and the extent of its road network. When road accident risk in Savona is assessed using resident population as a denominator, a potential overestimating of the local risk effect can affect the resulting indicators, because the denominator does not capture the large transient, non-resident road user population. Indicators based on road network length offer a more stable and geographically pertinent assessment of risk in such territories (Figure 3).

Figure 3 - *Monza e della Brianza road layout by OSM graph. Year 2023.*



Conversely, the province of Monza and Brianza exemplifies a different scenario. An extremely high urban concentration and pervasive urbanization characterize this territory, resulting in a dense road network, intensively utilized through all geographical area. In such a context, the indicator normalized by road carriageway length reaches maximum risk values. This suggests that the sheer density of traffic and interactions on a highly utilized, even if geographically limited, road network creates a heightened risk environment. While population-based indicators might also show high risk, the road length metric emphasizes the intensity of risk concentrated on the existing infrastructure. This is particularly sensitive for identifying areas where infrastructure is under significant pressure, a critical factor in densely urbanized settings. The contrasting findings for provinces like Savona and Monza and Brianza demonstrate that road length-based indicators are not only more stable but also capable of revealing different facets of road safety risk crucial for tailored policy interventions (Figure 4).

Figure 4 - *Savona road layout by OSM graph. Year 2023.*

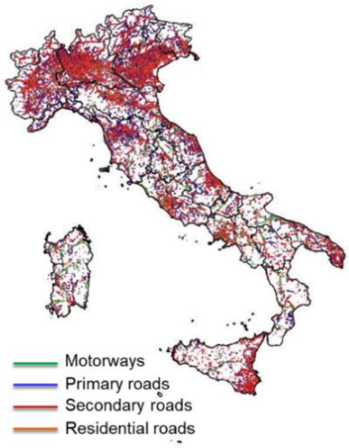


3.6. Towards the Measurement of Traffic Flows

The first result achieved with the measurement of road segment length allows a first step towards correlating road accidents and traffic flows for a more correct measurement of risks. It was essential to start from the knowledge of the length of the national road network by locality to reach the most frequently used indicator of "vehicles per kilometer" per road segments. The refinement process involves identifying road segments with intense traffic flows (using Point of Traffic - PoT data from OSM) and then constructing new synthetic indicators. This ongoing work aims to move beyond static denominators towards dynamic measures of exposure (Figure 5).

Figure 5 visualizes the spatial distribution of these Traffic Points (PoT) across the national territory. The map reveals a significant density of detection points along primary transport corridors, specifically motorways and major state roads and within key metropolitan areas such as Milan, Rome, and Naples. This spatial pattern confirms that the OSM dataset offers substantial coverage of the most intensively used segments of the network. By weighting road segments based on this PoT density, the model can differentiate between high-flow arteries and low-traffic rural roads, providing a crucial correction factor for estimating risk exposure more accurately than simple road length.

Figure 5 - *Point of Traffic - PoT data from OSM. Year 2023.*



4. Conclusions

This study demonstrates the significant potential of leveraging Big Data from OSM for a more nuanced and accurate territorial analysis of road accidents.

Integrating Big Data into official statistics requires continuous methodological adaptation to ensure consistency, quality, and representativeness. It is essential to support innovative data sources with a solid theoretical and statistical foundation to turn raw information into actionable knowledge for public policies. A key methodological contribution here is the innovative application of road carriageway length as a primary denominator, which, unlike traditional measures (resident population, vehicle fleet), is less affected by the distortions caused by human mobility. The inherent stability of the road network's length offers a more robust foundation for assessment. The project using OSM to investigate road accident patterns follows these principles, with ongoing activities expected to lead to further developments. Using non-traditional sources like OSM expands the possibilities for analysing road accidents across space and time. By combining open data, geospatial technologies, and advanced statistical methods, researchers can produce more timely and detailed territorial statistics. This study also shows that OSM-based Big Data can support synthetic indicators that offer clear benefits for policymakers. These tools empower a more evidence-based approach to road safety management, facilitating better-targeted prevention programs. The indicators allow proposing an efficient allocation of resources for infrastructure improvements and effective strategies for addressing unsafe driving practices too. Differences in high-risk province rankings, along with cases like Savona and Monza and Brianza, show the need for indicators that capture real exposure more accurately. Statistical agencies, researchers, and the open-data community must work together to

create shared standards and validation tools, aiming to estimate actual traffic flows on the national road network and to calculate more reliable accident probabilities.
Researchers who apply this methodology in other contexts should adapt the bridge coding table to local classifications, align the chosen OSM snapshot with the accident reference period, and use both advanced GIS skills and statistical rigor to manage the heavy computational work behind national-scale analyses.

References

BROCCOLI M., BRUZZONE S. 2023. Road accidents in Italy: New indicators, at province level, based on geographic information system open data. *Statistics, Technology and Data Science for Economic and Social Development Book of short papers of the ASA Bologna Conference 2023* - Supplement to Vol. 35, No. 3.

BROCCOLI M., BRUZZONE S. 2021. *Use of the open street map to calculate indicators for road accidents on the Italian roads. Updating with 2017 data*. Istat, Rome, Experimental Statistics. https://www.istat.it/en/archivio/257384

BROCCOLI M., BRUZZONE S. 2019. *Use of the open street map to calculate indicators for road accidents on the Italian roads. Year 2016*. Istat, Rome, Experimental Statistics. https://www.istat.it/en/archivio/231740

DE MURO P., MAZZIOTTA M., PARETO A. 2011. Composite Indices of Development and Poverty: an Application to MDGs. *Soc Indic Res*, Vol. 104, pp. 1–18.

Istat 2024. *Road Accidents in Italy. Year 2023*. Press Release.
https://www.istat.it/en/press-release/road-accidents-2023/

HAKKERT A.S., BRAIMAISTER L. 2002: *The uses of exposure and risk in road safety studies*. Number: R-2002-12, SWOV Institute for Road Safety Research, The Netherlands. Leidschendam 2002

ZILSKE M., NEUMANN A., NAGEL K. 2011. Open Street Map for traffic simulation. In *Proceedings of the 1 st European state of the map: OpenStreetMap conference*. - Wien: OpenStreetMap Austria u.a., 2011. - pp. 126–134.

Marco BROCCOLI, Istat - Istituto Nazionale di Statistica broccoli@istat.it
Silvia BRUZZONE, Istat - Istituto Nazionale di Statistica bruzzone@istat.it

THE NEW TRAJECTORIES OF ARTIFICIAL INTELLIGENCE A MULTIDIMENSIONAL ANALYSIS

Giuseppe Lecardane

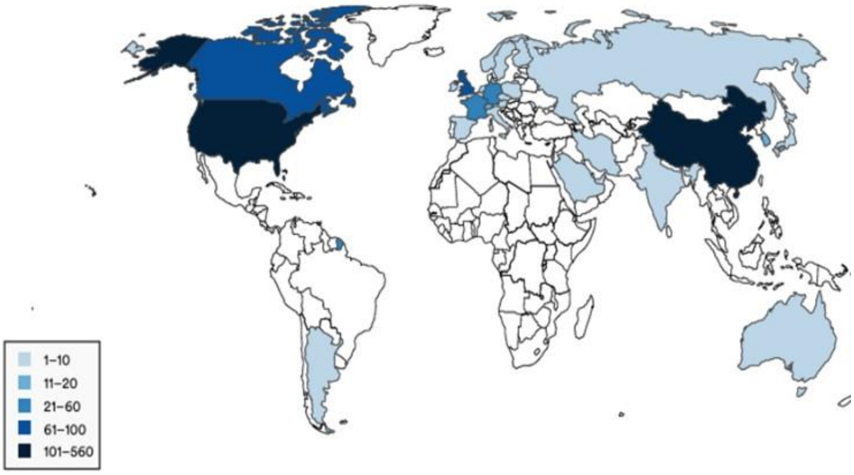
Abstract. Contemporary technological progress in the fields of statistics, artificial neural networks, and machine learning has generated remarkable advancements in Artificial Intelligence (AI) innovation. The proliferation of AI technologies is fundamentally transforming production methodologies, business paradigms, and organizational structures within both private sector enterprises and public administration institutions.
The present work examines principal indicators for the analysis and development of this novel advanced technology in the European context. Given the multidimensional characteristics of the study, following the analysis and description of individual dimensions, a multivariate analysis is conducted for territorial and temporal comparisons. This investigation offers significant contributions to understanding the contemporary and prospective AI landscape, with substantial ramifications for strategic decision-making across social, economic, technological, and industrial sectors.

1. Global and European Landscape in AI Development

The analysis of the global landscape reveals a significant concentration of relevant AI model development in advanced economies (Figs. 1 and 2) [1]. The United States, Canada, and several European countries exhibit the highest density of AI models, highlighting their role as technological leaders. China presents a moderate but significant concentration, while Europe shows a heterogeneous distribution with particular concentration in Nordic and Western countries[2].

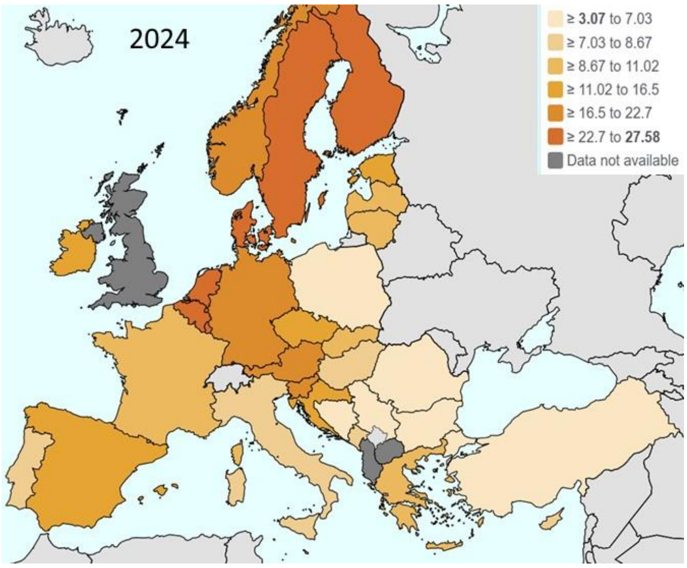
[1] EUROSTAT. 2025. *Use of artificial intelligence in enterprises*. Statistics Explained Statistics Explained (https://ec.europa.eu/eurostat/statisticsexplained).
[2] STANFORD UNIVERSITY HAI. 2025. *Artificial Intelligence*. Index Report 2025.

Figure 1 – *Number of notable AI models by geographical area, 2003-24 (Sum).*



Source: Epoch AI; Chart: 2025 Index AI report.

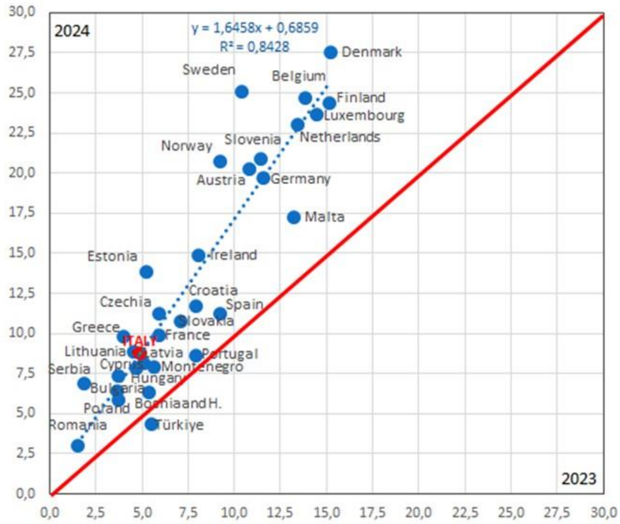
Figure 2 – *Enterprises using AI technologies, 2024 (%).*



Source: Eurostat data processing.

Analyzing temporal dynamics within the EU landscape, the 2023-2024 transition diagram reveals distinctive patterns of growth in AI adoption (Fig. 3). The positive correlation (R^2=0.8428) between adoption levels in the two consecutive years indicates structural stability among technological leaders. Nordic countries (Denmark, Sweden, Finland) and Benelux nations (Netherlands, Belgium, Luxembourg) maintain leadership positions with adoption rates exceeding 20%.

Figure 3 – *Transition diagram - Enterprises using AI technologies, 2023 and 2024 (% of enterprises).*

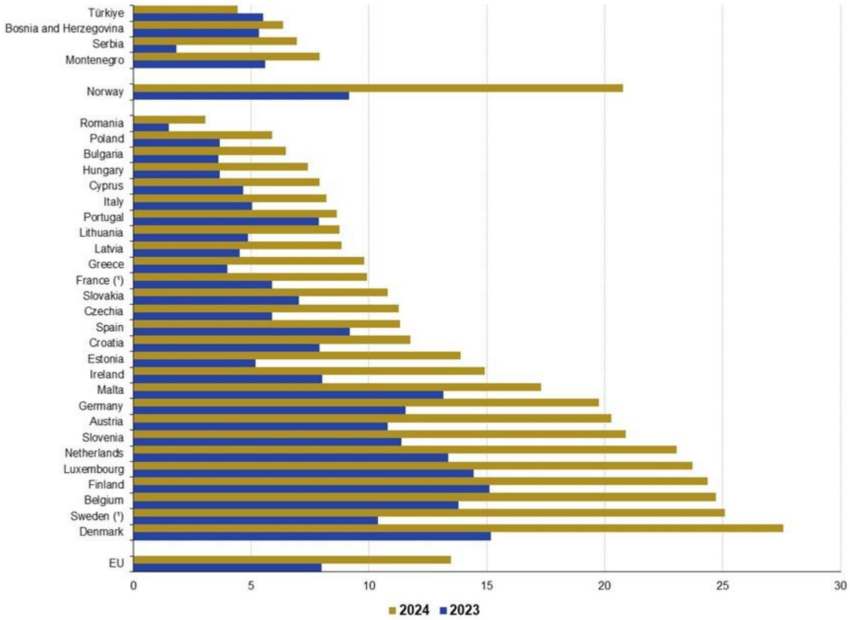


Source: Eurostat data processing.

Particularly significant is the positioning of countries: those above the regression line show accelerated growth in 2024, while those below the line exhibit more contained growth relative to the general trend. The EU average stands at approximately 13-14% in 2024, representing an increase from 2023. Figure 4 shows an overview of the evolution of AI investments in European enterprises between 2023 and 2024, confirming the important role of Nordic countries as AI technological leaders in Europe.

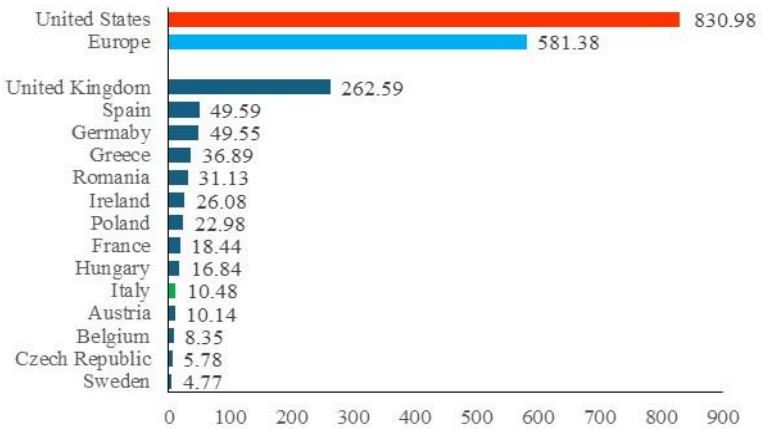
Public spending on AI contracts shows diversified trends (European Court of Auditors, 2024): the United States maintains consistent investment around 850 billion dollars (2023), while Europe presents significantly lower but growing values, with the United Kingdom (262.59 billion), Spain (43.59 billion), and Germany (49.55 billion) as the main investors (Fig. 5).

Figure 4 - *Enterprises using AI technologies, 2023 and 2024 (% of enterprises).*



Source: Eurostat data processing.

Figure 5 – *Public spending on AI-related contracts in select countries, 2023 (in billions US dollars)*



Source: Eurostat data processing.

2. Strategic Trajectories of EU Enterprises for AI Development through Exploratory PCA Analysis

2.1. Preliminary PCA analysis

The analysis was conducted on seven key indicators related to the adoption and development of artificial intelligence (AI) in European countries for 2023 (Tab. 1). Preliminary statistical validation confirmed the adequacy of the data for PCA analysis (Johnson, R. & Wichern, D.W. 2007). Examination of descriptive statistics immediately reveals the heterogeneity of the analyzed indicators in terms of means and variability (Tab. 2). Skewness indices highlight distributions deviating from normality: *AI Investments* (2.45) and *AI Patents* (2.24) with positive skewness. Conversely, the *Digital Skills* indicator (0.127) has an almost symmetric distribution. The leptokurtic distributions (kurtosis>3) for *AI Investments* (6.03), *AI Patents* (5.47), and *AI Startups* (3.91) indicate greater extreme concentrations with increased tails and outliers, while the platykurtic distributions (kurtosis<0) for *Digital Skills* (-1.28) and *AI Readiness Index* (-1.23) show opposite distributional patterns.

Table 1 – *AI indicators for European countries. 2023.*

Indicators		Source
B	AI Readiness Index	*AI Singapore (AISG)* [3]
C	AI Investiments	*European Court of Auditors and Commission EU*
D	Enterprises using AI	*Eurostat*
E	AI Startup	*Eurostat*
F	Digital Skill	*Eurostat*
G	AI Patents	*Epo Patent Index*
H	AI Publications	*Eurostat*

The combination of positive skewness and leptokurtosis in innovation indicators (*Investments*, *Patents*, and *Startups*) signals the presence of significant outliers representing European excellence hubs. Therefore, descriptive statistics make data normalization methodologically necessary to ensure analytically valid and interpretable results and justify the application of PCA as a dimensionality reduction technique (Abdi, H. & Williams, L. J. 2010).

[3] AI Singapore (AISG) constitutes a comprehensive national artificial intelligence program initiated by the National Research Foundation (NRF) with the objective of advancing Singapore's national capabilities in artificial intelligence research and development. AI Singapore. (2023). About AI Singapore. https://aisingapore.org/

Table 2 – *Descriptive statistics (a.v. and %).*

Descriptive statistics	AI Readiness Index	Investments in AI	Enterprises using AI	AI startups	Digital skills	AI patents	AI publications
Mean	65.5	962	7.76	125.0	71.7	308	1042
Median	63.2	520	6.70	95.0	70.1	195	780
Dev. Std	9.27	1156	4.30	104	9.53	279	785
CV%	14.15	120.16	55.41	83.2	13.3	90.58	75.33
Asymmetry[4]	0.234	2.45	0.548	1.97	0.127	2.24	1.74
Kurtosis[5]	-1.23	6.03	-0.817	3.91	-1.28	5.47	3.16

Pearson correlation analysis (Tab. 3 and Fig. 6) reveals significant and coherent relationships among most indicators, with coefficients ranging between 0.425 and 0.610 (all p<0.05).

Particularly strong is the correlation between the *AI Readiness Index* and *Enterprises using AI* (r=0.590, p=0.002), as well as between *AI Investments* and *AI Startups* (r=0.478, p<0.001).

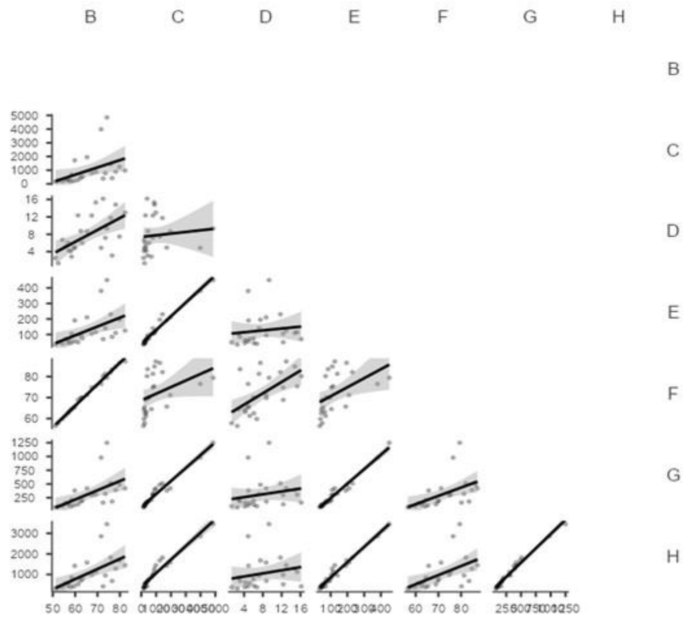
Table 3 – *Correlation matrix.*

Indicators	r	AI Readiness Index	Investments in AI	Enterprises using AI	AI startups	Digital skills	AI patents
Investments in AI	*r*	0.425 *	—				
	p	*0.034*	—				
Enterprises using AI	*r*	0.590 **	0.102	—			
	p	*0.002*	*0.626*	—			
AI startups	*r*	0.509 **	0.478 ***	0.125	—		
	p	*0.009*	*< .001*	*0.551*	—		
Digital skills	*r*	0.496 **	0.381	0.610 **	0.468 *	—	
	p	*0.003*	*0.060*	*0.001*	*0.018*	—	
AI patents	*r*	0.560 **	0.478 **	0.200	0.580 **	0.517 **	—
	p	*0.004*	*0.003*	*0.338*	*0.001*	*0.008*	—
AI Publications	*r*	0.581 **	0.567 **	0.207	0.482 **	0.537 **	0.589 **
	p	*0.002*	*0.002*	*0.320*	*0.002*	*0.006*	*0.002*

[4] Fisher's Asymmetry Index, $\gamma_1=(1/n)\ \Sigma[(x_i-\mu)/\sigma]^3$, measures how much the distribution deviates from symmetry. A value of 0 denotes perfect distributional symmetry; positive values (>0) indicate positive skewness with an extended right tail, negative values (<0) indicate negative skewness with an extended left tail.

[5] Pearson's Kurtosis coefficient, $Kurt = \Sigma(x_i-\mu)^4/n\sigma^4$, measures the degree of kurtosis as the arithmetic mean of the fourth moments of the standardized variable $Z=(x-\mu)/\sigma$.

Figure 6 – *Correlation graph.*



Furthermore, the analysis demonstrated the suitability of the data for PCA application through several statistical tests (Tabs. 4 and 5). Bartlett's Sphericity Test showed highly significant results (χ^2=382, df=21, p<0.001), confirming that variables are sufficiently correlated to justify factor analysis (Bartlett, M.S. 1950). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy reached an overall value of 0.771, considered "good" according to standard criteria, with individual values for each indicator ranging between 0.633 and 0.915, all above the acceptability threshold of 0.5 (Kaiser, M.W. & Olkin, I. 1978). Therefore, following preliminary statistical analyses, values were preventively normalized to ensure analytically valid and interpretable results using PCA.

Table 4 – *Bartlett's Sphericity Test.*

χ^2	gdl	p
382	21	< .001

Table 5 – *KMO Sampling Suitability Measure*.

Indicators	MSA
Global	0.771
AI Readiness Index	0.633
Investments in AI	0.773
Enterprises using AI	0.656
AI Startups	0.915
Digital skills	0.675
AI patents	0.893
AI publications	0.898

2.2. PCA results

Analysis of initial eigenvalues revealed a two-component structure that collectively explains 93.0% of total variance (Tab. 6 and Fig. 7). The first component explains 68.6% of variance (eigenvalue=4.803) while the second component contributes an additional 24.4% (eigenvalue=1.707). The scree plot clearly confirms this bi-factorial structure, showing a distinct *elbow* after the second component, with all subsequent eigenvalues below 1.0. This distribution indicates a strong concentration of variability in the first two factors, suggesting the existence of two fundamental latent dimensions in the European AI landscape.

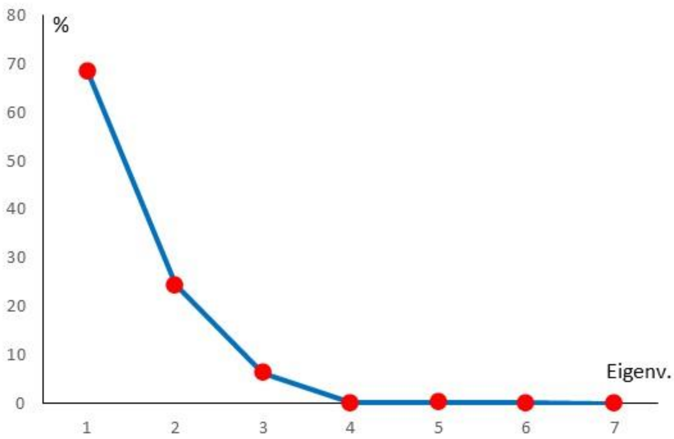
First Component: *Research and Innovation Ecosystem*. The first component is characterized by high loadings for indicators related to investments and research output: *AI Investments* (0.985), *AI Startups* (0.978), *AI Patents* (0.963), and *AI Publications* (0.958). This represents the "*Research and Innovation Ecosystem*" dimension in artificial intelligence, characterizing countries by their capacity to develop and sustain technological innovation, scientific research, and entrepreneurship in AI.

Second Component: *Diffusion and Digital Capacity*. The second component is dominated by the *AI Readiness Index* (0.888), *Digital Skills* (0.907), and *Enterprises using AI* (0.845). This represents the "*Diffusion and Digital Capacity*" dimension, reflecting the degree of AI penetration in the economic and social fabric of European countries.

Table 6 – *Initial Eigenvalues.*

Components	Eigenvalues	% of Variance	Cumulative %
1	480.323	686.175	68.6
2	170.663	243.805	93.0
3	0.44733	63.905	99.4
4	0.01826	0.2609	99.6
5	0.01688	0.2411	99.9
6	0.00608	0.0869	100.0
7	0.00158	0.0226	100.0

Figure 7 – *Cumulative Plot.*



2.3. Cluster results

PCA biplot visualization and K-means clustering highlighted positioning and segmentation in the European landscape, identifying two distinct clusters reflecting differentiated strategies in the European context (Figs. 8 and 9):

Cluster 1: *Technological and Innovation Leaders*

Includes countries from central-northern Europe such as France, Germany, Netherlands, Denmark, Finland, Sweden, and Belgium. These countries position themselves positively on both dimensions, showing both a strong research and innovation ecosystem and high AI adoption.

Cluster 2: *Countries in Transition*

Primarily comprises countries from central-eastern and southern Europe (Italy, Spain, Poland, Romania, Bulgaria, etc.) that show lower performance, particularly in the first component related to innovation. The validity indices of the two clusters, Silhouette and DB, show good separation between the two clusters (Tab. 7)[6].

Figure 8 – *K-means Plot.*

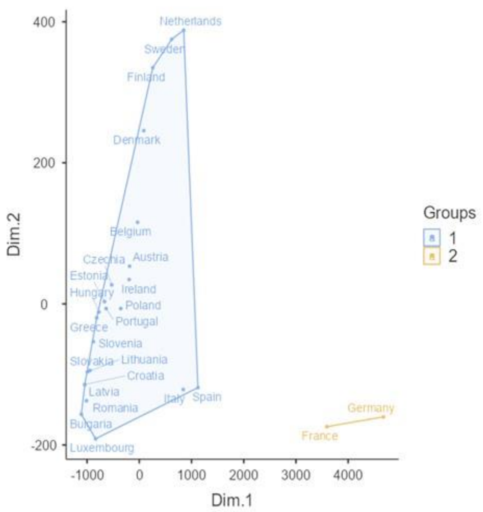


Figure 9 – *PCA – Biplot.*

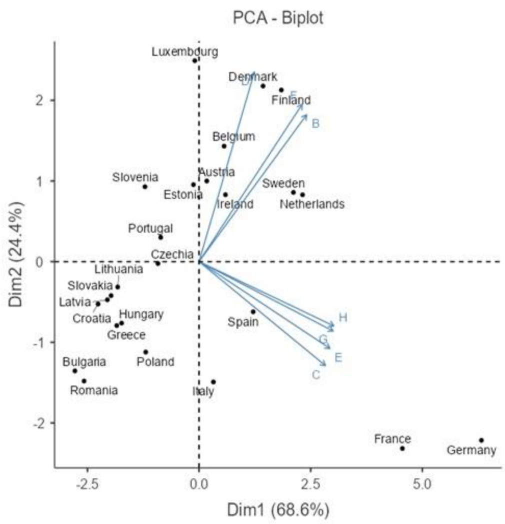


Table 7 - *Validation Indices.*

Index	Value
Silhouette	0,687
DB (Davies-Bouldin)	0,356

The strong correlation between dimensions, as indicated by the convergent arrows in the biplot, suggests that AI development requires a systemic approach that integrates technological capabilities and application diffusion.

This analysis provides an empirical framework for understanding different AI development trajectories in Europe and identifies key dimensions on which policymakers should focus efforts to promote digital competitiveness.

[6] The *Silhouette* and *DB (Davies-Bouldin) Validation indices* are used to objectively evaluate the quality of data partitioning into clusters. The *Silhouette index* calculates the average distance between a point and all other points within its own cluster, and the average distance between the point and all points located in different clusters. A *Silhouette* value close to 1 indicates good cluster separability, while a value close to -1 indicates poor separability. The *DB index* calculates the average distance between cluster centers and measures how compact and separated the clusters are. A value below 1 indicates good separation.

3. Conclusions

The principal component analysis conducted on seven key artificial intelligence indicators in European countries revealed a bidimensional structure that characterizes the European AI landscape. The results highlight the existence of two fundamental strategic trajectories that distinguish European countries in their approach to AI development.

The first dimension, *AI Innovation and Research Ecosystem*, reveals how certain countries have developed an integrated system that fuels research and development in the AI sector.

The second dimension, *Diffusion and Digital Capacity*, emphasizes the importance of the capacity for adoption and implementation of AI technologies at a systemic level.

Cluster analysis identified two distinct groups of countries: the *Innovation Leaders* (primarily Nordic and central-northern European countries) that excel in both dimensions, and the "Countries in Transition" (predominantly central-eastern and southern Europe) that present significant margins for improvement, particularly in the innovation ecosystem, Italy is part of this cluster. The presence of these two distinct clusters indicates a significant gap between leading countries and those lagging behind, suggesting that AI development requires a dualistic approach through differentiated policies to promote technological convergence at the European level.

On one hand, it is necessary to build solid foundations of research and innovation through targeted investments supporting technological entrepreneurship. On the other hand, it is fundamental to develop adoption capabilities and digital skills to ensure effective diffusion of AI technologies in the real economy.

The strong correlation between the two dimensions (cumulative variance of 93%) indicates that the most competitive countries in AI are those that successfully integrate innovative capabilities and application diffusion. This suggests that public policies should adopt a systemic vision, avoiding fragmented approaches that privilege only one of the two aspects.

In conclusion, PCA analysis provides a strategic map of the European AI landscape, highlighting how success in this sector depends on the ability to simultaneously orchestrate technological innovation and digital transformation of the socio-economic fabric (Lecardane, G. 2024). Countries aspiring to leadership positions will necessarily need to invest in both dimensions, developing robust innovation ecosystems while concurrently promoting widespread diffusion of AI skills and applications.

Countries in transition require comprehensive strategies that build both research infrastructure and digital transformation capabilities, rather than isolated interventions in specific areas (European Commission, 2025).

The bidimensional structure identified through PCA analysis provides important insights for European AI policy development.

References

ABDI H., WILLIAMS L. J. 2010. *Principal component analysis*. Wiley Interdisciplinary Reviews: Computational Statistics, Vol. 2, No. 4, pp. 433-459.
BARTLETT M.S. 1950. *Tests of Significance in Factor Analysis*. British Journal of Psychology, Vol. 41, No. 1, pp. 109-120.
EUROPEAN COMMISSION. 2025. *European approach to artificial intelligence*. Shaping Europe's digital future.
EUROPEAN COMMISSION. 2024. *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (AI Act)*. Official Journal of the European Union.
EUROPEAN COURT OF AUDITORS. 2024. Special report 08/2024: EU Artificial intelligence ambition. Luxembourg: Publications Office of the European Union.
EUROSTAT. 2025. *Use of artificial intelligence in enterprises*. Statistics Explained Statistics Explained (https://ec.europa.eu/eurostat/statisticsexplained).
JOHNSON R. WICHERN D. W. 2007. *Applied Multivariate Statistical Analysis (6th ed.)*. Upper Saddle River, NJ: Pearson Prentice Hall.
KAISER M. W., OLKIN, I. 1978. *Factor analysis*. New York: McGraw-Hill
LECARDANE G. 2024. *Official statistics measure up to the challenges of artificial intelligence*. Work presented at the Fourteenth Italian Statistics Day. Dipartimento di Scienze Politiche Sociali, Università degli Studi di Catania.
STANFORD UNIVERSITY HAI. 2025. *Artificial Intelligence*. Index Report 2025.

Giuseppe LECARDANE, Istat, lecardan@istat.it

ON THE USE OF RANDOM FOREST TO IMPUTE CATEGORICAL VARIABLES BEYOND THE SAMPLE

Ilaria Bombelli, Romina Filippini, Simona Toti

Abstract. The new paradigm of the Official Statistical production is based on a system of registers, resulting from the integration of administrative and survey data. Administrative data provide a complete enumeration of the units they cover: however, this population usually represents only a specific subset of the statistical target population, and such a subset is typically not obtained through a probabilistic sampling design. Similarly, survey data also cover only a subset of the population of interest, but they are made representative through the use of weights. Another common limitation of administrative sources is the delay in data availability.

In this context, generating a complete and consistent dataset is a critical task, which requires the implementation of specific procedures to account for delayed data and to impute missing values. One possible strategy is to use survey data as the target variable. However, this raises the issue of how to properly incorporate weights into the estimation model.

An example of a mass imputation approach is provided by the official estimates of the Attained Level of Education (ALE) adopted by the Italian National Institute of Statistics (Istat) for all the resident population in Italy. The official procedure is based on the estimation of different log-linear models.

In this application, we focus on the use of Random Forest (RF) to leverage the opportunities presented by Machine Learning (ML) technique on using all available information, including longitudinal data and variables with many categories. This approach allows for capturing complex relationships between variables, which is often challenging to incorporate comprehensively using standard methods.

The results are evaluated by focusing on different scenarios, each related to a level of available information, and focusing on three population subsets, each characterized by a different pattern of available information.

1. Introduction

Nowadays, Official Statistics is required to produce reliable and timely data to analyse and describe phenomena related to demography, economy, environment, and other aspects of society. Such data can be useful for decision makers and stakeholders to clearly understand problems and plan necessary actions. At the same time, it is required to satisfy several constraints. On the one hand, more and new statistics are demanded, requiring considerable resources, both monetary and in terms of effort. On the other hand, survey units are less likely to spend their time in filling in questionnaires, leading to high non-response rates. Therefore, Official Statistics must also address the issue of respondent burden.

In this framework, Official Statistics can exploit new data sources, such as administrative data, and new data collection methods. These alternatives can be used for statistical purposes. Moving from considering only traditional data sources and collection methods to being able to combine both traditional and new ones implies an important change in the paradigm. The new paradigm of the Official Statistical production is based on a system of registers, resulting from the integration of administrative and survey data.

The integration of administrative and survey data addresses several challenges. First, both administrative and survey data typically cover only a portion of the target population. Indeed, administrative data provide a complete enumeration of the units within the population they capture, but this population usually represents only a specific subset of the statistical target population. Similarly, survey data also cover only a subset of the population of interest, but representativeness is ensured through the application of sampling weights, derived from the survey design, and through calibration techniques. Another typical issue concerns the timing of data availability, as administrative sources are often provided with a lag time. Moreover, while survey data are specifically designed for statistical purposes, administrative data are not; therefore their integration requires additional effort, particularly in the harmonization of target parameter and the definition of target populations.

Consequently, within this framework, generating a complete and consistent dataset becomes a critical task. This requires the implementation of various procedures to predict delayed data and impute missing values. One possible strategy is to use survey data as the target variable.

An example of applying a mass imputation approach is provided by the official estimates of the Attained Level of Education (ALE) adopted by the Italian National Institute of Statistics (Istat) to cover the whole resident population in Italy.

The official procedure consists in estimating various log-linear models. In this paper, we propose an alternative approach, implementing a ML technique, specifically Random Forest (Breiman, 2001), to impute ALE. Indeed, ML approach allows to leverage all available information, including longitudinal data and variables with many categories. Moreover, these methods are able to capture complex relationships between variables that are often difficult to model comprehensively using traditional methods. However, the use of ML methods brings up the issue of how to appropriately incorporate sampling weights into the estimation process.

The remainder of the paper is organized as follows. Section 2 provides a general framework for ALE imputation, describing the data availability and their structure and detailing the official procedure based on log-linear models. Section 3 presents an overview of the Random Forest (RF) model. In Section 4, we illustrate the application of RF for ALE imputation, including a discussion on the incorporation of sampling weights and variable selection. Finally, Section 5 summarizes the main findings and discusses relevant issues and further developments.

2. Framework on Attained Level of Education

Many statistical outputs on the Italian resident population produced by Istat originate from the Base Register of Individuals (BRI), a statistical register that provides the most comprehensive coverage of individuals and contains core demographic information for each unit, gathered from various sources. Since a substantial amount of information related to education is available at the micro level from administrative sources, it is of interest to produce register-based statistics on ALE rather than survey-based. However, a key limitation of administrative data is that they are not fully up to date: for reference year *t*, administrative information on ALE is only available up to year *t-1*. Moreover, administrative sources cover only a specific subset of the target population.

To this end, a strategy has been implemented to estimate ALE for all individuals included in the BRI. This approach allows to produce official ALE statistics at different levels of detail and aggregation.

As ALE for the reference year *t* is observed only for a sample of individuals, interviewed for the Italian Permanent Census, imputations are required in order to enrich the BRI with the information on ALE, for units beyond the sample. This can be achieved by exploiting the high amount of administrative information, available at a micro level.

2.1. Data Description

For each unit of the reference population, a set of core variables gathered from different sources is available. This information is age, gender, citizenship, place of birth, and place of residence.

As regards the topic of education, longitudinal information is available from administrative sources, in particular, the Ministry of Education, University and Research (MIUR) provides information about ALE and course attendance for people entering a study program after 2011 and covers the period from 2011 to *t-1* (scholar year *t-1/t)*.

For individuals who have not attended any school course since 2011 and are therefore not included in the MIUR dataset, data from the 2011 Census are used to fill the gap. However, a residual subset of individuals—accounting for about 5% of the reference population—is not covered by any administrative sources on ALE. This subgroup consists of individuals who either entered Italy after 2011 or were present in Italy in 2011 but not captured by the Census, and who have not enrolled in any study program recorded by MIUR.

Another source containing information on ALE is the so-called APR4-module. This is the registration and cancellation form that must be filled in when applying for a change of residence, either for a new registration in Italy from abroad and/or when changing one's usual residence within the country. ALE information on APR4 is self-declared and classified into 4 aggregated categories.

Finally, an important source of information on ALE comes from the survey. Since 2018, Istat has collected ALE data through the Italian Permanent Census sample survey. As a result, directly observed ALE data for the reference year *t* are available for approximately 5% of the population, known as the Master Sample (MS).

Since Istat classification for ALE estimates is based on 8 categories, all ALE variables (with the exception of APR4) are harmonized and reclassified accordingly: 1 – Illiterate, 2 - Literate but no formal educational attainment, 3 - Primary education, 4 - Lower secondary education, 5 - Upper secondary education, 6 - Bachelor's degree or equivalent level, 7 - Master's degree or equivalent level, 8 - PhD level.

The different availability of information on ALE determines the partition of the target population into three subgroups, each characterized by a different pattern of available information.

A. Individuals who enrolled in a school course covered by MIUR in at least one school year after 2011, are characterized by a high amount of information available, including demographic information, longitudinal information on ALE, and school enrollment characteristics.
B. Individuals not in MIUR, interviewed in the 2011 Census, are characterized by the availability of demographic information and information on ALE referring to the year 2011.
C. Finally, individuals not in MIUR nor in the 2011 Census are associated only with demographic information, while ALE data are available only for a subset of them through the APR4-modules.

For each subgroup of individuals, ALE at year *t* is available for a representative sample interviewed in the MS.

2.2. *Official procedure*

The official procedure to impute ALE relies on the use of different log-linear models. Specifically, for individuals attending a school course in year *t-1*, the model predicts ALE given the year of the course attended the previous year. On the other hand, for individuals not enrolled in any school course covered by MIUR, due to the MIUR under-coverage, it is necessary to resort to sample survey data.

The conditional probabilities of each ALE category, are estimated for each profile, and the predicted class of each individual is then obtained by randomly drawing from the estimated probability distribution (Di Zio *et al.*, 2019).

3. Random Forest Algorithm

The Random Forest algorithm, introduced by Breiman (2001), is a general-purpose ML method suitable for both classification and regression tasks. It is an ensemble learning technique that generates multiple independent decision trees and combines their predictions using bagging (bootstrap aggregating). Specifically, each tree is constructed using a different bootstrap sample of the data and, at each node split, the best predictor is selected from a randomly chosen subset of features.

The final output of the RF model is obtained by aggregating the predictions of the individual trees. In regression tasks, this involves averaging the outputs of all trees. In classification tasks, the final prediction is typically based on a majority vote, although it is also possible to obtain the class probabilities instead of just the predicted class labels.

In our case, the classification forests were constructed using the Gini index as the splitting criterion. Class predictions were derived from the estimated class probabilities. Particularly, the predicted class of each individual is obtained by randomly drawing the class from the probability distribution of the classes related to each individual.

To implement RF in R, we used the ranger package (Wright & Ziegler, 2017). To achieve optimal predictive performance with RF, it is important to tune the hyperparameters — parameters that must be set before training and are not learned from the data (Probst *et al.*, 2019). Among the various hyperparameters, we focused on tuning: *mtry*, which controls the number of variables randomly selected as candidate features at each split; *min.node.sez,* which determines the minimum number of observations required in a terminal node (i.e., the minimum size of the leaves).

The results presented in the following sections were obtained using the tuneRanger R package (Probst *et al.*, 2018), which provides automated tuning of the hyperparameters for the ranger implementation of RF.

4. Application

The experiment in the present paper considers data from the Emilia Romagna Region, with the aim of producing a complete and consistent estimate of ALE for the reference year t=2022. To this purpose, a RF model is trained on the MS survey to predict the variable ALE for all residents in the Region.

The prediction of ALE for the units beyond the sample can be obtained from RF mainly in two ways: using the majority voting approach, i.e., predicting the class with the highest probability; predicting the class randomly drawn from the RF estimate probabilities.

Since our goal is to generate predictions that closely matches the distribution of ALE in the MS, the results are evaluated by assessing the degree of alignment between the predicted distribution and the benchmark distribution provided by the MS. As a consequence, we adopt the second approach as it provides a higher coherence at the macro-level (De Fausti *et al.*, 2022).

The majority voting approach, indeed, allows for good performance at the micro-level. However, it completely reduces all probabilistic information to a single category, thus losing the shape of the predicted distribution. In contrast, the random-draw approach allows the shape of the predicted distribution to be preserved, avoiding the distortion induced by the majority voting and preserving the variability (see, for example, Little and Rubin, 2002). In this sense, the second approach allows for better performance at the macro-level.

4.1. Sampling weights usage

The probabilistic sampling design of the MS assigns to each unit a sample weight that quantifies the number of population units that the unit represents.

Standard statistical model requires to include sampling weights at the likelihood level. In the ML framework, different possibilities are explored to consider the sample weights; we focus on two options: a) Ignoring them; b) Considering them by building the augmented version of the sample, by repeating each record a number of times given by the sampling weight of the record.

Option a) ignores the statistical probabilistic nature of the sample; in option b), the input dataset is augmented by replicating each row according to its sampling weight. In this way, the augmented MS mimics the target population structure.

In the initial experiments with RF, the same set of input variables used in the log-linear models is considered. Specifically, these variables are described in Section 4.2 and summarized in Table 2, Scenario 1.

Results are evaluated at the macro-level by comparing the distribution of ALE across the whole population with the benchmark, represented by the ALE distribution in the augmented MS, corresponding to the weighted ALE distribution. Performance is measured by computing the difference between the percentage relative frequencies of the predicted distributions and those of benchmark distribution.

Table 1 – *Absolute (a.v.) and percentage (%) distribution of ALE in the Augmented MS (benchmark). Differences between predicted and benchmark distribution.*

ALE	Augmented MS		Estimated – Augmented MS		
	a.v.	%	Log-Linear	RF Ignoring weights	RF Augmented data
1	16717	0.4%	0.05%	-0.01%	-0.02%
2	130504	3.2%	-0.21%	-0.25%	-0.25%
3	585713	14.2%	0.25%	0.09%	0.05%
4	1135472	27.4%	0.02%	-0.13%	-0.11%
5	1571471	38.0%	0.27%	0.32%	0.33%
6	187616	4.5%	-0.14%	0.00%	0.00%
7	486049	11.7%	-0.25%	-0.05%	-0.03%
8	23777	0.6%	0.00%	0.03%	0.02%
	Sum of absolute values		1.20%	0.88%	0.81%
	Mean of absolute values		0.15%	0.11%	0.10%

The results in Table 1 show that applying RF models leads to better performance, measured in terms of the sum of absolute differences, compared to log-linear models. The improvement is particularly evident for ALE classes 3 (Primary education), 6 (Bachelor's degree), and 7 (Master's degree). Since both models rely on the same set of input variables, the performance gain can be mainly attributed to the ability of RF to capture non-linear relationships.

Moreover, the inclusion of sampling weights in the construction of the augmented MS leads to slightly improved results.

Overall, the predictions obtained with the RF that incorporates sampling weights (applied to the augmented MS) are more consistent with the benchmark distribution than those obtained without using weights. Specifically, the sum of the differences between the estimated and benchmark distributions is 0.88% when weights are ignored, and 0.81% when weights are included.

Therefore, in the following analyses, the reported results refer to RF models trained on the augmented MS data.

4.2. Variables selection and encoding

After establishing the approach for accounting for sample weights and using the augmented MS as input data, we explored the full space of potential predictors to identify those to include in our model. We implemented three distinct scenarios to evaluate the effect of adding predictors and different levels of aggregation for predictor categories. These scenarios are designed to assess the contribution of additional data and the potential of ML techniques to capture complex relationships.

The first scenario (Scenario 1) includes the same variables used in the log-linear regression model. This allows for a direct comparison of performance between the methods, highlighting the ability of ML techniques to capture non-linear relationships. This feature set includes demographic characteristics, *i.e.*, age, gender, citizenship, and province of residence; educational data for the school year *t-2/t-1*, *i.e.*, ALE, school enrollment status, year of attendance, type of upper secondary school.

The second scenario (Scenario 2) includes the same variables used in Scenario 1, with more granular levels of detail for the variable Type of upper secondary school: from 3 to 28 categories.

Table 2 – *Description of input variables: type, presence of missing values and, encoding used in the RF. An: "X" indicates the inclusion of each variable in the scenarios.*

Variable	Type	anyNA	Encoding	Scen. 1	Scen. 2	Scen. 3
Age	Numerical		Ordinal	X	X	X
Gender	Binary		Dummy	X	X	X
Italian or not	Cat. (2 classes)		Dummy	X		X
Province of residence	Categorical		Dummy	X	X	X
Subgroup	Cat. (3 classes)		Dummy	X	X	X
ALE from Administrative Sources (MIUR or Cens11)	Cat. (8 classes+1)	X	Ordinal	X	X	X
ALE from APR4	Cat. (4 classes+1)	X	Ordinal	X	X	X
Year of attendance	Cat. (32 classes+1)	X	Ordinal	X (*t-1*)	X	X (*t-1 ; t-2*)
Evening school	Cat. (2 classes+1)	X	Dummy			X (*t-1 ; t-2*)
Type of attendance	Categorical	X	Dummy			X *t-1 ; t-2*)
Type of primary/lower secondary school	Cat. (2 classes+1)	X				X (*t-1 ; t-2*)
Type of upper secondary school	Cat. (2 classes+1)	X	Dummy	X (*t-1*)		
Type of upper secondary school	Cat. (28 classes+1)	X	Dummy		X	X (*t-1 ; t-2*)

Finally, Scenario 3 is implemented to fully exploit the potential of ML methods. This subset incorporates additional information on school attendance (type of primary and lower secondary school, evening school, and type of attendance) as well as data from previous years.

Table 2 provides a synthetic overview of the three scenarios.

In the three scenarios, nominal categorical predictors are converted into dummy variables (in R factors), and missing data are replaced with an “out-of-range” value, namely 99 or “99”.

A comparison of the results across the different scenarios (Table 3) shows that, as the amount of available information increases (from Scenario 1 to Scenario 3), overall prediction performance improves accordingly, with Scenario 3 yielding the best results.

Table 3 – *Summary statistics of differences between % relative frequency distribution obtained by RF model and the % relative frequency distribution of benchmark distribution, separately by subpopulations.*

	Overall	Pop A	Pop B	Pop C
Augmented data – scenario 1				
Sum of absolute values	0.81%	1.92%	0.79%	1.74%
Mean of absolute values	0.10%	0.24%	0.10%	0.22%
Augmented data – scenario 2				
Sum of absolute values	0.76%	1.90%	0.77%	1.97%
Mean of absolute values	0.09%	0.24%	0.10%	0.25%
Augmented data – scenario 3				
Sum of absolute values	0.69%	1.79%	0.78%	2.80%
Mean of absolute values	0.09%	0.22%	0.10%	0.35%
Augmented data – scenario 3 Fixed variables				
Sum of absolute values	0.82%	1.79%	0.84%	1.55%
Mean of absolute values	0.10%	0.22%	0.10%	0.19%

This trend is evident when focusing on population A, characterized by more informative profiles. For population B, the predictive performance of the RF remains stable across the three scenarios. In contrast, for population C, the trend is the opposite of that observed for population A—that is, prediction performance worsens as the amount of information increases. This occurs because the third scenario introduces additional variables that are informative only for population A and B. In RF, the trees are built by randomly selecting a subset of predictors from the full set (which is larger in Scenario 3) to evaluate at each split. As a result, the probability of selecting the informative predictors—those useful for estimating population C based only on demographic information—decreases from Scenario 1 to Scenario 3.

Indeed, the tuning across the three scenarios selects a high number of predictors: in all cases, the *mtry* value is very close to the total number of covariates, and the leaf size is very small. This indicates that the tuned RF trees are very deep and make use of nearly all the available covariates.

In the latest experiment, demographic information was constrained to be included at every split. The results of the RF model built in this way are comparable to those obtained in the three scenarios, with the best performance in predicting population C. However, in this case, the *mtry* value is much lower than the total number of predictors—approximately equal to the square root of the number of predictors.

5. Discussion

In this work, we showed how the use of ML technique can help for imputation purposes. We proposed to impute the ALE for units beyond MS using RF model.

The evaluation of the results was carried out at a macro level, using as a benchmark the distribution derived from the augmented MS with sampling weights.

When comparing ALE predictions from the official LL procedure with those generated by the RF model—based on their divergence from the benchmark distribution—the RF model shows superior performance, particularly when trained on the augmented sample.

In the second part of the analysis, we explored the possibility of incorporating all available information into the model. ML models generally offer the advantage of handling a high number of input features, including fully disaggregated categorical variables. This flexibility represents a key strength of ML approaches compared to traditional statistical models.

We designed three scenarios to evaluate the contribution of additional information and to assess the ability of the RF model to leverage this information across three subpopulations, each characterized by distinct information profiles. Overall, the results show a clear trend: the difference from the benchmark distribution decreases with a progressively closer alignment as we move from the least informative to the most informative scenario.

However, the improvement manifests differently across subpopulations: there is a clear gain for the population with more populated regressors (subpopulation A), relative stability for the group with moderately informative profiles (subpopulation B), and a marked deterioration for the subpopulation with only demographic information (subpopulation C). In fact, the increase in the number of regressors makes the selection of the informative predictors for ALE in subpopulation C less likely during the tree-splitting process. Even for subpopulation C, the results are comparable to the overall ones, because the increase in the number of regressors is balanced by the hyperparameter *mtry*, which is appropriately tuned.

Usually, the suggested default value for *mtry* is close to the square root of the number of regressors (see, for example, Breiman, 2002). However, as the author suggested, when very noise regressors are present, the *mtry* parameter should be higher. In our application, the patterns of available information differ substantially across subpopulations. For sub-population C only three covariates are informative while the remaining regressors are essentially noise. For this reason, the tuning process indicates that the optimal value of *mtry* is close to the total number of regressors. However, if the trees are forced to consider, at each split, the non-missing regressors of subpopulation C among the randomly selected candidates, then the overall prediction performance becomes comparable to that of the other RF. In this case, prediction for subpopulation C improves significantly, with the optimal *mtry* value close to the square root of the number of regressors.

In conclusion, this research shows that the suggested value of *mtry* cannot be applied without taking into account the structural characteristics of the data. While it is a good choice when information is uniformly distributed across the population and no noisy regressors are present, it becomes inadequate in scenarios characterized by heterogeneous information patterns and the presence of noisy regressors, as in the case of our study.

The flexibility of the RF structure, which can adapt to different prediction scenarios through the specification of different hyperparameters, makes the tuning phase essential and allows for the semi-automatic handling of prediction problems in presence of highly heterogeneous populations.

In general, when the population of interest is highly heterogeneous in terms of available information, ML provides a valuable tool, i.e., the ensemble model. The idea of an ensemble model is to combine different ML techniques, each one specifically suited for a subpopulation, to obtain a final prediction for each unit. This approach can be useful in our application. As a further development, we plan to implement additional ML models, better suited for subpopulations B and C, such as XGBoost (Chen and Guestrin, 2016), and Neural Network (Rumelhart *et al*., 1986). Finally, we aim to combine these different ML models into a single ensemble model (see, for example, Dietterich, 2000).

References

BREIMAN L. 2001. Random forests, *Machine learning*, Vol. 45, pp. 5–32.

BREIMAN L. 2002. *Manual On Setting Up, Using, And Understanding Random Forests V3.1.*

CHEN T., GUESTRIN, C. 2016. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, pp. 785–794.

DE FAUSTI F., DI ZIO M., FILIPPINI R., TOTI, S., ZARDETTO D. 2022. Multilayer perceptron models for the estimation of the attained level of education in the Italian Permanent Census, *Statistical Journal of the IAOS*, Vol. 38, No. 2, pp. 637-646.

DIETTERICH T.G. 2000. Ensemble Methods in Machine Learning. In *Multiple Classifier Systems. MCS 2000. Lecture Notes in Computer Science*, Vol. 1857. Springer, Berlin, Heidelberg.

DI ZIO M., FILIPPINI R., ROCCHETTI G. 2019. An imputation procedure for the Italian attained level of education in the register of individuals based on administrative and survey data. *Rivista di Statistica Ufficiale*, Vol. 2, No. 3, pp. 143-174.

LITTLE R.J.A, RUBIN D.B. 2002. *Statistical Analysis with Missing Data*. Wiley, New York.

PROBST P., WRIGHT M., BOULESTEIX A. 2018. Hyperparameters and Tuning Strategies for Random Forest, *Data Mining and Knowledge Discovery*.

PROBST P., BOULESTEIX A., BISCHL B. 2019 Tunability: Importance of Hyperparameters of Machine Learning Algorithms, *Journal of Machine Learning Research*, Vol. 20, pp. 1-32.

RUMELHART D. E., HINTON G. E., WILLIAMS R. J. 1986. Learning representations by back-propagating errors. *Nature*, Vol. 323, No. 6088, pp. 533–536.

WRIGHT M.N., ZIEGLER A. 2017. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R, *Journal of Statistical Software*, Vol. 77, No. 1, pp. 1–17.

Ilaria BOMBELLI, Istat, ilaria.bombelli@istat.it
Romina FILIPPINI, Istat, filippini@istat.it
Simona TOTI, Istat, toti@istat.it

WORKING ANYTIME, WORKING ANYWHERE AT THE ITALIAN NATIONAL INSTITUTE OF STATISTICS (ISTAT)[1]

Chiara Limiti, Giuseppina Sola

Abstract. The COVID-19 pandemic prompted a widespread shift toward remote work, significantly altering traditional work arrangements and locations. This transformation has had a profound and enduring impact on the operational modalities of Italian Public Administrations (PAs). This study investigates the smart working model implemented by the Italian National Institute of Statistics (ISTAT), with the aim of assessing the frequency and intensity of both in-office and remote work practices. The novelty of this work lies in its focus on the evolving organizational frameworks within the public sector -such as workflow management, team dynamics, leadership styles, and institutional culture - where smart working remains a developing paradigm. By offering quantitative insights into the application of smart working, this study contributes to the broader discourse and supports future research in this domain.

1. Introduction

The COVID-19 pandemic catalyzed the adoption of remote work, exposing both its potential and limitations, even within traditionally rigid public sector structures. Smart working has since evolved from an emergency response into a strategic lever for modernizing public administration in the digital era.

Evaluating its impact requires a multidimensional approach, encompassing productivity, institutional efficiency, transparency, service delivery, and employee well-being. In the Italian context, where public administration plays a pivotal role in socio-economic governance, such analysis must also address how the sector addresses the digital divide and supports the development of digital infrastructure and policy. Furthermore, it is imperative to comprehend the manner in which smart working contributes to the enhancement of work-life balance, employee satisfaction, and organisational culture, enhancing public service effectiveness.

In the following discussion, the characteristics and differences between remote working, smart working, teleworking, will not be examined in depth. For a more detailed analysis of these concepts, the reader is referred to the relevant specialist

[1] This article is the result of the common contribution of the joint study of the two authors.

literature on the subject (ILO, 2020, Pareschi, 2021, CIPD, 2008). We shall instead refer to the concept of smart working or "*lavoro agile*" as introduced into the Italian legal system by Art. 18, Law No. 81 of May 22nd 2017 as a "*method for the execution of the employment relationship established by agreement between the parties, even with forms of organization by phases, cycles and objectives and without precise constraints of time or place of work, with the possible use of technological tools for the performance of the work activity. The work activity is performed, partly inside company premises and partly outside without a fixed location, within the limits of maximum duration only of daily and weekly working hours, deriving from the law and collective bargaining*". In brief, smart working is subject to regulation through a voluntary agreement that is embedded within the broader employment relationship. This agreement confers upon employees the autonomy to determine their working hours and location, subject to general arrangements established with the employer. In contrast, remote work is formally defined in the employment contract itself – or subsequent amendments – and involves fixed specifications regarding the place of work, the tools to be used, and the working schedule.

The paper is structured as follows. The next section reviews the organizational measures adopted by the Italian National Institute of Statistics (ISTAT) to comply with the legislative framework governing the implementation of smart working. The third session outlines the data sources and methodological choices made during the data processing. The fourth section presents the results the empirical findings, with particular attention to gender-related outcomes. The final section presents potential avenues for future research on working time.

2. Smart working in the years 2023 and 2024: ISTAT internal provisions

In the aftermath of the COVID-19 pandemic, ISTAT implemented a hybrid work model grounded in the principle of prevalence, mandating that 51% of workdays be conducted in-person and 49% remotely. Between 2023 and 2024, the Institute experimented with two scheduling methodologies—monthly and bi-monthly—to optimize flexibility, work-life balance, productivity, and resource management without overburdening facility managers.

In detail, as a result of the various internal regulations that followed, three distinct regulatory phases were observed:

a. Jan–Feb 2023: A bi-monthly cap of 20 smart working days was introduced, with a mixed modality (16 full days + 4 half-days). "Fragile"[2] and special teleworking employees were granted full remote work exemptions.
b. Mar 2023–Apr 2024: A monthly cap of 10 days was adopted (8 full + 2 half-days). Decentralized work from one ISTAT office different from the one where the assigned structure is located was enabled, and exemptions for vulnerable employees were maintained. Additional smart working days were allowed in cases of workplace unavailability.
c. May 2024–Feb 2025: Two profiles were defined—ordinary (20 days/2 months) and enhanced (24 days/2 months), the latter for employees with documented personal or health-related constraints. Mixed modality, location-based exceptions and workplace unavailability were upheld. A medical review process was introduced to authorize temporary full remote work for vulnerable employees.

To summarize, the Institute's smart working model offers high flexibility in terms of when and where work is done. This innovative model reorganizes work and promotes a healthy work-life balance while adapting to a changing work world. This model constitutes a marked deviation from the paradigm of more flexible telework. In contrast to traditional teleworking practices - typically characterized by home-based work with fixed schedules - the adopted framework imposes only two specific constraints on employees' autonomy in managing their work-life balance: (1) a cap on the number of days that may be performed in smart mode, and (2) a defined contactability window, which delineates the time frame during which employees are expected to be reachable for communication and meetings. Nonetheless, the scheduling of in-office and smart working days, as well as the contactability band, are not imposed; rather, they are defined in consultation with the service manager, upon the proposal of the employee(s). Moreover, given the absence of compulsory return days in the workplace, work performance in smart mode is not constrained by specific days of the week or limited by the enjoyment of other institutions of absence. It is particularly salient to note that no regulations have been established to prevent or limit work. Specifically, since no rules have been determined that prevent and/or restrict smart working in relation to vacation or leave, it is possible to work agile in the days before or after vacation is taken, without the obligation to make any return to headquarters.

To ensure the continuity of statistical production processes and the organization's general functioning, ISTAT provided laptops and connectivity tools, enhanced the

[2] The definition of the category of "fraigile workers" was entrusted to a Decree of the Ministry of Health dated Feb. 4, 2022, which includes patients with situations of severe immune system impairment, patients with at least 3 or more concomitant serious diseases among those indicated by the same DM, and people with documented exemption to vaccination.

technological infrastructure, and acquired and deployed software tools — specifically, technologies for video conferencing. The organization also offered support to employees through various telematic channels to ensure proper configuration and training. ISTAT has adopted advanced solutions for secure data and information access in order to comply with security standards. Virtual Desktop Infrastructure (VDI) and Virtual Private Networks (VPN) were configured, and additional servers were set up to provide the aforementioned services. ISTAT is implementing a progressive migration to the cloud of applications, software and documents with a view to fostering collaboration among workers, even when they are working remotely.

ISTAT employees like hybrid work - i.e., work performed in a mixed mode between presence at the office and one of the forms of remote work[3]. On 31 December of each year, only 1% of the in-service workforce, respectively amounting to 1,851 and 1,843, choose voluntarily not to use any form of remote work. In contrast, 90% of employees confirmed their choice to join smart working. There are no significant differences in the choice of doing part of one's work in agile mode by gender, age groups and occupational position[4]. Considering the employees in service on 31 December of the year, with respect to gender, women choosing to work in agile mode in the two years considered exceed 92% of the female employees in service; men exceed 88% of the employees in service (Figure 1).

With respect to age groups, the percentage of employees who voluntarily decide to take advantage of smart working, while maintaining values above 80% of the current workforce, decreases as age increases (see Figures 2).

Reading the figure without knowing the composition of the workforce could lead to the assumption that older employees, free from the responsibilities of caring for

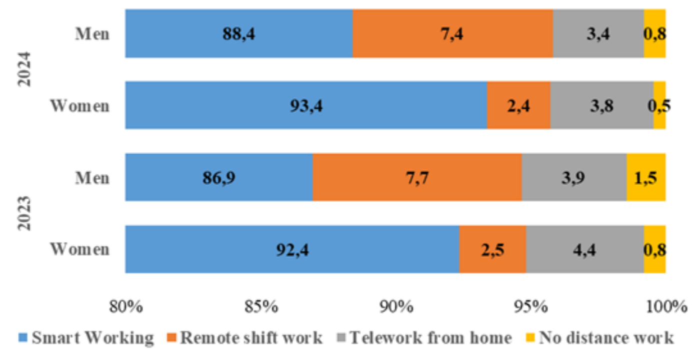
---

[3] Smart working adoption in Italy was significantly lower than in other E.U. Countries. According to Eurostat data, in 2023, a mere 4.4% of Italian workers had the capacity to perform at least half of their weekly hours in a remote capacity. This figure stands in stark contrast to the European average of 9%. Following Italy, Cyprus (3.9%), Hungary (3.7%), and Romania (1.2%). Finland (21.7%) and Sweden (14.3%) are at the top of the list (cfr. https://data.europa.eu/). As reported by the Smart Working Observatory of the Politecnico di Milano, in the 2024, 61% of Italian public administration organizations had adopted smart working initiatives. This percentage increases to 70% considering all remote working opportunities (cfr. https://www.osservatori.net/).

[4] In the framework of Italian administrative law, public sector employment is typically categorized into executive personnel and non-executive personnel. Executive personnel (or *personale dirigente*) are entrusted with managerial, organizational, and decision-making responsibilities, often acting as the heads of administrative units or departments.. In contrast, non-executive personnel (*personale non dirigente*) perform operational, technical, or clerical functions under the direction of executives, without holding autonomous decision-making authority. This structural distinction reflects the broader principle of the separation between political direction and administrative management, as established by legislative reforms such as Legislative Decree No. 165/2001, which delineates the respective roles and responsibilities within the public administration.

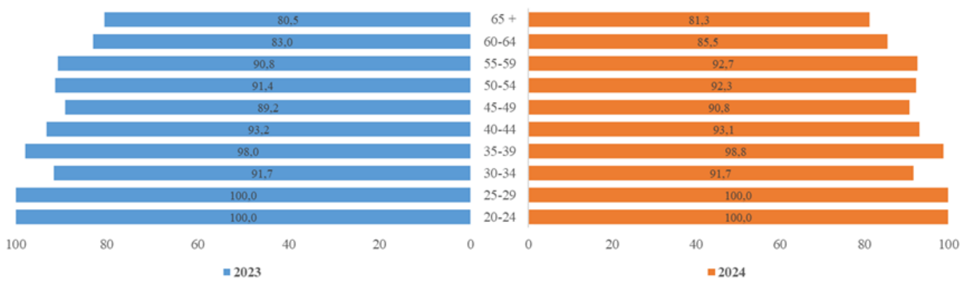
young children and elderly parents, prefer to work on-site rather than remotely. The hypothesis is considerably weakened by the fact that the staff who take advantage of the other flexible work modes, namely special teleworking and remote shift performance, which likewise do not involve onsite presence, belong to the more adult age group.

Figure 1 - *Employees in service at 31 December by types of work, gender and year. (% values).*



Source: authors' elaborations based on Istat administrative data

Figure 2 - *Smart workers in service at 31 December by class of age and year (% values).*



Source: authors' elaborations based on Istat administrative data

Regarding occupational position, we note that almost all executive personnel (nearly 99% in 2024) and non-executive staff (91% in 2024) adhere to smart working.)[5].

The adoption of smart working across all gender, age groups and professional categories is a significant indicator of organizational well-being. It promotes fairness, inclusion, autonomy and the empowerment of people in the workplace without reducing the effectiveness of management and productivity. Access to agile working is shown to strengthen the perception of organizational justice, reduce the risk of discrimination and promote a more equitable working environment. This voluntary choice fosters enhanced involvement and participation of all employees in business processes, thereby stimulating clear objectives, direct communication, and consistency. Relational dynamics based on trust, responsibility and autonomy are developed and help overcome the traditional opposition between managers and employees. The agile management style of these leaders demonstrates that leadership can be executed with a high degree of flexibility. By focusing more on collaboration and less on physical presence, managers set a new example and prove the model works, integrating it into the organization's strategy. When all staff adopt smart working, it helps effective change management and new digital and organizational policies. It can optimize processes and cut overheads (e.g. offices, travel) without risk of internal resistance and process fragmentation.

3. Working at ISTAT: methodology and data source

The underlying assumption of this analysis is that work performance can be categorized into four distinct types: work in the presence at one of the Institute's Roman or territorial offices (i.e. office work), smart working, vacation and other paid leave. The term "vacation" encompasses vacation days, compensated time off and compensated absences for increased services rendered. The category designated as "other paid leave" encompasses absences resulting from leave or those situations explicitly addressed and legally protected by current legislation (i.e. Law No.104/92 permission, parental care absence, etc.).

The calculation of working days was performed by aggregating the justification codes, attributable to the previously mentioned four modes, recorded by the URBI personnel management system adopted by the Institute to certify work performance

[5] Executive personnel includes 1st and 2nd level administrative managers, technical directors and service managers. non-executive is composed of the remaining staff classified in the profiles of researcher and technologist and that of civil servant, collaborator and technical and administrative worker. In 2023, executive personnel is 70 and non-executive personnel is 1,781. In 2024, these two distinct group are 71 and 1,772.

(attendance/absences). The analysis exclusively considers personnel in service as of the last day of the month analyzed, thus excluding personnel on secondment, detached, or absent due to other types of expectations.

In analyzing the data, it was necessary to take into account the following elements. Firstly, the work performance rendered in person can be greater or less than 7 hours and 12 minutes, which is the contractual daily working time stipulated by the CCNL of the "Education and Research" sector. The hourly flexibility recognized by the CCNL and the current internal regulation of working time allows the individual employee to determine, based on their needs, the amount of daily hours effective worked. The hours worked in reality are allocated to a "virtual hour bank," which is accessible to employees in accordance with the provisions concerning the legal classification profile (researcher/technologist or technical-administrative staff). Consequently, in the subsequent analysis, any day on which the employee's timecard records 7 hours and 12 minutes is designated as an in-office day[6]. Secondly, some paid leave can be utilized on an hourly basis. Finally, internal regulations have established a mixed mode for the implementation of smart working. That said, in order to calculate the daily hours of effective work, the office work and the mixed smart working day were parameterized at 7 hours and 12 minutes.

In addition, it was decided to count smart working justification codes in the month they were used, regardless of whether internal regulation defined use on a month or two-month basis. This choice facilitates a monthly comparison of the data, but may obscure the effects of seasonal absence due to traditional vacation periods.

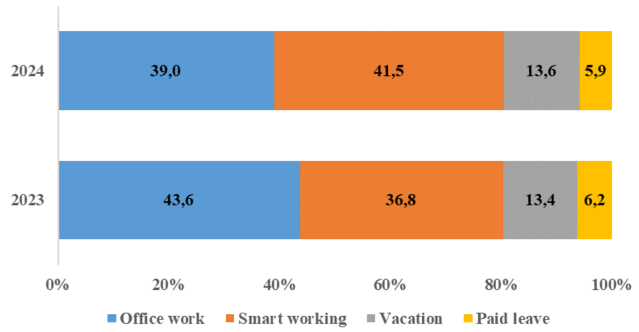
4. Working anytime at ISTAT

ISTAT employees are increasingly comfortable with hybrid work arrangements, whereby job duties are performed partly on-site at one of the institution’s locations and partly through smart working. In 2023, out of a total of 456,526 workable days, approximately 37% were worked remotely, 44% in-office, and the remaining 19% were not worked due to vacation, leave, or illness. In 2024, out of a total of 457,595 workable days, the proportion of days worked remotely increased to 41% (a rise of 5 percentage points compared to the previous year), while the share of office-based

[6] On average, employees working on-office tend to exceed the standard daily working time of 7 hours and 12 minutes. In 2023, the raw count of badge swipes certifying physical presence at the workplace amounted to 198,913 days. When adjusted for standard working time, this figure corresponds to 199,173 days - indicating an excess of 260 full working days. Similarly, in 2024, the raw count of on-office attendance days was 178,106, which, once standardized, equates to 178,531 days - reflecting an additional 425 full working days.

work declined to 39%. Absences, including vacation and other forms of paid leave, remained stable at 19% (see Figure 3).

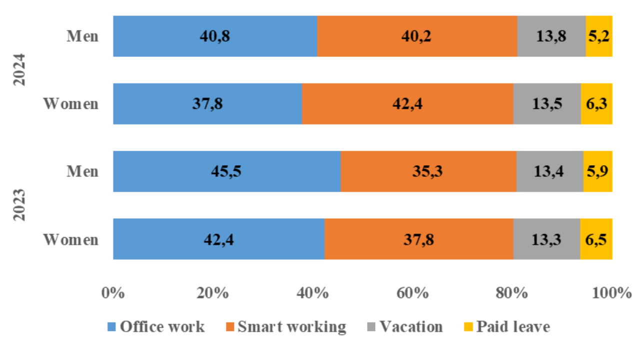
Figure 3 - *Annual days worked by employees in service by types and year (% value).*



Source: authors' elaborations based on Istat administrative data

The percentage composition of annual work performance by gender, relative to total working days, shows that in both years analysed, female employees used smart working more than their male counterparts (see Figure 4).

Figure 4 - *Annual days worked by employees in service by type, gender and year (% values).*



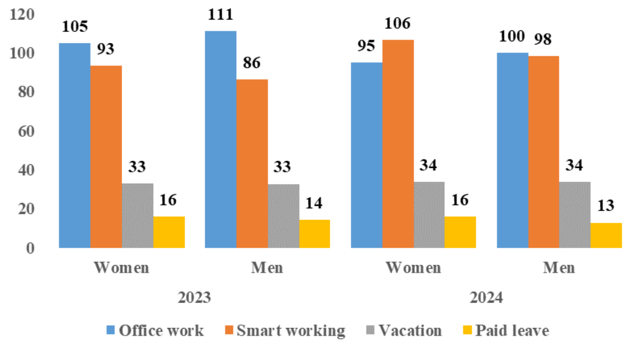
Source: authors' elaborations based on Istat administrative data

In 2023, out of 274,182 workable days for women, 38% were worked remotely, 42% in-office and the remaining 19% were accounted for by absences due to vacation, leave, or illness. In contrast, of the 182,343 workable days for men, 35% were worked remotely, 45% in-office and 19 % were similarly attributed to absences.

In 2024, among the 276,364 workable days for women, 42% were worked remotely, 38% in-office and 19% were lost to vacation, leave, or illness. For men, out of 181,231 workable days, 40 % were worked remotely, 41% in-office and the remaining 19% were due to absences.

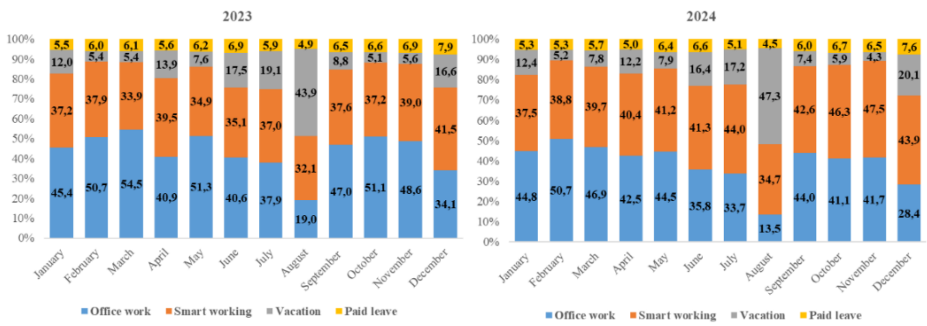
Average days show similar trends for the two genders: performance in smart working increased in 2024 compared to 2023 (+14% for both) (see Figure 5) with monthly variations showing peaks in certain months (Figure 6); office work and other paid leave decreased in 2024 compared to 2023.

Figure 5 - *Annual days worked by employees in service by type, gender and year (average values).*



Source: authors' elaborations based on Istat administrative data

Figure 6- *Annual days worked by employees in service by type, month and year (percentuage values*).



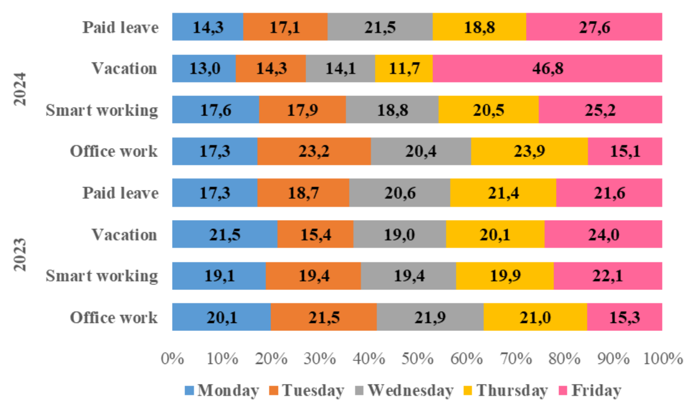
Source: authors' elaborations based on Istat administrative data

The monthly distribution of working time reveals, as expected, fluctuations with peaks in absences during certain months. In both years analyzed, the lowest levels of on-site attendance were recorded in the months traditionally favored by Italians for holidays - namely July, August, and December. During these months, the number of vacation days increases significantly.

The months with the highest proportion of office attendance are February, March, and May. These are periods in which ISTAT employees are particularly engaged in ensuring the timely release of official statistics and related products. As a result, recourse to overtime becomes necessary, including extended working hours beyond the standard 7 hours and 12 minutes, as well as work on holidays (Saturdays and Sundays) through extraordinary openings of office locations. Finally, it is worth noting that the use of vacation days affects only on-site attendance, while the proportion of smart working remains relatively stable throughout the months considered.

The cultural model of "working anytime" has become so deeply embedded in the organizational culture of the Institute that there are no designated days for either on-site or smart working (see Figure 7). An analysis of the weekly distribution of annual work performance across the two years under comparison reveals that ISTAT employees have selected their workdays based on the optimal reconciliation of professional and personal responsibilities.

Figure 7 - *Annual days worked by employees in service by types, week-day and year (percentuage values).*



Source: authors' elaborations based on Istat administrative data

In 2023, Monday emerged as the most preferred day for on-office work (27.7%), whereas in 2024, Tuesday held that position (26%). Conversely, the lowest levels of office attendance were recorded on Tuesday in 2023 (11%) and on Thursday in 2024 (4.7%).As regards smart working, in the two years the preferences are reversed: if in

2023 the preferred days to carry out their work performance in agile were Monday (29 %) and Friday (25 %), in 2024 the choice is reversed. The days when "like" to work in agile less are, in both years, on Thursday and Tuesday.

As was to be expected, the preferred days for the enjoyment of vacation and compensatory absences for higher performances are Friday and Monday (long weekend effect).

5. Conclusion

ISTAT employees are increasingly looking for flexible working arrangements, such as remote working (from home and anywhere), hybrid models and flexible hours, to improve work-life balance. As evidenced by the absence of preferred days for working in presence or agile, ISTAT employees seem to take full advantage of not being obliged to come to the office every day to work from their usual desk, but rather to be able to achieve their goals wherever they want, without location constraints. In this way, they can have the flexibility to design their days so that their private and professional lives can be lived to their full potential and coexist in a flexible and non-conflicting manner.

The hybrid work model currently being consolidated at the Institute necessitates a profound cultural shift in the conceptualization of work. The physical office must increasingly serve as a space for collaboration and interpersonal engagement, technologies must facilitate collaboration and thus guarantee maximum flexibility and mobility. Employees must be empowered and managed through goal-oriented approaches, grounded in mutual trust between colleagues and with management, in order to foster both productivity and well-being.

In the context of this transformation, the Institute must address not only the transition from the physicality of individual workstations to shared spaces (e.g., desk-sharing, co-working) or virtual environments (e.g., cloud-based systems), but also the need to establish clear boundaries between professional and private life. This includes, for example, the introduction of virtual time-sharing mechanisms aimed at safeguarding individual health and well-being. Moreover, it is essential to develop new strategies to strengthen professional relationships and to prevent physical distance from undermining the quality of interpersonal interactions, the overall social climate within the Institute, and organizational well-being.

To support people in changing their working methods and in the use of digital tools, ISTAT has organized training courses and implemented a broad change management program with the aim of raising awareness among people about the correct management of work-life balance and providing them with the necessary digital skills and soft skills. The Institute has also promoted courses to improve organizational culture and leadership style. At the same time, a section of the Institute's intranet has been developed for continuous training on the use of technological systems.

However, today the Institute should invest in knowledge, that is, carry out an internal survey to allow workers to report the possible negative effects of smart working actually experienced. The ideal would be to carry out an internal survey on an annual basis, given that the latest available data on the matter can be extracted from a survey carried out by ISTAT (CUG)[7] back in 2021. Several studies (Li *et.al*, 2024; Milea, 2024) highlight that many workers report difficulties in separating private life from work time. The absence of clear boundaries can lead to hyperconnection and overworking phenomena, with a consequent increase in work-related stress and fatigue. Another of the most frequently reported risks is the isolation of workers. Remote working reduces opportunities for informal discussion, with negative effects on collaboration and psychological well-being, on the perception of opportunities for professional growth. All this can compromise the sense of belonging, work engagement and the quality of professional relationships, the ability to attract and retain the Institute.

References

CIPD. 2008. *Smart working: how smart is UK plc?* London: Chartered Institute of Personnel and Development. Available at: http://www.cipd.co.uk/.

ILO. 2020. *Defining and measuring remote work, telework, work at home and home-based work*, Geneva. Available at: https://www.ilo.org/.

LI *et al*., 2024. Innovation and challenges of artificial intelligence technology in personalized healthcare. *Scientific Reports (Nature)*, Article number: 18994.

MILEA A. 2024. Digitalization of Work and Its Impact on Worker Safety and Health. In MOLDOVAN L. and GLIGOR A. (Eds) *The 17th International Conference Interdisciplinarity in Engineering. Inter-ENG 2023*. Lecture Notes in Networks and Systems, Vol. 928, Springer, pp. 15-32.

PARESCHI L. 2021. Is It Smart Working? An Analysis of the Public Discourse about Teleworking in Italy during the Covid-19 Pandemic, *PuntOorg International Journal*, Vol. 6, No. 2, pp. 133–157.

Chiara LIMITI, Istat, chiara.limiti@istat.it
Giuseppina SOLA, Istat, giuseppina.sola@istat.it

[7] Between 8 February and 15 March 2021, the CUG conducted a survey among Istat employees to collect information on well-being at work and identify possible areas of hardship. 1,315 employees participated in the survey, with a response rate of 65.9%.

METHODOLOGICAL EVALUATION OF RESPONDENT DRIVEN SAMPLE IN ISTAT LGB EXPERIMENTAL SURVEY

Eugenia De Rosa, Francesca Inglese, David Trambusti

Abstract. The Respondent Driven Sampling (RDS) technique was first applied by Istat (the Italian National Institute of Statistics) in the 2022 “Survey on Labour Discrimination against LGB (lesbians, gay and bisexuals) people not in Civil Union”. The RDS is a valuable approach for studying populations that are difficult to reach, such as LGB people, thanks to its robust theoretical basis. However, the validity of the samples it produces depends on strict assumptions about network structure, the recruitment process, and the sample/population. Furthermore, its implementation is particularly sensitive to operational constraints, including privacy concerns. This work provides a methodological evaluation of the RDS sample obtained in the aforementioned survey. The aim is to identify critical issues related to design choices, implementation limitations and other factors, such as network characteristics and recruitment dynamics. The analyses focused on sample convergence and dependence on seeds, along with potential sources of recruitment bias, including network bottlenecks and homophilic behaviour among participants. The results highlighted several factors that undermined the inferential validity and representativeness of the sample. However, the Italian experience demonstrates the RDS’s ability to engage with populations that are usually under-represented in probability-based surveys. It also contributes to the wider debate within official statistics on the use and enhancement of non-probability sampling methods, and on combining these with probability-based techniques.

1. Introduction

In 2022, Istat implemented for the first time the Respondent-Driven Sampling (RDS) technique in its web-based version (WebRDS) (Wejnert and Heckathorn, 2008) within the framework of the *Survey on Labour Discrimination against LGB people not in Civil Union* (De Rosa and Inglese, 2018).

RDS represents an advanced variant of snowball sampling and, under specific theoretical conditions, offers significant inferential advantages, as it can asymptotically approximate a simple random sample. It is a network-based sampling, that integrates peer-to-peer recruitment with a Markov process model: the sampling begins with the identification of initial participants (*seeds)* from within the target population. Seeds are tasked with recruiting individuals from their own social networks and belonging to the same population, who subsequently recruit further participants, generating chains of recruitment (Heckathorn, 1997; 2002).

RDS is particularly valuable for studying hard-to-reach populations, especially when conventional sampling methods risk excluding marginalised groups or are ineffective. Nevertheless, the validity of RDS relies heavily on strict assumptions regarding network structure, recruitment processes, and the sample/population. Specifically, the network must be connected and free of bottlenecks, characterised by reciprocal ties. Recruitment must be random among peers, with a fixed and limited number of recruits. For statistical inferences to be valid, the sample must reach equilibrium, and the size of the network must be accurately reported and remain stable. In practice, these conditions are often only partially satisfied.

The recruitment process is susceptible to bias from significant variations in personal network size, compounded by structural bottlenecks and behavioural homophily. Furthermore, researchers' limited control over sample composition can result in a final sample that does not sufficiently capture the heterogeneity of the target population. Consequently, a rigorous evaluation of the realized RDS samples is essential to assess data quality and ensure reliable inference. Measurement errors in the self-reporting of personal network size (degree) constitute a critical issue, as they can compromise the validity of the weighting process.

This article analyses the RDS data from the Istat survey of LGB people. Section 2 outlines the context of the survey and explains how RDS was implemented. Section 3 delineates the methodological framework for evaluating sample quality and presents a selection of key findings. The final section offers concluding remarks and proposes areas for potential intervention to improve the design and implementation of RDS in future studies.

2. The application of WebRDS in the Survey on Labour Discrimination against LGB people (not in Civil Union)

Between 2018 and 2023, Istat collaborated with UNAR (National Anti-Discrimination Office) to address the lack of statistical information on LGBT+ populations. The joint project, "Labour Discrimination against LGBT+ People and Diversity Policies in Enterprises", included three targeted surveys. The first two surveys were designed to reach two complementary target groups (mainly LGB people in Civil Union and LGB people not in Civil Union) combining standard and non-standard sampling techniques (De Rosa, 2024; De Rosa and Inglese, 2024). The third survey addressed trans and non-binary people.

The WebRDS approach was cautiously applied to the survey addressed to LGB people not in Civil Union (De Rosa and Inglese, 2018). Initial recruitment via RDS was insufficient to reach the desired sample size, so after several weeks and with a limited number of established recruitment chains, a convenience sampling strategy was introduced to complete the data collection.

The choice of the RDS technique was preceded by a formative study conducted to assess whether the target population was sufficiently networked. This involved a review of academic and grey literature, analysis of existing data, and qualitative research through interviews with stakeholders, key informants, and experts in LGBT+ issues and discrimination. Consultations with LGBT+ associations further explored levels of community belonging, inter-association networking, and informal networks relevant for recruitment feasibility.

The WebRDS design, detailed also in the Data Protection Impact Assessment (Art. 35, GDPR), included: partnership with LGBT+ associations, seeds selection by the association, anonymous web-based self-administration questionnaire on discrimination, and peer recruitment (De Rosa *et al*., 2020). Around 50 LGBT+ associations supported the survey, signing data protection agreements with Istat and in charge of seed identification. Each association selected up to 10 individuals belonging to the population of interest (LGB) seeds based on socio-demographic grid (that included sex, age, region and sexual orientation). Additional selection criteria requested these individuals had strong social connectivity and high motivation to support the study.

For data collection, each association was assigned a unique survey link, enabling monitoring of referral chains without disclosing individual identities. Respondents entered the survey via an "accession module," which provided project information and determined eligibility (e.g., aged 18+, living in Italy, not in Civil Union, and personally knowing the recruiter). Eligible individuals submitted an email address to receive the survey link. Eligibility was confirmed in the initial section of the main questionnaire through a question on current sexual orientation; those who identified as "Other" or selected "Prefer not to say" were screened out. To preserve confidentiality, data from the accession module and the survey were stored on separate servers, and email addresses were encrypted. Network size questions essential for RDS estimation were included in the main questionnaire, asking how many homosexual or bisexual people the respondent knew, and how many they had contacted in the past month. Upon completing the questionnaire, respondents were invited to recruit up to four LGB people from their personal networks by sharing a system-generated referral link, available directly on the screen and sent via email. Recruitment could occur via email, SMS, or messaging apps. No incentives were offered, neither monetary nor of any other nature. Recruitment chains were tracked using anonymous unique codes to monitor the structure of referral waves.

During the data collection process, which started last week of January, Istat researchers closely monitored recruitment indicators such as chain length, demographic composition, and association participation. Despite these efforts, by late April it became evident that the RDS was not functioning effectively. Consequently, from April 26, 2022, the survey was opened to a convenience sample.

The final sample comprised 1,159 LGB individuals: 730 recruited via RDS and 429 via convenience sampling. The main findings of the study were published in May 2023 (Istat-UNAR, 2023).

3. Quality assessment of RDS sample: analysis and main results

Validating an RDS sample requires verifying its quality against the method's underlying assumptions. Given that recruitment occurs through social networks, a comprehensive evaluation is essential to identify and differentiate all potential sources of error that could increase sampling variance and introduce bias.

Convergence analysis is a key step in this process. This verifies whether the sample has reached saturation, which occurs when the composition of the observed characteristics stabilises and becomes independent of the initial seeds. Failure to achieve converge undermines the validity of the inference and the reliability of the results. However, convergence alone does not guarantee that the sample is representative, as structural or behavioural biases may still occur during recruitment (Wejnert, 2009; McCreesh *et al*., 2012; Yamanis *et al*., 2013). Furthermore, variability in network size can introduce recruitment bias by generating unequal inclusion probabilities. Individuals with larger networks are more likely to be recruited, which can lead to the over-representation of well-connected subgroups. Conversely, individuals with smaller networks - often from marginalised or socially isolated groups - are less likely to be included and may be under-represented, thus limiting the diversity captured in the sample.

Convergence must, therefore, be supported by additional diagnostic indicators, such as recruitment bottleneck measures and homophily indices. These tools help determine whether the sample is deep and diverse enough to reflect the target population. Thus, the validity of the RDS assumptions is crucial for data interpretation: if they are met, the results have inferential value and can be generalised; otherwise, the results remain descriptive and not representative.

3.1. RDS sample: preliminary results and seeds contribution

The RDS sample, derived from the Istat survey, comprises 730 individuals who display certain characteristics. More than 40% of respondents are seeds, but only 34% of these generate chains. These chains tend to be short, averaging just 1.76 waves, with only one reaching nine waves. Around 370 recruits (88% of the total recruits) are concentrated in the first three waves.

This preliminary overview highlights the key issues in the RDS procedure. The final composition is significantly affected by the non-random selection of seeds and their uneven contribution to the recruitment process. Furthermore, the limited expansion capacity of the recruitment chains implies partial coverage of the target population, which may indicate a structural failure in the RDS recruitment process, as there are fewer than the four to five waves typically considered to be adequate.

The differential contribution of each seed to the overall recruitment effort was analysed, taking into account the percentage distribution of seeds per recruitment wave generated. Recruitment capacity varies considerably among seeds: 20.3% generated only one recruitment wave, while just 5% produced chains extending between the third and fifth waves (Table 1).

Table 1 – *Distribution of seeds in the recruitment process by waves generated.*

Wave	0	1	2	3	4	5	9
% of seeds	65.9	20.3	8.4	1.9	1.6	1.6	0.3

Furthermore, to evaluate the effectiveness of the initial seeds, an overall seed productivity index was calculated using the formula:

$$\text{Productivity index } = \frac{n_{RDS}}{s_{RDS}} - 1 \qquad (1)$$

where, n_{RDS} is the final sample of respondents (730) and s_{RDS} is the number of initial selected seeds (311). The seed productivity index is 1.35, indicating low recruitment efficiency. When non-productive seeds - those that did not recruit any participants - are excluded, the index rises to nearly 4. This suggests that recruitment may have been driven by a small number of highly productive seeds with specific, distinctive traits.

Subsequent analysis evaluates the contribution of the initial seeds to recruitment, employing both descriptive statistical test[1] and a multivariate model.

Firstly, the distribution of 106 productive and 205 non-productive seeds were analysed by network size. Due to substantial heterogeneity in the self-reported values, the variable was recoded into quartile-based categories. A chi-square test confirms a statistically significant difference in network size between productive and non-productive seeds (test statistic: 18.2; p-value: 0.0004). The proportion of productive seeds with a network size greater than 20 is nearly twice as high as that of non-productive seeds in the same category. Conversely, the proportion of non-productive seeds is higher for small networks and decreases for larger networks (Table 2).

[1] Asterisks indicate the level of statistical significance (p-value): * for p-value < 0.05, ** for p-value < 0.01, *** for p-value < 0.001. The method used in all contingency tables to obtain this result is the bootstrap simulation.

Table 2 – *Distribution of productive and non-productive seeds by network size class (% by column).*

Network size	Productive seeds	Non-productive seeds
≤ 5	10.4 **	24.4 *
< 5 - ≤ 10	18.9	25.4
< 10 - ≤ 20	29.2	28.8
> 20	41.5 **	21.5 *

Secondly, a logistic regression model was applied to further investigate the determinants of recruitment success, The dependent variable was defined as binary, taking the value 1 if the seeds were productive, and 0 otherwise. The model included the following auxiliary variables: employment status (employed, unemployed, inactive); geographical area (five municipality categories); the logarithm of network size; and participation (at the time of the survey) in LGBT+ associations or groups (yes or no).

The model achieved a classification accuracy of 69.5% and an area under the ROC curve (AUC) of 69%, indicating an acceptable level of discriminative power. Only three variables were found to be statistically significant predictors of seed productivity ($p < 0.05$): employment status, log-transformed network size, and participation in LGBT+ associations or groups (Table 3).

Table 3 – *Odds ratio of the logit model.*

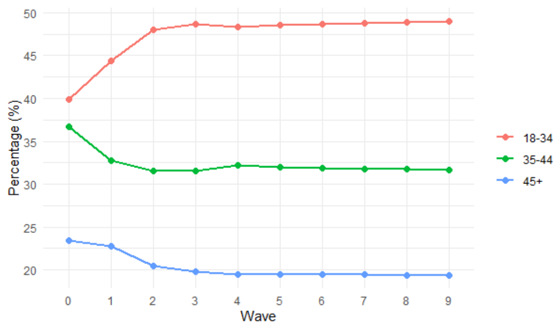
Variable	Odds ratio	p-value	significance
(Intercept)	0.11	0	***
Employment status:			
employed (reference)	–	–	
unemployed	0.2	0.032	*
inactive	1.38	0.445	
Municipality:			
≤ 5000 (reference)	–	–	
5,001-20,000	1.7	0.319	
20,001-50,000	0.5	0.259	
50,001-150,000	0.78	0.632	
> 150,000	1.37	0.513	
Log(network size)	1.33	0.025	*
Being part of associations or other groups	2.31	0.014	*

As shown in table 3, unemployment significantly reduces the likelihood of successful recruitment, making unemployed individuals around five times less likely to be recruited. Conversely, active engagement in LGBT+ associations or groups more than doubles the likelihood of successful recruitment. Having a larger social network also has a positive, albeit more modest, influence. No statistically significant association was observed between municipality category and seeds productivity.

3.2. *RDS validation: sample convergence and recruitment analyses*

To examine sampling dynamics, a convergence analysis of the cumulative frequencies of participants by age class across successive waves was carried out. An equilibrium plot (Figure 1) was used for this purpose to assess the stability of the sample over time. Considering a variation of ±2% between waves, it appears that all categories reach equilibrium by the second or third wave. However, this apparent stability does not guarantee true convergence of the RDS sample, since most respondents are concentrated in the first three waves (88% of recruits), and the proportions flatten quickly. Therefore, this result may be interpreted as reflecting the initial composition of the sample rather than as evidence of genuine convergence in the RDS process with respect to the age variable.

Figure 1 – *Equilibrium plot by age class.*



This interpretation is reinforced by comparing the seeds and recruits. The chi-square test reveals significant differences in age composition between the two groups (test statistic: 19.14; p-value: 0.0003). Younger respondents are over-represented among recruits (55.8%) compared to the initial group of seeds (39.9%). Conversely, all older age groups are under-represented among recruits. This decline is particularly pronounced for the 35–44 and 45–54 age groups (Table 4).

Table 4 – *Distribution of seeds and recruits by age group (% by column).*

Age class	Seeds	Recruits
18-34	39.9 **	55.8 **
35-44	36.7 *	27.9
45-54	18 *	11.5 *
55+	5.5	4.8

The recruitment process was analysed to identify potential sources of bias related to the high variability in network size, bottlenecks in network structure, and participants' homophilous behaviour.

Network size was assessed to determine whether individuals with larger or smaller personal networks were disproportionately represented within the sample. The analysis revealed systematic imbalances in the composition of the sample with respect to this variable, which appear to be closely associated with two structural dimensions: respondents' age and municipality of residence.

A chi-square test confirms that network size varied significantly with age (test statistic: 18.23; p-value: 0.033). Younger people predominate in small networks, while older people tend to be more visible in larger ones (Table 5).

Table 5 – *Distribution of RDS sample by network size and age group (% by row).*

Network size	Age class			
	18-34	35-44	45-54	55+
≤ 5	52.9	31.1	13.6	2.4 *
< 5 - ≤ 10	54	25.7	13.9	6.4
< 10 - ≤ 20	48.6	34.1	14	3.4
> 20	38.6 *	36.7	15.8	8.9 *

RDS preferentially captured certain subgroups. In essence, an individual's likelihood of being included in the sample was not random but was significantly influenced by their network size, which in turn correlates strongly with their age and where they live.

Larger municipalities have more extensive social networks, whereas very small networks prevail in municipalities with fewer than 5,000 inhabitants (test statistic: 40.26; p-value: 0.000065). This result highlights the difficulty of recruiting in small municipalities, where recruitment chains are quickly exhausted, and the risk of oversampling "super-spreaders" in large cities. Respondents living in larger municipalities tend to report wider and more heterogeneous personal networks, while small municipalities are characterized by a predominance of very limited networks (Table 6).

Table 6 – *Distribution of RDS sample by network size and municipality (% by row).*

Network size	Municipality (for number of inhabitants)				
	≤ 5000	5,001-20,000	20,001-50,000	50,001-150,000	> 150,000
≤ 5	13.1 *	21.4	17 **	22.3	26.2 ***
< 5 - ≤ 10	9.6	13.4	9.6	27.8 *	39.6
< 10 - ≤ 20	7.3	16.2	10.6	20.1	45.8
> 20	7	15.8	6.3 *	18.4	52.5 **

Structural barriers in the recruitment network were examined using a Bottleneck Indicator (BI) calculated for three age class, based on their representation in the sample generated from productive seeds only (Table 7). For each seed, the BI takes into account three components: (i) the size of the recruitment chain, (ii) the proportion of the target group within the chain, and (iii) the absolute deviation of this proportion from the group's share in the overall sample| pij - pj |. The bottleneck indicator is then computed as the weighted mean of these absolute deviations, following the formula below:

$$BI = \sum_{i=1}^{S} w_i \sum_{j=1}^{H} \left| p_{ij} - p_j \right|, \tag{1}$$

where: S = number of chains (seed), i=1,…,S; H = number of groups, j=1,…,H; p_{ij}= proportion of group j in chain i; p_j= proportion of group j in the total sample; w_i= weight of the chain i.

Table 7 - *Recruitment bottlenecks - proportion of age groups.*

Age group	Proportion in RDS sample	Bottleneck indicator (weighted with number of recruits)
18-34	52.9 %	30.9 %
35-44	29.9 %	24.1 %
45+	17.2 %	20.1 %

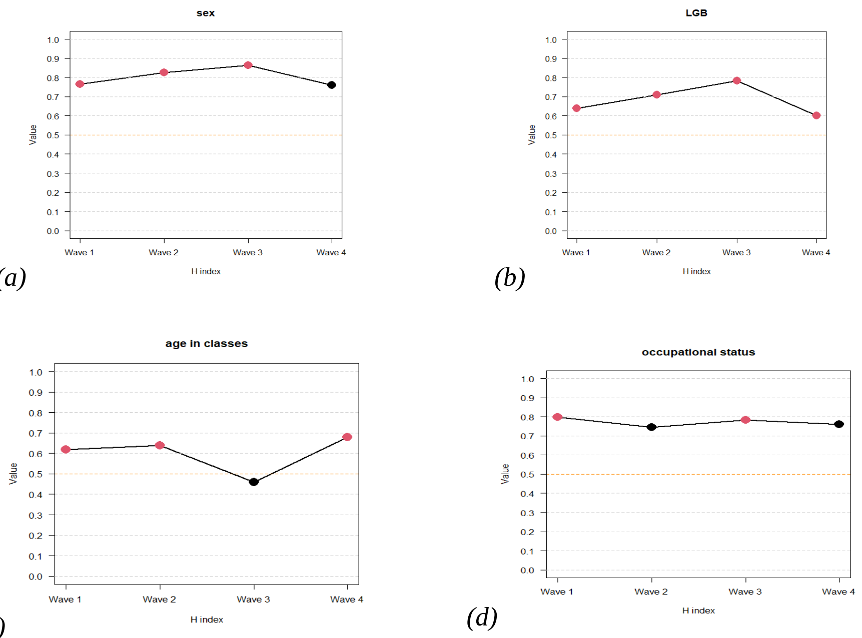
The results in table 7 indicate the presence of structural barriers in the network. The 18–34 age group, despite being the largest in the sample (52.9%), shows a relatively high BI (30.9%), suggesting that younger respondents are concentrated in specific chains rather than evenly distributed. The 35–44 age group has a moderate BI (24.1%), indicating a more uniform distribution along the chains. The 45+ group, although representing only 17.2% of the sample, exhibits a BI of 20.1%, revealing that older participants tend to cluster in certain chains, pointing to structural constraints in reaching this group. Overall, these findings highlight that structural barriers in the network affect both the largest and smallest age groups, while the intermediate group experiences a more balanced distribution. Such barriers may hinder the even diffusion of certain groups and restrict population mixing. This affects the representativeness of the sample and introduces substantial selection bias.

The homophilic tendencies of the RDS participants were evaluated by analysing recruiter-recruit similarity to understand the extent to which participants' recruited individuals similar to themselves. The homophily index (H) was calculated for different distributions - sex, sexual orientation, age class, employment status, educational level, income class, family size and municipality - on the transition matrix M of dimension k×k:

$$H = \frac{\sum_{i=1}^{k} m_{ii}}{\sum_{i=1}^{k}\sum_{j=1}^{k} m_{ij}} \quad (2)$$

In the formula, *i* and *j* denote row and column respectively; the numerator is the sum of the frequencies on the main diagonal of the transition matrix and the denominator is the sum of the total frequency of the matrix. The analysis revealed a significant homophily effect in the recruitment patterns of the following variables: sex, sexual orientation, age, employment status and level of education. This suggests the formation of homogeneous social clusters, which limits the diversity of the sample and likely results in the over-representation of certain subgroups. Significant results are shown in Figure 2, where red dots indicate significance ($p < 5\%$) in the chi-square test or Fisher's exact test if not performable due to a lack of data, while black dots indicate non-significance. In wave 3, approximately 90% of the sample recruited individuals of the same sex, over 80% recruited individuals of the same sexual orientation, and over 70% recruited individuals of the same occupational status. A significant result emerges for the age variable in waves 1 – 2.

Figures 2 - *H index for sex and sexual orientation (LGB group) (Figure (a) and (b)); H index for age in 3 classes and occupational status (employed vs not) (Figure (c) and (d)).*



(a) *(b)* *(c)* *(d)*

4. Final remarks

The RDS sample from the Istat-UNAR survey revealed several critical limitations. The non-random selection of seeds resulted in significant imbalances, and the sample did not reach stationarity, with the initial selection of seeds continuing to influence its final composition. Recruitment chains rarely reached sufficient depth, which limited coverage and left substantial portions of the target population under-represented. Structural bottlenecks and homophilic behaviours restricted recruitment, resulting in uneven representation. Recruitment process integrity was compromised by the fact that participants often failed to recruit the maximum number of peers. The limited effectiveness of seeds selection was likely due to several factors: many nominated seeds did not participate, suggesting self-selection and suboptimal choices by associations. Privacy constraints prevented the research team from training or monitoring seeds, and procedural limitations - such as recruitment via email only and the absence of incentives - further weakened the development of recruitment chains. In future studies, seed selection should be considered a fundamental design element of RDS and supported by robust communication strategies, such as engaging LGBTQ+ influencers. The formative study is equally important and should not be considered marginal; prior data and formative findings are essential for understanding the target population and designing network models that promote diverse recruitment. Finally, the high variability found in the self-reported network size variable highlights the potential for measurement error, and underlines the need to test and refine these questions to ensure consistent and accurate interpretation by respondents.

References

DE ROSA E., INGLESE F. 2018. Diseguaglianze e discriminazioni nei confronti delle persone LGBT: quale contributo della statistica ufficiale?, *Rivista italiana di economia demografia e statistica RIEDS*, Vol. LXXII, No. 4, pp.77-88.

DE ROSA E., DE VITIIS C., INGLESE F., VITALINI A. 2020. Il web-respondent driven sampling per lo studio della popolazione LGBT+, *Rivista italiana di economia demografia e statistica RIEDS*, Vol. LXXIV-1 January-March 2020, pp. 66-73.

DE ROSA, E. (Eds) 2024. *Discriminazioni lavorative nei confronti delle persone LGBT+ e le diversity policy" Istat: Roma*. https://www.istat.it/wp-content/uploads/2024/07/Discriminazioni-lavorative-persone-LGBT-diversity-policy-Ebook.pdf

DE ROSA E., INGLESE F. 2024. Sampling and civil society engagement in surveying hard to reach populations, *Rivista italiana di economia demografia e statistica RIEDS*, Vol. LXXVIII-3 July-September 2024, pp. 135-147.

ISTAT-UNAR. 2023. *L'Indagine Istat-UNAR sulle discriminazioni lavorative nei confronti delle persone LGBT+ (non in unione civile o già in unione)*. Anno 2022.

HECKATHORN D. D. 1997. Respondent-driven sampling: A new approach to the study of hidden populations, *Social Problems*, Vol. 44, No. 2, pp. 174-199.

HECKATHORN D. D. 2002. Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations, *Social Problems*, Vol. 49, No. 1, pp. 11-34.

MCCREESH N., FROST S. D. W., SEELEY J., KATONGOLE J., TARSH M. N., NDUNGUSE R., WHITE R. G. 2012. Evaluation of respondent-driven sampling, *Epidemiology*, Vol. 23, No. 1, pp. 138-147.

WEJNERT C., HECKATHORN D. D. 2008. Web-based network sampling: Efficiency and efficacy of respondent-driven sampling for online research, *Sociological Methods & Research*, Vol. 37, No. 1, pp. 105-134.

WEJNERT C. 2009. An Empirical Test of Respondent-Driven Sampling: Point Estimates, Variance, Degree Measures, and Out-of-Equilibrium Data. *Sociological Methodology*, Vol. 39: pp. 73-116.

YAMANIS T. J., MERLI M. G., NEELY W. W., TIA, F., MOODY J., TU X., GAO E. 2013. An empirical analysis of the impact of recruitment patterns on RDS estimates among a socially ordered population of female sex workers in China. *Sociological Methods & Research*, Vol. 42, No. 3, pp. 392-425.

Eugenia DE ROSA, Istat, eugenia.derosa@istat.it
Francesca INGLESE, Istat, fringles@istat.it
David TRAMBUSTI, Istat, david.trambusti@istat.it

AN ANALYSIS OF THE RESPONSE BEHAVIOUR OF FARMS IN THE FARM STRUCTURE SURVEY[1]

Ornella Mobilia, Annalisa Pallotti, Francesca Rossetti

Abstract. The 2023 edition of the Farm Structure Survey is a sample survey that includes approximately 100,000 agricultural and livestock units registered in the Farm Register. This survey was conducted in accordance with Regulation (EC) No 1091/2018 of the European Parliament and of the Council, following an organizational model based on collaboration among several private entities.
Thanks to an agreement with Istat, the Agricultural Assistance Centras (CAA) played a crucial role in the survey activities, acting as intermediaries in the data collection for various agricultural surveys. Operators affiliated with CAAs have a privileged relationship with farms due to their ongoing activities and possess specialized training in agricultural matters.
Data collection was carried out using a combination of two survey techniques to optimize costs and minimize the statistical burden on respondents. Farms were given the option to complete the questionnaire online via CAWI (Computer Assisted Web Interviewing). Non-responding units were subsequently contacted by an operator who administered the questionnaire using CAPI (Computer Assisted Personal Interviewing).
The objective of this work is to analyse the possible links between the characteristics of farms and their response behaviour, also in relation to the possible effects of actions taken to encourage participation.

1. Introduction

In the context of the 2023 edition of the Farm Structure Survey (Eurostat, 2022), which covered around 100,000 agricultural and livestock holdings from the National Farm Register, this study examines the effectiveness of collaborative data collection strategies in the agricultural sector. Conducted under an European Regulation (EC), the survey employed an innovative organisational model involving multiple private actors. Central to this model were the Agricultural Assistance Centres (CAAs) (Poti B., 2011), which acted as trusted intermediaries in the collection of farm data through a formal agreement with Italian National Institute of Statistics (Istat). Thanks to their specialised training and long-standing relationships with local farms, CAA operators

[1] The article was only possible thanks to the joint work of the authors. In particular, Annalisa Pallotti wrote sections 2, 4 and 6, Francesca Rossetti wrote sections 1 and 5 Ornella Mobilia wrote section 3.

played a pivotal role in ensuring effective engagement with respondents. This article explores the relationship between farm characteristics and survey response behaviour, paying particular attention to the impact of targeted actions designed to increase participation. We propose a statistical analysis based on the logistic model to assess the propensity of farms to cooperate in the survey in relation to the survey characteristics of the targeted actions.

The document is structured as follows: Sections 2 and 3 outline the survey objectives and design, with a particular focus on the survey network and data collection methods; section 4 presents a detailed analysis of response rate trends over the course of the data collection period; section 5 describes the results of the application of the logistical model; finally, section 6 provides concluding remarks.

2. Farm Structure Survey

Farm Structure Survey represents one of the main structural statistical sources on the Italian agricultural system. It is a sample survey carried out every three years, conducted in the intermediate years between the decennial general agricultural censuses. This periodicity allows for dynamic monitoring of the evolution of farms between two full censuses, ensuring continuity of statistical information.

The 2023 edition involved 109,229 agricultural and livestock holdings, selected from the Farm Register, and was conducted in compliance with Regulation (EU) No. 2018/1091 and its related Implementing Regulation of the European Commission (No 2024/2914), as part of the system of harmonized surveys established at the EU level. The data collected are crucial for supporting impact analyses of the Common Agricultural Policy (CAP), environmental sustainability, and rural development, providing analytical insights in the inter-census period.

The main objectives of the survey are to analyze the structural evolution of farms, through the collection of information on their economic and physical size, production type, labor force, use of natural resources, and agro-environmental practices. The variables collected were designed to provide a multidimensional reading of the agricultural holding, covering both structural aspects (land area, livestock, labor, machinery) and farming techniques (crop types, fertilization methods, livestock systems).

The methodological design included a stratified and regionally representative sampling plan, ensuring sufficient coverage for disaggregated data analysis. Particular attention was paid to improving the quality of the questionnaire and simplifying the electronic data collection interface. The questions were adapted linguistically and functionally to make them easier for respondents to understand and reduce non-response rates.

In this context, a well-structured and target-specific data collection strategy proved essential to ensure the timeliness of information, alignment with current sector dynamics, and the statistical robustness required for analytical and policy-making purposes.

3. Network and Data Collection Techniques

The organizational structure of the survey was based on an integrated collaboration model between ISTAT and the Agricultural Assistance Centers, formalized through the signing of an operational agreement that assigned the CAA the role of official data collection bodies on behalf of the Institute. The CAA, entities recognized by the Ministry of Agriculture and accredited by the National Agricultural Information System (SIAN), act as certified intermediaries between the State and farms for the management of the farm register and access to key instruments of the Common Agricultural Policy (MASAF, 2023).

Their role within the data collection network is not only logistical but also methodological: thanks to their familiarity with local farming practices, in-depth knowledge of regulatory definitions, and technical expertise, the CAA contribute to coherent, standardized, and highly reliable data collection.

For the first time, during the 2020 Agricultural Census, the network of Agricultural Assistance Centers (CAAs) was employed as the primary data collection channel, producing positive outcomes in terms of territorial coverage and data quality (Istat, 2023). Owing to their longstanding relationships with farms and specialized expertise in agricultural matters, CAA operators served as qualified intermediaries, enabling comprehensive, detailed, and accurate data gathering across the national territory. In the aftermath of the COVID-19 emergency, in fact, reliance on administrative records had resulted in an over coverage of the farm register, thus underscoring the necessity of field verification to determine the actual number of active agricultural holdings. In this context, the CAA network—characterized by its technical competence and in-depth sectoral knowledge—proved instrumental in achieving the cognitive goals of the Census, further consolidating its collaboration with Istat.

In the 2023 edition of the survey, their involvement enabled the combination of operational efficiency and territorial coverage, ensuring high data quality nationwide. The network deployed included 6 lead CAA organizations, 1,819 local offices, and 2,290 trained operators, who maintained trusted relationships with local farms. The 6 lead CAAs each coordinated a group of affiliated CAAs, bringing the total number of national CAAs involved in the network to 28.

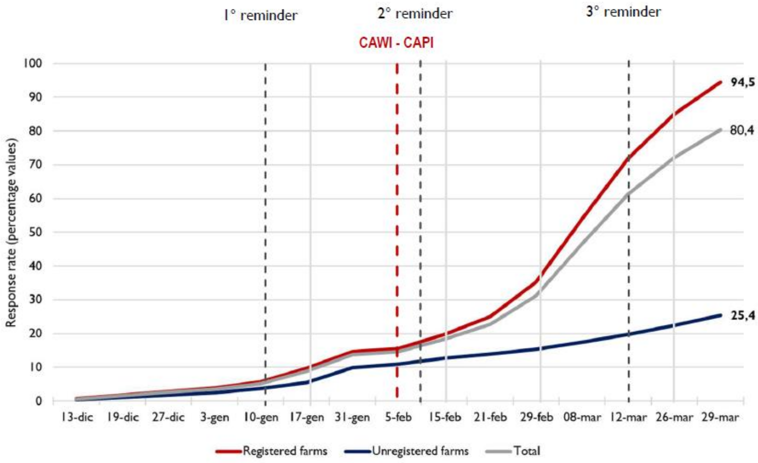
The sample of approximately 109,229 holdings was divided into two groups: farms with an active farm register, hereinafter “registered” (86,995 units registered with a participating CAA) and farms without an active register, hereinafter “unregistered”, (22,304 units either not associated with or not recorded in a participating CAA). The survey, conducted between December 2023 and March 2024, utilized two data collection modes: Computer-Assisted Web Interviewing (CAWI), for self-completion of the electronic questionnaire, and Computer Assisted Personal interviewing (CAPI) conducted by CAA operators. While the first group had access to both modes sequentially, the second group could only use the CAWI mode. The initial communication inviting units to participate in the survey also includes a description of the technique(s) through which they can contribute. This enables registered units to decide whether to wait for the CAA support available during the second part of the survey period.

This multimodal approach made it possible to optimize participation, reduce respondent burden, and ensure the quality of the data collected.

4. Trends in Survey Response Rates During Fieldwork

During the field data collection period, farm response rates were continuously monitored. Figure 1 illustrates the response rate trends in relation to actions taken to encourage survey participation, distinguishing between the two groups defined by the survey design: registered and non-registered units.

Figure 1 – *Response rates by reminder and technical changes.*



Source: our elaboration on data from Farm Structure Survey 2023.

The response curve showed a steady upward trend from the beginning of data collection, with a marked acceleration following the first scheduled reminder sent on 10 January. By 17 January, the overall response rate had increased by 5.3 percentage points, and by 5.8 points among registered farms. The reminder stated that registered farms could respond via CAWI mode until 21 January[2]. Following the reminder, a notable increase was observed across all series, particularly among registered farms.

From 5 February, with the launch of the CAPI phase, the implementation of the mixed CAWI-CAPI technique for registered farms marked a qualitative leap in data collection. As of 15 February, the response rate stood at 19.8% for registered units and 12.8% for non-registered ones.

The second reminder, sent on 12 February, positively influenced survey performance—especially for registered farms: by 29 February, the observed response rates had reached 35.2% for registered farms and 15.3% for non-registered farms.

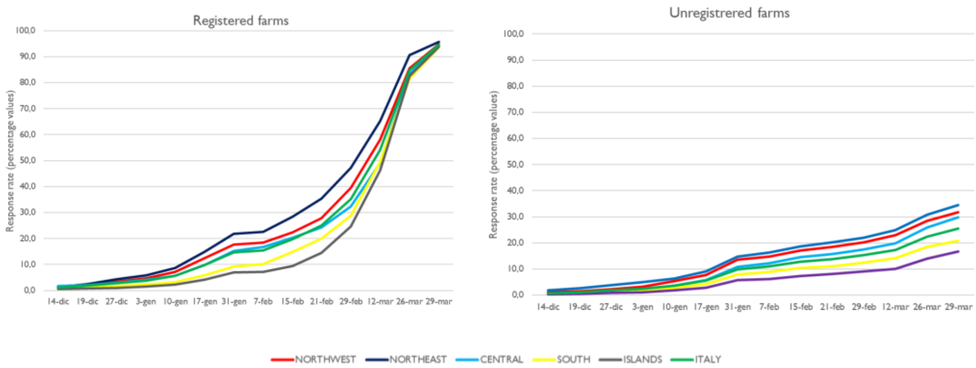
The final reminder, issued on 12 March, was instrumental in maximizing coverage, particularly among registered farms. The ability of CAAs to rapidly mobilize local structures contributed to the final increase, in line with the closure target set for the end of March.

By 29 March, the end of the data collection period, the overall response rate had reached 80.4%, with a peak of 94.5% among registered farms and a final response rate of 25.4% among non-registered farms.

The trend in response rates over time, revealing consistent patterns across different geographic areas, but with higher rates in the North than in the South. This pattern is evident in both the registered and unregistered groups, which, as previously noted, follow distinct trajectories. A comparison of the two sample segments shows that the introduction of the CAPI phase for registered farms, during the second part of data collection, marks a turning point in the trend lines, as the gap in response rates between geographic areas gradually narrows over time. By the end of the data collection period, response rates among registered participants had converged, with differences reduced to within 1.2 percentage points. The intervention of CAA operators helped offset the lower survey participation rates among firms in the South and the Islands, contributing to improved data quality in these areas. In contrast, among unregistered participants, disparities in response propensity persisted throughout the period, with a gap of 15.1 percentage points observed between the North-East and the final phase (Figure 2).

[2] The start of the CAPI phase, initially scheduled for 22 January, was postponed to 5 February.

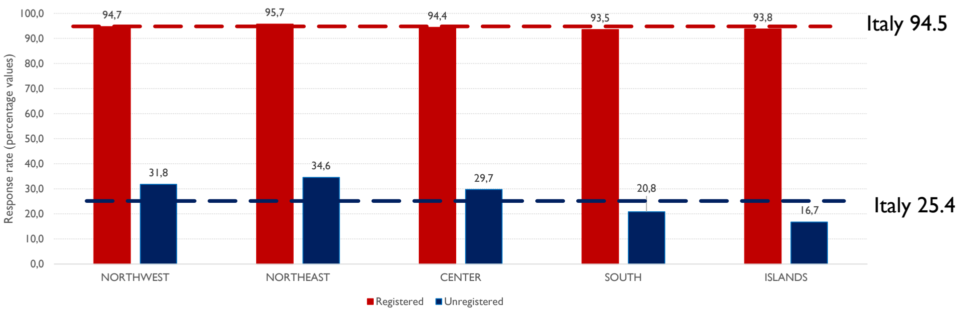
Figure 2 – *Response rates by subsample and geographic area*.



Source: our elaboration on data from Farm Structure Survey 2023

Farms with a company file achieve the highest response rates in all geographical areas, with clear peaks in the north-east of the country. These areas therefore appear to be characterised by a greater formalisation of agricultural businesses. Companies without a registered file, on the other hand, show lower response rates everywhere, with the greatest critical issues being highlighted in the Islands and the South. This data could reflect structural differences in farms, with fewer formalised companies in these areas (Figure 3).

Figure 3 – *Territorial distribution of response rate.*



Source: our elaboration on data from Farm Structure Survey 2023.

5. Exploring Predictors of Participation through Logistic Modelling

Analysis of the response rates revealed varying levels of participation in the survey among registered farms and the rest of the sample. This characteristic also determined assignment to one of two survey designs, using either dual CAWI-CAPI or CAWI-only techniques.

To further analyse the determinants of farm participation in the survey, a logistic model was applied (Long et al., 2014; Agresti, 2013). In the model the dependent variable is having or not having collaborated in the survey. The explanatory variables used in the model are the survey technique and the variables that provide a measure of the farms, such as labour units per year (ULA), standard farm production value in euros (SO), utilised agricultural area in hectares (UAA) and Standard livestock units (UBA). The model also considers the geographical area and membership of one of the CAA involved in the convention. The values entered the model by the variables are shown in the Table 1.

Table 1 – *Explanatory variables and mode encoding applied*[3].

Explanatory variables	Code
Survey techniques	ind_tecnica=0 for CAWI; ind_tecnica =1 for CAPI
Labour Units per Year – ULA	ULA_TOT_ind1= until 1; ULA_TOT_ind2=from 1 to 9; ULA_TOT_ind3= over 10
Value of the company's standard output - SO in euros	SO_ITA_ind1= Up to 2000; SO_ITA_ind2=From 2.000 to 15.000; SO_ITA_ind3=from 15.000 to 100.000; SO_ITA_ind4=from 100.000 to 1.000.000; SO_ITA_ind5=over 1.000.000
Utilised Agricultural Area – SAU	SAU_ind1= Until 0.99; SAU_ind2=From 1 to 1.99; SAU_ind3=From 2 to 4.99; SAU_ind4=From 5 to 19.99; SAU_ind5=over 20
Standard livestock units – UBA	UBA_ind1= Until 1.99; UBA_ind2=From 2 to 4.99; UBA_ind3=From 5 to 9.99; UBA_ind4=From 10 to 19.99; UBA_ind5=From 20 to 49.99; UBA_ind6=over 50
Geographical area	ripa_NO=North-West; ripa_NE=North-East; ripa_CE=Centre; ripa_SU=South; ripa_IS=Islands for Geographical area
CAA of membership	caa_1=1; caa_2=2; caa_3=3; caa_4=4; caa_5=5; caa_6=6

Source: our elaboration on data from Farm Structure Survey 2023

[3] The CAAs that contributed to the data collection process, listed in alphabetical order, are: CAA CAF Agri, CAA CIA, CAA Coldiretti, CAA Confagricoltura, CAA Degli Agricoltori, CAA Delle Venezie. To safeguard the confidentiality of individual performance—which is not pertinent to the objectives of this paper—we have opted to refer to them using random numerical codes.

Table 2 reports the Chi-square statistics along with the corresponding p-values, highlighting variables that are not statistically significant in the model. It also includes the odds ratios, which represent the logistic regression coefficients for the explanatory variables[4].

Table 2 – *Explanatory variables, Chi-square statistics with p-value associated.*

	Parameter	odds ratio	Pr > Chi-sq
	Intercept	105,0	<.0001
Survey techniques	ind_tecnica	1,2	<.0001
Work Units per Year – ULA	ULA_TOT_ind2	1,8	<.0001
	ULA_TOT_ind3	1,0	<.0001
Value of the company's standard output – SO	SO_ITA_ind2	1,2	0,3441
	SO_ITA_ind3	1,3	0,0092
	SO_ITA_ind4	2,3	<.0001
	SO_ITA_ind5	1,1	<.0001
Utilised Agricultural Area – SAU	SAU_ind2	1,0	0,2619
	SAU_ind3	1,2	0,8967
	SAU_ind4	1,4	<.0001
	SAU_ind5	0,8	<.0001
Standard livestock units – UBA	UBA_ind2	0,6	0,0021
	UBA_ind3	0,6	<.0001
	UBA_ind4	0,6	<.0001
	UBA_ind5	0,7	<.0001
	UBA_ind6	2,3	<.0001
Geographical area	ripa_NO	4,4	<.0001
	ripa_NE	2,2	<.0001
	ripa_CE	1,3	<.0001
	ripa_SU	105,0	<.0001
CAA of membership	caa_1	1,2	<.0001
	caa_2	1,8	<.0001
	caa_3	1,0	<.0001
	caa_4	1,2	<.0001
	caa_6	1,3	<.0001

Source: our elaboration on data from Farm Structure Survey 2023.

According to the observation of the p-values, it can be inferred that nearly all the independent variables considered in the analysis are statistically significant. Based on the multicollinearity analysis - which excluded the presence of critical collinearity among predictors - and the negligible difference observed in goodness-of-fit indicators when excluding non-significant variables, the model was retained in its

[4] The baseline values of each predictor are ULA_TOT_ind1 for ULA; SO_ITA_ind1 for SO; SAU_ind1 for SAU; UBA_ind1for UBA; Islands for Geographical area; caa_5 for CAA.

entirety. This choice ensures thematic coherence and structural consistency throughout the overall reading of the article. The overall fit of the logistic regression model was deemed acceptable. The model's effectiveness was confirmed by multiple concordance measures, each showing strong results (Hosmer et al., 2013). Detailed information is provided in the Appendix.

The model confirms that the assigned data collection method is the primary driver of survey participation: this variable is determined by the survey design and is closely linked to a farm's registration with a CAA involved in the survey network. In fact, the observed odds for CAA membership generally indicate a higher likelihood of participation among registered units compared to unregistered ones, with varying propensities depending on the specific CAA grouping to which they belong. The odds ratios associated with farm characteristics are all greater than 1, indicating that the propensity to participate increases with the size of the farm. The only exception is found in relation to the number of standard livestock units, for which the odds ratios are less than 1: here, the likelihood of participation decreases as the size class increases. Regarding geographic area, units located outside the island regions show a significantly higher propensity to participate—up to more than four times greater in the eastern regions.

6. Conclusion

The success of the survey strategy relies heavily on the unique strengths of the Agricultural Assistance Centres. Their direct knowledge of farms and widespread territorial presence play a critical role in fostering participation among statistical units. The professional network they represent—often composed of agronomists with deep expertise in the subject matter—enhances the quality of microdata collected from registered holdings. However, the inclusion of unregistered farms, often less inclined to engage due to their weaker links with CAAs, required a deliberate oversampling approach to mitigate risks of under-coverage and self-selection bias. The adoption of the CAWI technique further ensured broad accessibility, enabling participation from all farms regardless of registration status. Notably, the communication of the assigned survey technique from the outset can influence response behavior, particularly for CAPI among CAA-affiliated units. Ultimately, the use of a technically equipped and territorially embedded network remains essential for high-quality, inclusive data collection across the agricultural sector.

Appendix

Assessment and key indicators of the deployed logistic model

Multicollinearity analysis.

The correlation matrix computed for the variables listed in Table 1 shows coefficients ranging from –0.58 to 0.46, indicating the absence of strong linear associations among the variables. Furthermore, the assessment of multicollinearity through the Variance Inflation Factor (VIF) reveals values below the commonly accepted threshold of 5 for the majority of variables. Only a limited number of exceptions were observed, with VIF values falling within the moderate range of 6 to 8 (SO_ITA_ind3 = 6.09; SO_ITA_ind = 7.78; SAU_ind5 = 6.53), which do not raise substantial concerns regarding multicollinearity (see Table A1).

Table A1 – *Variance Inflation Factor values for all variables included in the model*.

Variables	VIF
ULA_TOT_ind2	1,72
ULA_TOT_ind3	1,18
SO_ITA_ind2	3,06
SO_ITA_ind3	6,09
SO_ITA_ind4	7,78
SO_ITA_ind5	2,81
SAU_ind2	1,62
SAU_ind3	2,44
SAU_ind4	4,68
SAU_ind5	6,53
UBA_ind2	1,02
UBA_ind3	1,03
UBA_ind4	1,10
UBA_ind5	1,19
UBA_ind6	1,55
ripa_NO	1,94
ripa_NE	2,12
ripa_CE	1,97
ripa_SU	2,13
caa_103	1,91
caa_124	1,26
caa_107	1,46
caa_129	1,24
caa_105	1,59

Source: our elaboration on data from Farm Structure Survey 2023.

Logistic model diagnostics.

The Akaike Information Criterion (AIC = 39192) and the Bayesian Information Criterion (BIC = 39441) suggest a reasonable balance between model complexity and goodness of fit, with lower values indicating better parsimony. The –2 Log-Likelihood value (–2 Log L = 39140) further supports the adequacy of the model in capturing the observed data structure.

In terms of explanatory power, the model yielded a Coefficient of Determination (R^2) of 0.39, indicating that approximately 39% of the variance in the outcome variable is accounted for by the predictors. Notably, the Adjusted R-squared value of 0.68 suggests a substantial improvement in model fit when adjusting for the number of predictors, highlighting the robustness of the model specification (see Table A2).

Table A2 – *Logistic model diagnostics: comparison between models using significant variables vs. models using all variables, with percentage differences in indicators.*

Fit indicator	Model using significant variables	Model using all variables	Percentage difference
AIC - Akaike Information Criterion	39188	39192	0,01
SC (BIC) - Schwarz Criterion (Bayesian Information Criterion)	39409	39441	0,08
-2 Log L - Negative Two Times the Logarithm of the Likelihood Function	39142	39140	-0,01
R-squared - Coefficient of Determination	0,39	0,39	0,00
Adjusted R-squared Adjusted Coefficient of Determination	0,68	0,68	0,00

Source: our elaboration on data from Farm Structure Survey 2023.

Predictive performance.

The model demonstrates excellent discriminative ability, as indicated by an AUC/c-statistic of 0.959, which is considered outstanding. Concordance statistics—Somers' D of 0.917 and Gamma of 0.918—further confirm that the predicted probabilities are well aligned with the observed outcomes. The Tau-a coefficient, equal to 0.239, is comparatively lower, as expected, since it accounts for the total number of possible pairs (see Table A3).

Table A3 - *Diagnostic indicators assessing the discriminative power and ordinal association of the logistic model.*

Diagnostic indicators	Value
Percentage of concordant pairs	95,8
Percentage of discordant pairs	4,1
Percentage of tied pairs	0,1
Somers' D	0,917
Gamma	0,918
Kendall's Tau - Tau-a	0,239
Area Under the Curve - AUC/c-statistic	0,959

Source: our elaboration on data from Farm Structure Survey 2023.

References

AGRESTI A. 2013. *Categorical Data Analysis.(*3rd) Hoboken, Wiley.

EUROSTAT. 2022. Farm Structure Survey 2020 – *Main Results*. Luxembourg: Publications Office of the European Union

HOSMER D. W., LEMESHOW S., STURDIVANT R. X. 2013, *Applied Logistic Regression* (3rd ed.). Hoboken, Wiley.

ISTAT. 2023. 7th General Census of Agriculture. *Final Results*. Rome: ISTAT.

LONG J.S., FREESE J**.** 2014. *Regression Models for Categorical Dependent Variables Using Stata*. College Station: Stata Press.

MINISTRY OF AGRICULTURE, FOOD SOVEREIGNTY AND FORESTS (MASAF). 2023. *National List of Authorized Agricultural Assistance Centers (CAA)*. Rome: MASAF

POTI B. 2011. The Organization of Services for Agricultural Enterprises: Evolution and Role of Agricultural Assistance Centers (CAA). In: *QA Journal of the Rossi-Doria Association*, No. 1, pp. 103–122. Rome: Donzelli.

Ornella MOBILIA, Istat, ornella.mobilia@istat.it
Annalisa PALLOTTI, Istat, pallotti@istat.it
Francesca ROSSETTI, Istat, frrosset@istat.it

DO THE RESPONSE RATES FOR SHORT-TERM BUSINESS SURVEYS HAVE A SEASONALITY?[1]

Katia Bontempi, Claudio Ceccarelli, Francesca Rossetti

Abstract. As with many statistical outputs produced by the Italian National Statistics Institute (ISTAT), short-term business statistics are subject to specific European regulations. These regulations define the structure of questionnaires and surveys, the scope of observation, the reference population, the sampling methodology, and the precision of the estimates. Within the specific context of enterprises, especially in the Italian production landscape, where small businesses are predominant, regulatory criteria and constraints necessitate the continuous inclusion of large enterprises in survey samples over time.
Through the time-series analysis of business response rates, this paper aims to verify the presence of seasonality in the response behaviour of businesses over time. Should seasonal factors emerge, it would be possible to consistently provide specific interventions modulated according to the phase in which business participation is most lacking. This would allow an optimization of the resources deployed and a reduction of the burden in periods when participation is highest.
The data analysed in this study come from the *The Business Statistical Portal*, i.e. the system that collectively manages all short-term business surveys. The information used, properly normalized and standardized, covers all the survey units involved in short term business surveys in the period 2016 to 2024.

1. Introduction

Short-term business statistics are governed by specific European regulations that define the structure of questionnaires and surveys, the scope of observations, the reference population, sampling methodology, and the accuracy of estimates. In the context of enterprises—particularly within Italy's production sector, which is predominantly composed of small businesses—regulatory requirements necessitate the repeated inclusion of large enterprises in survey samples over time.

The analysis presented in this article is aimed at investigating characteristics of enterprise participation in surveys conducted by the Italian National Institute of

[1] The article was only possible thanks to the joint work of the authors. In particular, Claudio Ceccarelli wrote section 1 and 6, Francesca Rossetti wrote section 2 and 4, Katia Bontempi wrote sections 3 and 5.

Statistics (ISTAT) and complements previous work conducted using the survival data analysis approach (Binci *et al.*, 2025).

The analysis carried out in this paper aims to verify the presence of seasonal components in the trend of response rates by applying time-series analysis and checking whether different behaviours can be observed between surveys with differing characteristics. The data used cover all the survey units involved in two specific short-term surveys, Employment in large enterprises and Retail trade, over the period 2016-2024.

Section 2 details the characteristics of the short-term business surveys and describes the characteristics of the platform used for the data collection, “The Business Statistical Portal” (BSP). Section 3 outlines the characteristics of the business surveys being analysed. Section 4 presents the methodology underlying the work. The principal results are illustrated in Section 5, while Section 6 provides a summary of the findings and discusses future research directions.

2. Centralized Management of Short-Term Business Surveys

Short-term Business Statistics (STS) are economic indicators collected on a monthly or quarterly basis, essential for tracking business cycles and supporting monetary and fiscal policies. These indicators are regulated by Regulation EC No. 2152/2019 amending Regulation EC No. 1158/2005 and all further implementing and amending regulations. Furthermore, STS form part of the National Statistical Programme, which establishes the rules for data collection and management.

The introduction of a centralized access system through the Business Statistical Portal (BSP) and a dedicated Contact Centre has optimized enterprise support by providing standardised responses, reducing the burden on respondents, and improving communication management (Fazio *et al.* 2013). This innovation has fostered a virtuous cycle in business collaboration, resulting in significant increases in response rates for the surveys considered. In particular, the BSP allows for an integrated and harmonized management of all survey phases, enabling shared access to information based on user profiles.

To ensure high-quality and timely data, communications with enterprises have been improved through a revision of the ISTAT information letter, focusing on deadlines, along with a notification system that includes certified emails (PEC), regular emails, and telephone reminders and a compliance penalty system has been introduced (Binci *et al.*, 2025).

The use of the BSP has standardised the data collection processes, allowing the comparability of enterprises’ response behaviour. At the same time, the standardisation of the STS makes a long series of data available for each of the surveys (Bellini *et al.* 2019).

3. The Short-Term Business Surveys selected for the analysis

The analysis focuses on two surveys that fall within the scope of the STS and share similar methodological characteristics: Monthly Survey on Employment, Working Hours, Wages, and Labour Costs in Large Enterprisesand and Monthly Retail Trade survey. These surveys were selected because they allow for a comparative evaluation of response rates across two different sampling designs (see Table 1).

Table 1 – *Selected short-term business surveys: main characteristics.*

Survey	Observation field	Sampling design	Statistical unit
Employment, Working Hours, Wages, and Labour Costs in Large Enterprises (OCC)	Enterprises with at least 500 employees	Census survey for enterprises with at least 500 employees	Functional unit[2]
Retail trade (DETT)	Enterprises with main economic activity in sec G of the Nace Rev. 2 classification	Stratified random sampling for enterprises with less than 50 employees - Census survey for enterprises with at least 50 employees	Enterprises

Istat

3.1 The Monthly Survey on Employment, Working Hours, Wages, and Labour Costs in Large Enterprises Survey

The Monthly Survey on Employment, Working Hours, Wages, and Labour Costs in Large Enterprises (OCC) is a critical statistical tool designed to monitor labour demand in enterprises with employees.

The OCC survey is part of the National Statistical Program (IST-00050), and subject to Legislative Decree No. 322/1989, which mandates enterprises to provide accurate labour data. The results are processed to generate monthly index numbers,

[2] Functional unit: the unit within an enterprise that groups together all parts contributing to the conduct of an economic activity at the class level (four digits) of the NACE Rev.1 classification. It is an entity corresponding to a system of information that allows at least the value of production, intermediate consumption, staff expenses, operating results, employment, and gross fixed investments to be provided or calculated for each economic activity unit.

published in the IstatData[3] database, while aggregated annual data appear in the Italian Statistical Yearbook (ASI). The survey targets large enterprises in industry and services, specifically those classified under Ateco 2007 sectors B to S, excluding temporary employment agencies, with at least 500 employees; the enterprises surveyed number is around 1,600. Key variables recorded include employee positions, workforce inflow and outflow, hours worked, paid but non-worked hours, unpaid strike hours, and wage guarantee fund hours. Data on solidarity contract hours, detailed wage components, social charges, and job vacancies at the end of each quarter are also collected. These indicators are analysed by qualification categories: clerical staff, manual workers, and executives.

The OCC survey is part of an integrated system alongside the Quarterly Survey on Job Vacancies and Hours Worked and the Quarterly Survey on Employment, Wages, and Social Charges, ensuring comprehensive coverage of employment trends. By integrating data coming from different sources, these surveys help policymakers track employment dynamics, wage trends, and labour costs, contributing to informed economic decisions.

3.2 Retail trade survey

The monthly retail trade survey (DETT) is a sampling-based statistical survey aimed at monitoring the performance of the retail sector across Italy. The survey is conducted in strict adherence to European regulatory frameworks, notably the EU Regulation 2019/2152, issued by the European Parliament and Council on November 27, 2019, replacing the earlier Regulation (EC) No. 1165/1998 and subsequent updates and is part of the National Statistical Program (IST-00151).

The survey is designed to capture essential economic trends in consumer purchasing behavior, providing data-driven insights for policymakers, businesses, and researchers. The methodology ensures statistical representativeness through a cut-off sampling strategy: for enterprises with fewer than 50 employees, a random selection is used, while enterprises with 50 or more employees are included in the survey by census.

The survey focuses on approximately 8,000 enterprises engaged in retail commerce, forming a longitudinal study where certain units overlap over time to maintain consistency and comparability. The largest enterprises, defined as those with more than 50 employees, are sourced from the Statistical Archive of Active Enterprises (ASIA) to ensure comprehensive sector coverage. By adopting a strategic sampling approach, ISTAT aims to create a dataset that accurately represents the full spectrum of retail businesses, from small independent retailers to large-scale commercial enterprises.

[3] https://esploradati.istat.it/

The information gathered includes total sales value, broken down by product category and sales method (such as brick-and-mortar stores, e-commerce platforms, door-to-door sales, and vending machines). Additionally, enterprises report the number of retail locations open to the public and provide year-over-year comparisons of sales performance for the corresponding months. The collected data undergoes extensive processing to generate monthly index numbers, which serve as a benchmark for analysing retail activity fluctuations. These results are subsequently published in the ISTAT press release, "Retail Trade," and made accessible via the IstatData database, ensuring transparency and availability for further analysis.

4. Applying ARIMA Models to Identify Seasonal Patterns

The application of seasonal adjustment techniques using an additive approach is a widely recognized method in time series analysis, aimed at eliminating seasonal fluctuations to isolate the underlying structure of a dataset. The seasonal adjustment of time series data with intra-annual frequency is based on the assumption that such series can be represented as a combination of distinct orthogonal components: the cycle-trend (CT) component, which reflects the medium- to long-term trend of the time series, unaffected by short-term fluctuations; the seasonal component (S), which recurs throughout the year and captures fluctuations attributable to meteorological, customary, or legislative factors and the irregular component (I), which results from erratic factors.

Under the additive model assumption, the data series (Y) can be expressed as:

$$Y = CT + S + I \tag{1}$$

Once the individual components are identified using appropriate statistical techniques (seasonal adjustment procedures), the seasonally adjusted series is obtained by subtracting the seasonal component from the original time series:

$$Dest(Y) = Y - S = CT + I \tag{2}$$

This method allows for the isolation of the underlying trend and irregular fluctuations, ensuring a more accurate representation of the fundamental dynamics within the dataset. In this study, the first step—identification of seasonality—was implemented to assess whether seasonal fluctuations were present. The procedure was intentionally halted before the second phase, as the goal was not full seasonal adjustment but rather verification of the existence of seasonal influences in the dataset. To fulfil this purpose, an autoregressive integrated moving average

(ARIMA) model was employed. This parametric approach utilizes filter-based techniques to process time series data, capturing underlying statistical dependencies and preparing the dataset for further modelling. The ARIMA model consists of three distinct components: AR (Autoregressive), which captures dependency on past observations within the time series; I (Integrated), which determines the number of differencing steps required to transform the series into a stationary format; and MA (Moving Average), which models dependencies based on past error terms (Dagum *et al.* 1999).

The key preprocessing step involved the application of first-order differencing (LAG(1)), which serves multiple statistical purposes: stabilizing the mean of the series over time, eliminating non-stationary trends or systematic components, and preparing the dataset for subsequent ARIMA modelling. These steps are critical, as stationarity is a fundamental requirement for the efficacy of ARIMA models in forecasting and trend analysis.

$$LAG(1) = Y_m - Y_{m-1} \tag{3}$$

A central objective of this study was to determine whether the removal of the trend component would reveal residual seasonal structures in the autocorrelation analysis. This approach is particularly useful in cases where apparent seasonal patterns are suspected but not explicitly confirmed. By the application of first-order differencing in order to eliminate the trend component and to analyse residual autocorrelation, researchers can assess whether additional adjustments or modelling refinements are necessary to capture persistent seasonal effects in the dataset.

In what follows, we propose applying a time series approach to analyse response rates observed over time (Ceccarelli *et al.*, 2016). These rates pertain to the DETT and the OCC surveys; to ensure comparability between the data collected in the two surveys, enterprises are used as the statistical units in this analysis. In the OCC survey, where the reference units are functional units, an occurrence of non-response is defined as the situation in which no functional unit has provided a response.

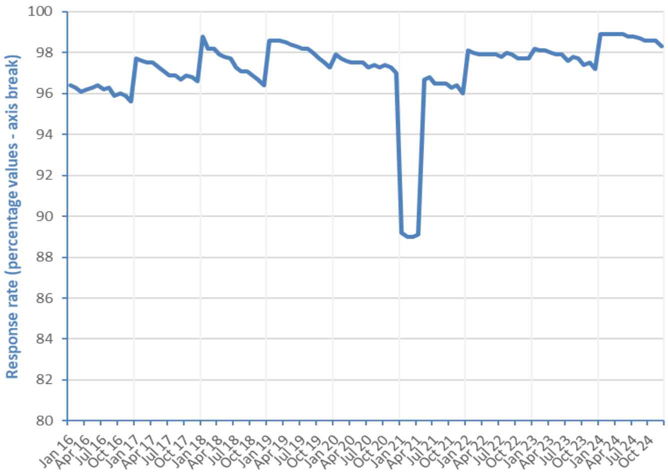
In every survey occasion, response rates are calculated as the ratio between the number of responding units and the number of units included in the sampling list. All analyses were undertaken using SAS Version 9.4.

5. Results

Both surveys considered in this study are characterised by a time series consisting of 108 observations, which represent the response rates recorded at each survey occasion between 2016 and 2024.

Figure 1 displays the time series of response rates for the OCC survey. The y-axis shows the response rate values on a scale starting at 80%, which better capture the fluctuations in the series. This survey has very high response rates, with a mean value of 97.2%, a minimum value of 89.0% (in March 2021) and a maximum value of 98.9% (from January to May 2024)[4].

Figure 1 – *Time series of response rates for the OCC survey. Years 2016-2024.*



Despite the evident stability in the response behaviour of the enterprises involved in the survey, a stepped profile can still be recognised, with a peak in January followed by a slight decrease in subsequent months each year[5].

The time series obtained by applying first-order differencing to the original time series is shown in Figure 2: the new series appears to be stationary, with values centred around zero, except for the peaks corresponding to January.

Table 2 shows the values of the autocorrelations of the observations with past data, from lag 1 to lag 24, for both the time series: the observed series and the first-order differenced series. Positive autocorrelations are coloured green, with the intensity of the colour decreasing as the autocorrelation value decreases. The observed series of response rates for the OCC survey shows relevant positive autocorrelation for lags up to eight, indicating that the series is non-stationary,

[4] The March 2021 drop in response rates is linked to COVID-19 emergency management, as it coincided with the final deadline for submitting previously suspended data, possibly impacting business participation.

[5] Please note that immediately prior to the data collection phase in January, the enterprises included in the sample receive an information letter from ISTAT, informing them of their involvement in the survey and providing all the details necessary for completing it.

whereas the autocorrelations between observations for the new time series confirm that the series is stationary and that a seasonality component is no longer present.

Figure 2 – *Time series obtained by applying first-order differencing to the response rates time series for the OCC survey. Years 2016-2024.*

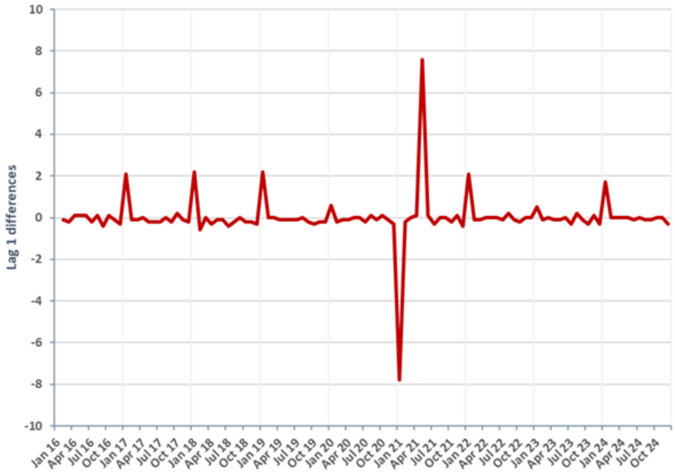


Table 2 – *Autocorrelations of the response rates and the first-order differenced time series for the OCC survey (lags 1 to 24).*

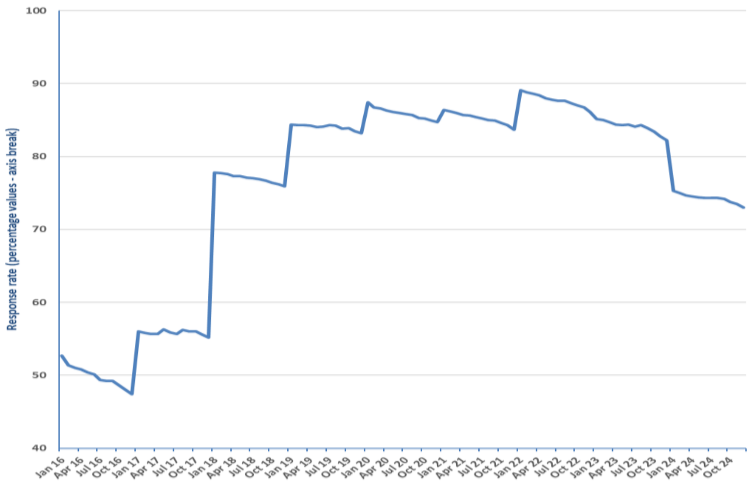
Time Series	Up to lag	Chi-sqr	DF	Pr > ChiSqr	Autocorrelations*					
					Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
Original time series	6	129.54	6	<.0001	0.789	0.577	0.366	0.168	0.148	0.141
	12	134.99	12	<.0001	0.135	0.132	0.086	0.038	0.000	-0.038
	18	140.11	18	<.0001	-0.047	-0.051	-0.060	-0.076	-0.102	-0.122
	24	162.41	24	<.0001	-0.131	-0.138	-0.157	-0.175	-0.188	-0.187
First-order differencing time series	6	20.52	6	0.0022	0.004	-0.011	-0.029	-0.423	-0.033	-0.004
	12	22.00	12	0.0375	-0.012	0.096	0.01	-0.022	0.003	-0.05
	18	22.27	18	0.2202	-0.016	0.007	0.017	0.025	-0.011	-0.027
	24	23.37	24	0.4983	-0.014	0.022	0.004	-0.012	-0.038	-0.075

Legend: 0.000-0.333 0.334-0.666 0.667-1.000

** Each column contains autocorrelations for lags multiples of 6 (e.g., "Lag 1" includes lags 1, 7, 13, and 19)*

The time series of response rates for the DETT survey is shown in Figure 3. The y-axis shows the response rate values on a scale starting at 40%, to better visualise the fluctuations in the series.

Figure 3 – *Time series of response rates for the DETT survey. Years 2016-2024.*

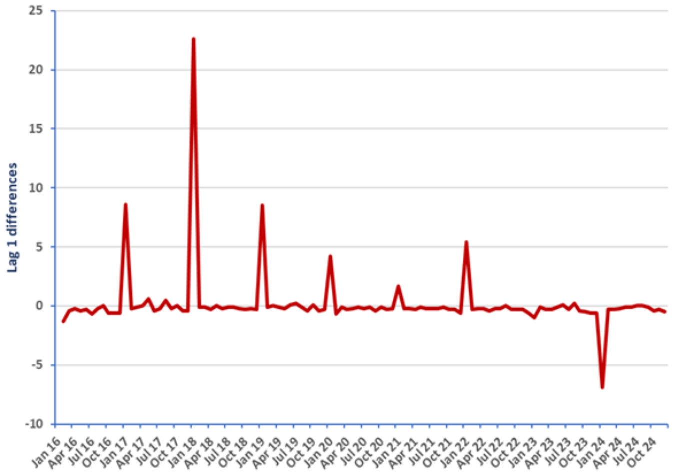


The series presents good response rates, especially since 2018, with an average value of 78.5%, a minimum of 47.5% in December 2016, and a maximum of 89.4% in January 2022. The series presents a stepped profile, more pronounced compared to the OCC response rates series. Once again, there is a peak in January[6], followed by a slight decrease in response rates in subsequent months each year.

Figure 4 displays the time series obtained by applying first-order differencing to the original one. The values in the new series fluctuate around zero, with very noticeable peaks occurring in January each year. This suggests that first-order differencing alone is insufficient to make the series stationary.

[6] Please note that immediately prior to the data collection phase in January, the enterprises included in the sample receive an information letter from ISTAT, informing them of their involvement in the survey and providing all the details necessary for completing it.

Figure 4 – *Time series obtained by applying first-order differencing to the response rates time series for the DETT survey. Years 2016-2024.*



As shown in Table 3, the original time series exhibits positive autocorrelations up to lag 24, confirming its non-stationary behavior. The autocorrelations of the transformed series still display significant values at lags 12 and 24, indicating the persistence of an annual seasonal component.

Table 3 – *Autocorrelations of the response rates and the first-order differenced time series for the DETT survey (lags 1 to 24).*

Time Series	Up to lag	Chi-sqr	DF	Pr > ChiSqr	Autocorrelations*					
					Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
Original time series	6	523.96	6	<.0001	0.965	0.929	0.893	0.858	0.822	0.785
	12	838.89	12	<.0001	0.749	0.713	0.676	0.639	0.603	0.567
	18	968.85	18	<.0001	0.519	0.472	0.426	0.381	0.338	0.293
	24	990.18	24	<.0001	0.250	0.208	0.166	0.124	0.084	0.045
First-order differencing time series	6	0.25	6	0.9997	-0.031	-0.023	-0.018	-0.007	-0.003	-0.02
	12	41.47	12	<.0001	-0.020	-0.005	-0.014	-0.025	-0.045	0.577
	18	42.07	18	0.0011	-0.037	-0.032	-0.035	-0.017	-0.005	-0.028
	24	50	24	0.0014	-0.021	-0.029	-0.018	-0.038	-0.070	0.220

Legend: 0.000-0.333 0.334-0.666 0.667-1.000

** Each column contains autocorrelations for lags multiples of 6 (e.g., "Lag 1" includes lags 1, 7, 13, and 19)*

6. Synthesis of Results and Research Outlook

The time series analysis approach applied to two different datasets yields seemingly different results. A previous study, which examined participation in

ISTAT short-term surveys using survival curve analysis (Binci *et al.,* 2025), observed that response burden acts as a deterrent to continued participation for small enterprises compared to larger ones, as the organizational structures of large enterprises allow them to bear such burden. Similarly, the results presented in this article suggest that the differences observed in the analysis of the two time series may be explained by differences in sampling and strategy design: the OCC survey involves only enterprises with more than 500 employees, whereas the DETT survey also includes smaller enterprises. However, the additional analysis presented in the appendix does not confirm this interpretation. Further research is needed—possibly extending the analysis to other surveys—to investigate the factors underlying the different seasonal patterns observed in response rates. Potential hypotheses to explore include characteristics of the types of enterprises in the samples, such as differences in the sectors in which they operate.

In the set of studies that Directorate for data collection **(**DCRD) has been carrying out for a number of years, the approach presented is aimed at bringing to light any respondent behaviour that, overall, may be repetitive over time regardless of the other characteristics that we have analysed in other works. Highlighting the non-stationarity of the series of respondents in the two surveys gives us the possibility to make targeted actions to support the survey and avoid costly and inefficient ones. It should always be stressed that the adjective 'costly' does not express a quantity whose unit of measurement is “euro” or “dollar”, but, in our case, “time”. This cost element enters predominantly into the surveys as it is the fundamental parameter in fieldwork and data collection. It is on this parameter that we must intervene to increase the quality of the information produced.

Appendix

We present the results of an additional analysis carried out following a suggestion from the referee, whom we thank. The aim is to assess whether the different seasonal patterns in response rates observed in the OCC and DETT surveys are influenced by differences in sampling design. In this follow-up, the analysis for the DETT survey refers to the same target group of enterprises involved in the OCC survey, focusing exclusively on enterprises with more than 500 employees. Table A1 displays the autocorrelation values for both the response rate series and the series obtained through first-order differencing, highlighting patterns similar to those calculated for the full sample (see Table 3); when assessing the robustness of the analysis, note that limiting the focus to enterprises with over 500 employees reduces observations from about 8,000 to 160.

Table A1 – *Autocorrelations of the response rates and the first-order differenced time series for the DETT survey – Only enterprises with 500+ employees (lags 1 to 24).*

Time Series	Up to lag	Chi-sqr	DF	Pr > ChiSqr	Autocorrelations*					
					Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
Original time series	6	491.07	6	<.0001	0.954	0.913	0.871	0.829	0.779	0.732
	12	741.04	12	<.0001	0.690	0.643	0.596	0.561	0.527	0.494
	18	856.00	18	<.0001	0.458	0.428	0.392	0.362	0.342	0.323
	24	911.44	24	<.0001	0.297	0.282	0.266	0.253	0.237	0.216
First-order differencing time series	6	4.840	6	0.565	-0.120	-0.072	-0.040	0.109	-0.092	-0.041
	12	24.560	12	0.017	-0.061	0.063	-0.179	-0.033	-0.039	0.346
	18	28.110	18	0.060	-0.109	0.039	-0.084	-0.022	-0.069	0.050
	24	47.100	24	0.003	-0.105	0.033	-0.091	0.035	0.005	0.339

Legend: 0.000-0.333 | 0.334-0.666 | 0.667-1.000

* *Each column contains autocorrelations for lags multiples of 6 (e.g., "Lag 1" includes lags 1, 7, 13, and 19)*

References

BELLINI G., BINCI S., DE GAETANO L., MONETTI F., PAPA P. 2019. Effects of Istat CDC (Centralised Data Collection) approach on the reduction of the Total Survey Error: experiences in data collection implementation field. In *Workshop on Statistical Data Collection 'New Sources and New Technologies'*.

BINCI S., BONTEMPI K., CECCARELLI C., MONETTI F., ROSSETTI F. 2025. Burden analysis for short-term business surveys respondents. *Rivista Italiana di Economia Demografia e Statistica* , Vol. LXXIX – No. 3 Luglio – Settembre 2025 pp. 251-262.

CECCARELLI C., GRAZIANI C., LORIGA S., SALVATORE M.A., SPIZZICHINO A. 2016. Gli errori campionari nelle serie storiche destagionalizzate: alcune analisi sui dati mensili della rilevazione sulle forze lavoro. *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXX – No. 4 Ottobre-Dicembre 2016 pp. 55-66.

DAGUM E, *et al.*1999. X11ARIMA version 2000: Foundations and User's Manual.

FAZIO N.R., MURGIA .M, NUNNARI A. 2013. The Business statistical Portal: a new way of organizing and managing data collection processes for business surveys in Istat. *Working Papers 14*, Seminar on Statistical Data Collection. Unece - Conference of European statisticians, 25-27 September 2013, Geneva, Switzerland.

Katia BONTEMPI, Italian National Institute of Statistics, katia.bontempi@istat.it
Claudio CECCARELLI, Italian National Institute of Statistics, clceccar@istat.it
Francesca ROSSETTI, Italian National Institute of Statistics, frrosset@istat.it;

MAFIA HOMICIDES AND LAW ENFORCEMENT

Luigi Maria Solivetti

Abstract. The concept of deterrence emerged during the Enlightenment owing to the works of Beccaria and Bentham, who posited that the realistic threat of punishment deters people from committing crimes. In the nineteenth century, however, under the influence of positivism, this concept was discarded, and the offender's constitution, mental disorders and/or socioeconomic conditions substituted deterrence as the primary crime determinants. Since the late 1960s, however, the deterrence theory has been revived, and new theoretical and empirical works have been dedicated to it. Despite this, the hypothesis that punishment is the key to crime control has not been consistently endorsed by empirical evidence. The present study intended to test this hypothesis by analysing the evolution of intentional homicide rates in Italy's Mezzogiorno. In the past, this region has been well-known for its much higher rates compared to other European countries and the rest of Italy. Importantly, Mezzogiorno has also been the cradle of the most famous and feared crime organisations, the Mafia-type gangs. Since the 1990s, however, the fight against Mezzogiorno's crime has benefited from more severe sanctions and better-organised enforcement. We analysed the impact of these changes using interrupted time series regression models on series spanning a 40-year period. Our findings support the hypothesis that more robust law enforcement significantly affects intentional homicide rates by making the threat of punishment more realistic.

1. Introduction: Naissance and decline of the concept of deterrence

The concept of deterrence – namely, discouraging a criminal act through fear of the consequences – has been characterised by a long history but also by an oscillating endorsement.

In the eighteenth century, Cesare Beccaria posited that a citizen, confronted with the choice between law-abiding and law-breaking, would inevitably choose the former as long as the government imposes on the latter a sanction as severe, sure and swift as to remove the advantage associated with the illicit opportunities. Ultimately, while self-interest, which resides in everyone, would urge individuals to take advantage of illicit opportunities, the threat of certain and swift punishment would restrain them from doing so.

Jeremy Bentham, in turn, affirmed that human behaviour is “under the governance of two sovereign masters, pain and pleasure”. Consequently, man will choose the course of action that has the greater sum of benefits over costs. Bentham, unlike Beccaria, tried to identify the various configurations of benefits and costs. Thus, people will be attracted by the benefits coming from crime – the pleasures of the senses, wealth, and power over other people – but will be restrained by the threat of costs such as imprisonment, loss of reputation, the feeling of guilt, etc.

Despite these differences, Beccaria and Bentham shared a typical Enlightenment tenet. Namely, that, since *all* men are endowed with free will and reason, *all* men can calculate the benefits and costs of crime and choose between law-abiding and law-breaking on that ground. This tenet has been subject to criticism during the nineteenth century and up to the present.

The so-called *moral statisticians*, such as André-Michel Guerry and Adolphe Quetelet, did not reject the free will tenet but thought that, at the level of large numbers, crime and other social pathologies should be regarded as the product of the socioeconomic environment rather than an individual choice. This shift in focus redirected criminological studies from deterrence to social conditions.

Later, positivist criminologists, such as Cesare Lombroso and his followers, posited, in conflict with the idea that all men are endowed with reason and free will, the existence of anthropological differences between criminals and non-criminals.

Other positivists, such as Enrico Ferri, shifted the root of crime from biological to psychological-social features but maintained the rejection of free will and supported the hypothesis of the heterogeneity of criminals and crime factors.

Late nineteenth-century contributions continued to focus on social conditions, but they followed two distinct criminological perspectives. The first one was inspired by Èmil Durkheim’s study on *anomie* and suggested that crime results from the breakdown of the social standards necessary for regulating human behaviour. Crime, being a social fact, does not stem from individual conscience but from previous social facts. All this left no room for the concept of deterrence.

The second perspective, exemplified by the work of Willem Bonger, advanced a Marxist theory of crime, in which crime emerged from the unequal distribution of resources and the egoistic impulses generated by a capitalist society. Consequently, punishment was regarded as class violence rather than as a tool of crime prevention.

In the twentieth century, the two aforementioned perspectives – namely, anomie and Marxist criminology – spawned new approaches. Anomie generated the *relative deprivation* theory, which posited that social pressure to succeed materially in the face of scarce legitimate opportunities leads to crime. For decades, deprivation theory has been the dominant frame of reference for criminological studies. Since it focused on culture and economic structure, this theory did not at all encourage an investigation of the role of deterrence.

In turn, Marxist criminology of the positivist era was followed by radical criminology. Radical criminology scholars have shared with their predecessors the assumption that crime was the outcome of capitalism. However, they have regarded the crime of the underprivileged classes not as a symptom of maladjustment but as an active protest against the system. Ultimately, in this perspective, punishment is not perceived as what counterbalances self-interest in breaking the rules but as the violence by which the dominant class preserves its supremacy.

Another influential twentieth-century theory, *labelling theory*, has regarded crime as a social product and the criminal as someone who has been accidentally and arbitrarily labelled as such. In this perspective, the purpose of the social sciences would have been to identify the paths and interactions leading to labelling rather than to study how deterrence could restrain crime.

2. The reemergence of deterrence

Despite the dominant role of relative deprivation theory, the late 1960s saw the publication of two seminal works on deterrence.

Gary Becker (1968), an economist, taking inspiration from Beccaria and Bentham, assumed that people will commit an offence if the utility of doing so exceeds the utility of not doing so. In this perspective, he hypothesised that the number of crimes committed by any person is a function of their probability of conviction, their punishment if convicted, and other variables, such as the income available to them through legal and other illegal activities.

Jack Gibbs (1968), a sociologist, believed that the only realistic approach to estimate the deterrence impact was to analyse the effect of the actual legal reactions to crimes in comparable social contexts, measuring these reactions in terms of severity and certainty of imprisonment. Gibbs' study, therefore, although of an empirical nature, was primarily inspired by Beccaria because, unlike Becker, it focused only on punishment, ignoring socioeconomic covariates.

Gibbs's and Becker's articles ignited great interest in testing the impact of deterrence. Deterrence has become the subject of numerous analyses, employing a variety of methods.

A substantial group of studies analysed deterrence in terms of perception, assuming that deterrence impacts crime as long as punishment is perceived as a real threat (Waldo and Chiricos 1972; Paternoster 1987). This approach implies focusing on the micro, individual dimension and making recourse to surveys directed to identify self-reported criminality.

Another substantial group of studies has used macro data. Within this group, it is possible to identify two distinct waves of studies (Nagin 2013). The first one has examined the relationship between deterrence and crime by comparing states and other territorial entities based on their levels of punishment and crime rates (Ehrlich 1973; Geerken and Gove 1977). Punishment has been measured by the clearance ratio, the ratio of prison admissions to reported crimes (i.e. certainty of punishment), and the median time served (i.e. severity of punishment).

A second wave of macro studies has utilised longitudinal data to analyse deterrence and crime across states or other territorial units and over time. In these studies, deterrence has typically been measured by imprisonment rates, clearance ratios or severity of sentence (Entorf and Spengler 2000; Abramovaite *et al.* 2023).

3. Deterrence literature: some considerations

Literature on deterrence and crime has yielded inconsistent findings throughout its long history, largely due to the diverse methods employed.

The original formulation of the theory in the eighteenth century relied on postulates (the rationality of man's actions and free will) similar to those supporting the portrayal of *homo oeconomicus*, and neither Beccaria nor Bentham thought it necessary to corroborate their hypotheses with empirical evidence. Following the resurgence of the deterrence concept, numerous studies have provided sophisticated equations of the deterrence-crime link without empirical analysis. Several other studies provided hypotheses in the form of nonmathematical conceptual theories.

Twentieth and twenty-first-century empirical studies have reached contradictory conclusions about the impact of deterrence, and their methods have often been criticised. Comparative analyses have frequently overlooked the multifaceted reality of punishment, e.g. the fact that statutory penalties do not always correspond to the penalties imposed by the judge, and the penalties imposed can differ significantly from the penalties actually served. Macro studies comparing capital-punishment states with non-capital-punishment ones have often neglected the infrequent occurrence of capital punishment or the severity of non-capital sanctions. Other macro studies have compared societies in terms of penalties and crime rates, but have ignored the relevance of extralegal factors. For instance, comparisons between countries would be biased by their socio-cultural differences. Lastly, other studies have compared societies distant in time. For instance, comparing crime rates in eighteenth-century England – when capital punishment could be imposed for more than 200 offences – with those in the same country in the twenty-first century is nonsensical because it involves comparing two incomparable social contexts.

Studies that avoided the previous weaknesses have not been immune to criticism. In particular, the numerous studies that measured deterrence by imprisonment rates inevitably obtained the combined effect of deterrence and incapacitation on crime rates rather than the effect of deterrence alone.

Regarding studies that focus on perceived deterrence, one cannot help but agree with their premise. Punishment cannot impact an individual's propensity to commit a crime unless it is perceived as a real threat. This perception differs for each individual. Therefore, micro studies would be potentially more accurate than macro ones. At the same time, it is also true that an analysis of deterrence centred on perception would imply either remaining within the boundaries of non-empirical models or making recourse to data from self-reported perceptions and self-reported criminality. And self-reported crime data present obvious weaknesses (Kleck and Sever 1980: 81 ff.). Ultimately, macro-level research is deemed superior to individual-level research that relies on self-reported data.

Having considered all the above, we believe that the most suitable method to analyse deterrence consists in:

- avoiding comparisons of contexts distant in time from each other;
- relying on objective facts, such as actual punishment and actual crime rates, more than on subjective interpretations of punishment and self-reported criminality;
- focusing on the impact on crime of specific changes in punishment within the same society: an approach meant to assure a substantial homogeneity of the extralegal factors and the legal system as a whole;
- using macro-level data;
- using panel models because they are intrinsically superior to cross-sectional models (Kleck and Sever 2018: 175).

4. Enforcement, deterrence, and Mafia homicides

Italy's homicide trend in Mezzogiorno[1] provides an excellent opportunity to test the effective impact of deterrence on crime. Differences in extralegal factors in Mezzogiorno are relatively limited. There are no differences across the Mezzogiorno regions in terms of the criminal justice system, while, in the past few decades, there have been significant nationwide changes in law enforcement and penalties. These changes provide an opportunity to analyse the deterrence-crime link. Lastly, time series concerning crime and enforcement are available. Mezzogiorno, the Mafia's original turf, has traditionally presented very high rates of intentional homicide (hereafter IH). In the early 1980s, a vast gap existed between the Mezzogiorno's IH rates and the rest of Italy's: 3 to 6 IHs per 100K population vs ~1. At that time, the average IH rate of the other West European countries was ~1.4, which was much lower than the rates in the Mezzogiorno but slightly higher than those in the rest of

[1] With "Mezzogiorno", we refer to the southern part of Italy, including Sicily but excluding Sardinia where Mafia-type organisations are sporadic.

Italy. Over the last years, however, Mezzogiorno seemed to have lost its criminal exceptionalism, no matter the Mafia-type families dominating the local context: 'Ndrangheta (Calabria), Camorra (Campania), Sacra Corona (Apulia), and Mafia proper (Sicily).

A factor in the alleged decline in homicidal violence might have been the more robust enforcement by the state. Since 1991, the Italian judiciary has performed an effective action against Mafia-type gangs by taking advantage of the (contentious) collaboration of Mafia's former affiliates (so-called *pentiti*), to whom reduced sentences and protection were granted. In 1992, after the Mafia killed two high-ranking magistrates overseeing anti-Mafia activities, new acts meant to strengthen the provisions to counteract Mafia-type crime were passed. Among other things, the acts expanded the possibility of seizing property and money of suspicious origin, serving as a financial and psychological deterrent to Mafia bosses. The new acts also provided a hard prison regime for Mafia-type criminals (1992). Since then, several Mafia bosses have been left behind bars until the end of their lives. Measures of law enforcement showed positive variations since 1992: Mezzogiorno's IH clearance rate grew from ~36% in the late 1980s to ~55% around 2020, while the ratio people-charged-with-IH / IH-number increased from 0.69 in 1991 to ~2 around 2020.

Ultimately, we hypothesise that:

H1. More robust law enforcement led to a decrease in Mezzogiorno's homicides by escalating the threat of punishment.

H2. Although since the 1990s there has been a decline in homicides in Italy as well as in Europe, the decline registered in Mezzogiorno was significantly higher than that of other regions. This would support the hypothesis of a causal link between tougher enforcement against the Mafia and the decline in homicides in Mezzogiorno.

5. Data and methods

Our response variable measures the time series of intentional, completed homicides (henceforth IH), calculated as $ln((IH_{it} + \overline{\overline{IH}})/population)$. We considered total IHs and the Mafia-type IHs (on average, 27.4% of Mezzogiorno's IHs). Mafia-type homicides are those classified by the police as "eminently characterised by the force of the criminal organisations' associative bond". The label potentially applies to various criminal organisations but was designed to target the Mafia. Owing to the ambiguous nature of some IHs, some Mafia-type IHs are not correctly classified. Hence, the total number of IHs is a variable that must be taken into account. IH data were recorded by the police and operated by Istat, the Italian Statistical Office.

The 1983-2022 time series concern the Mezzogiorno provinces, ranging from 30 to 33 according to the period.

To investigate these data, we employed interrupted time series analysis (ITSA), which is named so because the intervention is expected to "interrupt" the trend over time of the outcome variable (Shadish, Cook, and Campbell, 2002). As a model designed to evaluate the effectiveness of policy changes, ITSA employs an aggregate entity (e.g., hospital, city, region, or county) as the treatment unit and summary-level measures (e.g., mortality or crime rates) as the outcome.

When we have only the treatment group, a single treatment period, and a set of entities under study, the general ITSA regression model (Linden 2015) assumes the following form (1):

$$Y_{ti} = \beta 0 + \beta 1 T_{ti} + \beta 2 X_{ti} + \beta 3 X_{ti} T_{ti} + \varepsilon_{ti} \tag{1}$$

where Yti is the outcome variable for each time point t and each individual-level i, $\beta 0$ is the intercept or starting level of the outcome variable, $\beta 1$ is the slope of the outcome variable until the intervention, $\beta 2$ is the change that occurs in the period immediately after the intervention, $\beta 3$ is the difference between pre-intervention and post-intervention outcomes. Therefore, a significant $\beta 2$ indicates an immediate treatment effect, and a significant $\beta 3$ a treatment effect over time.

Significant $\beta 2$ and $\beta 3$ are not conclusive proof of a causal link between the intervention and the response. The intervention-response link might result from an unmeasured confounder. However, we assume that any time-varying confounder should exhibit relatively slower variations. Consequently, it would be distinguishable from the expected sharp variation following the intervention.

Nevertheless, to verify the robustness of the results, researchers usually resort to control series regarding subjects unaffected by the intervention. However, the changes introduced to better combat the Mafia (i.e. the *intervention*) were applied nationwide. Therefore, it would be impossible to find territorial units formally unaffected by the intervention. Still, we assumed that the changes in enforcement aimed at combating the Mafia were inconsequential where Mafia gangs were substantially absent. Therefore, we checked whether the fall in homicides has also been shared by provinces where Mafia gangs have been substantially absent. In practice, we compared the homicide trend in the 15 provinces (all belonging to Mezzogiorno) that registered the highest rates of Mafia-type conspiracy during the period preceding the intervention (1983-1991) with the trend in the 15 provinces with the lowest rates (all bar one outside Mezzogiorno). To do this, we used a two-group interrupted time series model that assumes the following form (2):

$$Y_{ti} = \beta 0 + \beta 1 T_{ti} + \beta 2 X_{ti} + \beta 3 X_{ti} T_{ti} + \beta 4 Z_i + \beta 5 Z_i T_{ti} + \beta 6 Z_i X_{ti} + \beta 6 Z_i X_{ti} T_{ti} + \varepsilon_{ti} \tag{2}$$

where, in addition to form (1), *Zi* is a dummy identifying the individual's assignment (treatment or control), and *ZiTti*, *ZiXti*, and *ZiXtiTti* are all interaction terms.

6. Results

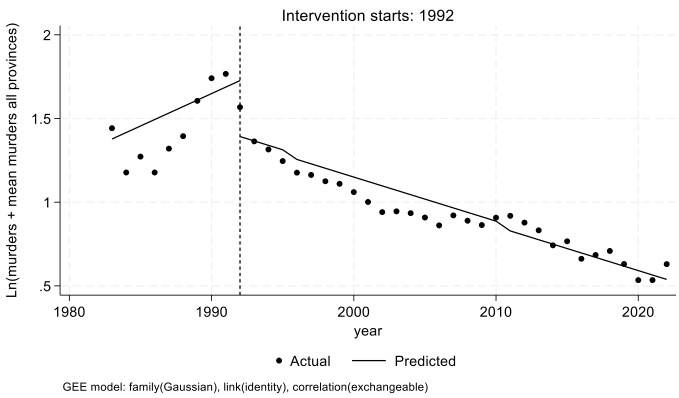
Table 1 and Figure 1 present the outcome of the ITSA[2] when the response variable is the time series of rates for all intentional homicides.

Table 1 - *Interrupted time series analysis (ITSA) with panel data. All Mezzogiorno provinces from 1983 to 2022. Homicide rates and intervention (1992).*

GEE population-averaged model	Number of obs	=	1,266
Group variable: Province	Number of groups	=	33
Family: Gaussian	Obs per group:		
Link: Identity	min	=	12
Correlation: exchangeable	avg	=	38.4
	max	=	40
	Wald chi2(3)	=	81.13
Scale parameter = 0.2999	Prob > chi2	=	0.000

Ln(homicides)	Coefficient	Robust std. err.	z	P>z
t	0.0388	0.0136	2.86	0.004
x 1992	–0.3341	0.0915	–3.65	0.000
x_t 1992	–0.0651	0.0146	–4.47	0.000
constant	1.3775	0.1276	10.8	0.000

Figure 1 - *Interrupted time series analysis (ITSA) with panel data. All Mezzogiorno provinces from 1983 to 2022. Homicide rates and intervention (1992).*



[2] We used the Stata module *xtitsa* by Linden (2015).

The intervention coincides with a sharp change in homicides. We observe that the slope of the time series is positive before the 1992 intervention (in the tables, t) and negative immediately after it (x), as well as in the long term (x_t) when compared to the pre-intervention period. All the coefficients are significant.

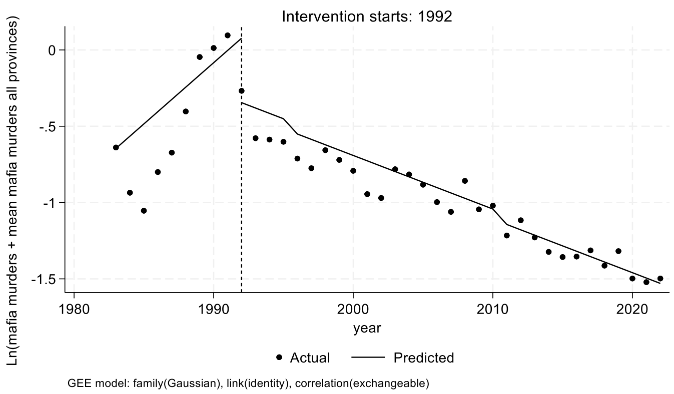
The outcome of the ITSA model when the response variable is the rates for Mafia-type intentional homicides (Table 2 and Figure 2) mirrors the outcome obtained with all intentional homicides. Again, there is a sharp fall in homicide rates immediately after the intervention and in the long term.

Table 2 - *Interrupted time series analysis (ITSA) with panel data. All Mezzogiorno provinces from 1983 to 2022. Mafia-type homicide rates and intervention (1992).*

GEE population-averaged model	Number of obs	=	1,266
Group variable: Province	Number of groups	=	33
Family: Gaussian	Obs per group:		
Link: Identity	min	=	12
Correlation: exchangeable	avg	=	38.4
	max	=	40
	Wald chi2(3)	=	81.13
Scale parameter = 0.9939	Prob > chi2	=	0.000

Ln(Mafia-type homic.)	Coefficient	Robust std. err.	z	P>z
t	0.0802	0.0293	2.74	0.006
x 1992	–0.4223	0.1757	–2.40	0.016
x_t 1992	–0.1153	0.0324	–3.56	0.000
constant	–0.6452	0.2325	–2.78	0.006

Figure 2 - *Interrupted time series analysis (ITSA) with panel data. All Mezzogiorno provinces from 1983 to 2022. Mafia-type homicide rates and intervention (1992).*



6.1 Robustness check

We assumed that the fall in homicide rates following the intervention could be attributed to the intervention itself. However, to verify the robustness of the results, we made recourse to the aforementioned two-group interrupted time series analysis. This shows (Table 3) that the changes regarding the controls were non-significant before and after the intervention (*t*, *x*, and *x_t*). Instead, the provinces with the highest rates of Mafia-type conspiracy presented an increase before the intervention (*z*), which was significantly higher than that of the controls (z_t). The Mafia-ridden provinces also exhibited a significant negative variation immediately after the intervention (*z_x*) and in the long term (z_x_t) when compared to the controls.

Table 3 – *Two-group interrupted time series analysis (ITSA) with panel data. All Mezzogiorno provinces from 1983 to 2022. Homicide rates and intervention (1992).*

GEE population-averaged model	Number of obs	=	1,200
Group variable: Province	Number of groups	=	30
Family: Gaussian	Obs per group:		
Link: Identity	min	=	40
Correlation: exchangeable	avg	=	40.0
	max	=	40
	Wald chi2(3)	=	108.03
Scale parameter = 0.1812	Prob > chi2	=	0.000

Ln(homicides)	Coefficient	Robust std. err.	z	P>z
t	0.0091	0.0117	0.78	0.4370
z	0.8227	0.1433	5.74	0.0000
z_t	0.0663	0.0216	3.06	0.0020
x 199*2*	–0.0205	0.0700	–0.29	0.7700
x_t 1992	–0.0171	0.0135	–1.27	0.2040
z_x 1992	–0.5917	0.1501	–3.94	0.0000
z_x_t 1992	-0.0930	0.0242	-3.85	0.0000
constant	0.6866	0.0860	7.98	0.0000

7. Conclusions

This study demonstrates that the fall in homicide rates corresponded to changes in crime policies. More effective enforcement and the parallel escalation in the probabilities of punishment and severity of penalties led to a marked decline in Mezzogiorno’s homicides as a whole and Mafia-type homicides.

At the end of the four-decade period considered, the fall in homicides was such that Mezzogiorno's rates became only fractionally higher than the rest of Italy's rates and neatly lower than the average IH rate for the other West European countries. This ended the long-lasting Mezzogiorno's exceptionalism in terms of homicidal violence.

From a theoretical perspective, the present analysis, based on actual changes in law enforcement and crime rates, demonstrates that a decrease in crime followed an increase in enforcement. The decrease in homicides was, in turn, non-significant in those provinces where the new anti-Mafia measures have been relatively inconsequential owing to a substantial absence of Mafia gangs.

All this evidence emerged from a macro-level analysis using panel data. A micro-level investigation would likely reveal a range of individual reactions to a change in enforcement. Any verification of the eighteenth-century scholars' tenet that deterrence affects crime because all men rationally calculate the costs and benefits of law-breaking is probably beyond the reach of an empirical macro-investigation such as the present one. However, this study's findings allow us to conclude at least that more robust enforcement results in an average decrease in homicide rates. This more robust enforcement encompasses, firstly, the positive variations in the homicide clearance rate and in the number of people charged with IH. The role of *pentiti* cannot be overlooked, although it is more difficult to quantify it. In any case, the *pentiti*'s contribution led to positive variations in the IH clearance rate and in the number of people charged with IH. The seizure of property and money of suspicious origin decreased the benefits in the Mafia's activities, but did not affect the *mafiosi*'s freedom. Instead, the positive variations in clearance rate and in people charged with IH had an incapacitation effect on criminals through incarceration, and it also produced a higher potential deterrence because it increased the risk of being brought to justice and receiving harsher penalties (e.g. hard prison regime). In turn, people belonging to criminal organisations are expected to be particularly sensitive to deterrence because they plan crimes "in a cool state of blood". Ultimately, everything suggests that the fall in homicides was also the outcome of the rational assessment of the risk of having to account for them.

References

ABRAMOVAITE J., BANDYOPADHYAY S., BHATTACHARYA S., COWEN N. 2023. Classical Deterrence Theory Revisited. *European Journal of Criminology*, Vol. 20, No 5, pp.1663–80.

BECKER G.S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy*, Vol. 76, No 2, pp. 169–217.

EHRLICH I. 1973. Participation in Illegitimate Activities: A Theoretical and Empirical Investigation. *Journal of Political Economy*, Vol. 81, No 3, pp. 521–65.

ENTORF H., SPENGLER H. 2000. Socioeconomic and Demographic Factors of Crime in Germany: Evidence from Panel Data of the German States. *International Review of Law and Economics*, Vol. 20, No 1, pp. 75–106.

GEERKEN M., GOVE W.R. 1977. Deterrence, Overload, and Incapacitation: An Empirical Evaluation. *Social Forces*, Vol. 56, No 2, pp. 424–47.

GIBBS J.P. 1968. Crime, Punishment, and Deterrence. *The Southwestern Social Science Quarterly*, Vol. 48, No 4, pp. 515–30.

KLECK G., SEVER B. 2018. *Punishment and Crime: The Limits of Punitive Crime Control*. New York and London: Routledge.

LINDEN A. 2015. Conducting Interrupted Time-Series Analysis for Single- and Multiple-Group Comparisons. *The Stata Journal*, Vol. 15, No 2, pp. 480–500.

NAGIN, D. 2013. Deterrence in the Twenty-First Century. *Crime and Justice*, Vol. 42, No 1, pp. 199–263.

PATERNOSTER R. 1987. The Deterrent Effect of the Perceived Certainty and Severity of Punishment: A Review of the Evidence and Issues. *Justice Quarterly*, Vol. 4, No 2, pp. 173–217.

SHADISH W.R., COOK T.D., CAMPBELL D.T. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. New York: Houghton Mifflin.

WALDO G.P., CHIRICOS T.G. 1972. Perceived Penal Sanction and Self-Reported Criminality: A Neglected Approach to Deterrence Research. *Social Problems,* Vol. 19, No 4, pp. 522–40.

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

CRIMINAL RECIDIVISM. TOWARDS RELIABLE AND TRANSPARENT PREDICTIVE MODELS

Flavia Tagliafierro, Claudio Caterino

Abstract. The prediction of criminal recidivism through machine learning (ML) models raises significant ethical, legal, and methodological challenges. This article promotes for a transparency and explainability-oriented approach by comparing three predictive models – logistic regression, random forest, and neural networks – applied to the COMPAS dataset of criminal history data, released by ProPublica a non-profit journalism organization in USA. To assess the coherence and readability of algorithmic decisions interpretability techniques such as SHAP values are employed. The analysis also considers the implications of adjusting the decision threshold to increase false positives for supportive – rather than punitive – purposes, emphasizing the greater ethical and social acceptability of such a strategy. The discussion is complemented by an overview of the regulatory developments in Italy and the European Union regarding the use of predictive technologies in the criminal justice system.

1. Introduction

Recidivism is broadly defined as the tendency of previously convicted individuals to reoffend, serving as a key indicator of risk, social dangerousness, and the effectiveness of penal and rehabilitative measures (Baratta, 1998; Lappi-Seppälä, 2003). Identifying the factors associated with reoffending as well as understanding their interactions can support prevention strategies, resource allocation, and evidence-based decision-making in the justice system.

Definitions and measurements of recidivism vary across jurisdictions. Some systems consider only new convictions while others include arrests or reports (Weatherburn *et al.,* 2003). In Italy, recidivism is regulated by articles 99–105 of the Penal Code which influence sentencing and enforcement. International organizations such as the United Nations Office on Drugs and Crime (UNODC, 2018) and the Council of Europe (SPACE, 2023) adopt different indicators depending on whether the focus is on policy evaluation, reintegration, or risk assessment.

Initial predictive efforts relied on static models using demographic and criminal history data (Burgess, 1928; Glueck and Glueck, 1950). Since the 1970s, attention shifted toward dynamic models that incorporate contextual and modifiable variables (Andrews and Bonta, 2010). In the last two decades, the development of machine learning (ML) has enabled more complex, yet less transparent, systems (Berk *et al*., 2018), underscoring the need for interpretable tools to guarantee transparency for legal actors and the public.

From a legal perspective, there is a clear divergence between common law and civil law systems. In Italy, Article 220 of the Code of Criminal Procedure prohibits expert assessments on personality or criminal propensity during trial, thereby restricting the use of AI tools to post-sentencing phases. Conversely, in common law systems such as the U.S., tools like COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) are routinely integrated into judicial decision-making in several States. Common law jurisdictions are increasingly adopting principles of fairness, accountability, and transparency (Citron and Pasquale, 2014), whereas civil law countries require a more cautious, rights-based approach (Floridi *et al.*, 2018). These institutional differences shape how ML systems are introduced into the justice sector and raise important questions about balancing predictive efficiency with the protection of individual rights.

The European Union's AI Act (Regulation EU 2024/1689), adopted on August 1, 2024, establishes a harmonized framework based on four levels of risk. Recidivism prediction tools categorized as "high risk" are subject to rigorous standards concerning transparency, reliability, fairness, and continuous oversight. While the AI Act seeks to facilitate the European digital single market, it also allows national authorities to tailor implementation to local legal traditions, ensuring respect for fundamental rights.

2. Data and Methods

This study relies on the COMPAS dataset published by a non-profit investigative journalism organization (Propublica, 2016), which contains information more than 7,000 individuals arrested in Broward County, Florida. It includes demographic data, criminal history, risk scores and a binary recidivism outcome within two years. We focused on key features: age, decile score, prior offenses, ethnicity (origins), gender, and offense type. Moderate linear correlations are observed between age and decile score (0.39), prior offenses and decile score (0.43), decile score and recidivism (0.35). The decile score, a composite risk indicator of recidivism correlates with reoffending as expected. The dataset showed issues of racial bias (Angwin *et al.,* 2016) and misclassification risks (Venkataraman, 2025). Instead of predicting individual risk scores, our aim is to investigate the relative importance of features and their contribution to the overall decision-making process. Despite ethical and legal constraints (Rudin *et al*., 2020), we argue that ML can elucidate complex crime patterns in data on convicts/detainees and guide targeted interventions.

Our approach combines a data-driven methodology using machine learning models like Random Forests and Neural Networks for predictive pattern detection, alongside interpretable models such as Logistic Regression to enhance transparency. Our goal is to critically evaluate the reliability and interpretability of these models.
All models were implemented in Python using specific packages: *sklearn.preprocessing, model_selection, sklearn.metrics, sklearn.linear_model, sklearn.ensemble, keras, sklearn.cluster, kmodes.prototypes, shap*.

Because the dataset contains features with different units of measure, variables were standardized to ensure comparability. In our preprocessing, we applied min-max normalization. Given potential heterogeneity in model performance, maximizing accuracy supports the identification of local explanations and the ranking of feature importance.

For the Logistic Regression model, we used default parameters from the "linear_model" package. The Random Forest model employed 100 trees from the "ensemble" package, with other hyperparameters left unchanged; increasing this number could improve performance but would increase model complexity. For example, a randomized search (“RandomizedSearchCV”) might identify a model with similar accuracy using 385 trees. The Neural Network is sequential, with three layers consisted of 16, 16, and 1 neuron(s); the first two layers used *Rectified Linear Unit* (ReLU) activation, while the final layer used *sigmoid function*. The optimizer was *Adam*, and the loss function was *binary_crossentropy*. The same test set served as validation.

We split the COMPAS dataset into training and test sets to evaluate model performance and to compute accuracy on the test data, in order to reduce the risk of overfitting. We focused on supervised learning approach for classification, where input features (labelled data) where mapped to discrete outcome categories. We used the test set to assess model capacity in the population (Hold-out procedure). We fitted our data with Random Forest, a non-parametric supervised learning method for non-linear relationships defined as an ensemble model, the result of aggregating a set of Decision Trees; it avoids local optima and correlation, Decision Trees limitations, through bagging and features selection. We also applied Neural networks (thousands of simple nonlinear models that work together, very difficult to interpret) and Logistic regression for comparing classification results, as they are representative in the trade-off between interpretability and accuracy (Wang *et al*., 2023).

With a binary target we analysed the classification report and the confusion matrix, focusing on the most important metrics: *accuracy, recall, precision, F1-score* and *specificity*. The confusion matrix is created by setting an assignment threshold. The decision threshold or cutoff point, generally set at 0.5, is a critical value used to convert the output of a classification model into a class prediction, since many algorithms return a probability score indicating the likelihood that an input belongs to the positive class, the threshold determines the cutoff point for classification. We obtained feature importance through the SHAP values related to each classifier, a powerful method for explaining the predictions of any machine learning model representing how much each feature contributes to a particular prediction for a given instance, given the expected baseline. We also analysed the classifiers behaviour in subgroups, to verify if accuracy is homogeneous in these subsets. To this end, we applied unsupervised classification to group the N units into clusters, to ensure that units within the same cluster displayed homogeneity.

We employed K-means to partition observations by minimizing within-cluster variance and maximizing separation between groups. To avoid instability associated with random initialization, we adopted the more robust K-means++ method. This approach improves convergence and clustering quality by selecting initial centroids with probabilities proportional to their squared distance from those already chosen, ensuring more diverse and informative starting points.

In order to select the optimal number of clusters K, we computed the Silhouette Score, which evaluates how well each unit fits its assigned cluster (cohesion) compared to the nearest alternative cluster (separation). The score ranges from –1 to 1, with higher values indicating better-defined clusters (1).

$$s(i) = b(i)-a(i) / \max\{b(i), a(i)\} \quad (1)$$

where a(i) is the average distance of each unit i to all the other units in its assigned cluster, and b(i) is the average distance of unit i to the units in the nearest cluster to which it was not assigned. The optimal K was selected by maximizing the silhouette score. A hierarchical solution with Ward's method (Ward, 1963) confronted at the same K produced similar values, supporting robustness, while K-means++ achieved slightly higher silhouette values, indicating improved stability and separation. By clustering similar instances, we identified subgroups differentiated by accuracy, and we analysed each cluster in terms of label imbalance, evaluation metrics, and SHAP-based feature importance (Lundberg and Lee, 2017). SHAP provides a linear explanation model in which binary variables attribute an effect to each feature; the sum of these contributions approximates the original model's output f(x) (2).

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i \, z'_i \quad (2)$$

where z'_i in $\{0,1\}$, M is the number of simplified input features, i ∈ R, f is the original prediction model to be explained and g the explanation model.

We compared subgroup metrics with the overall situation on the entire test set to underline differences and define each subgroup peculiarity: the idea is that the defined subgroup may highlight regions where the problem is more or less accurate.

3. Outcomes

The models show (Table1) accuracies of 0.68 (Logistic Regression), 0.67 (Neural Networks), and 0.63 (Random Forest), with the latter performing worst.

To assess overfitting, performance variability was evaluated via Cross-Validation (CV), which revealed a standard deviation (std) of 0.02 across all models, thus excluding significant overfitting.

Table 1 – *Classification report, main metrics for each class (negative, positive). Logistic Regression, Neural Networks, Random Forest models; COMPAS dataset.*

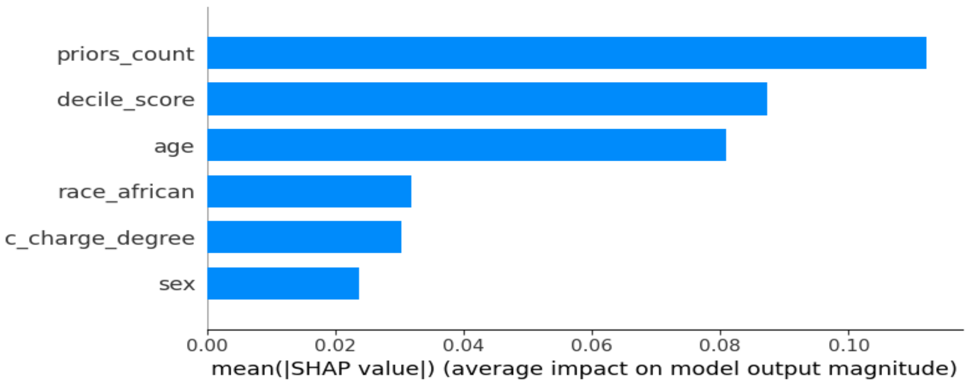
Models	Precision	Recall	F1-score	Support
Logistic Regression				
negative	0.69	0.78	0.73	1,189
positive	0.68	0.57	0.62	976
Accuracy			**0.68**	2,165
Macro avg	0.68	0.67	0.67	2,165
Weighted avg	0.68	0.68	0.68	2,165
Neural Networks				
negative	0.66	0.82	0.73	1,189
positive	0.69	0.50	0.58	976
Accuracy			**0.67**	2,165
Macro avg	0.68	0.66	0.66	2,165
Weighted avg	0.68	0.68	0.67	2,165
Random Forest				
negative	0.66	0.69	0.67	1,189
positive	0.60	0.56	0.58	976
Accuracy			**0.63**	2,165
Macro avg	0.63	0.63	0.63	2,165
Weighted avg	0.63	0.63	0.63	2,165

Note: negative (N) corresponds to non-recidivist; positive (P) corresponds to recidivist

All models show higher F1-scores for the negative class, as recall is greater for non-recidivists. Conversely, precision exceeds recall for the positive class, indicating better control of false positives than false negatives. Given the higher risk of false positives in criminal recidivism, prioritizing precision reduces the chance of mislabelling non-recidivists as recidivists. In the COMPAS tool's trade-off, avoiding missed recidivists is considered more acceptable than penalizing non-recidivists. The *Receiver Operating Characteristic-Area under the Curve (*ROC-AUC) values (0.74 for Logistic Regression and Neural Networks, 0.68 for Random Forest) support these findings; 0.74 is considered excellent for complex classification tasks. Selecting an optimal threshold requires analysing the confusion matrix to balance FP and TP where T and F denote true and false predicted values, respectively.

We obtained feature importance through the SHAP values (Figure 1), in general model agnostic while the actual computation for different model types often uses specific approximations or exact algorithms. The explainer for Random Forest is specifically designed for tree-based models and ensembles of trees (model specific) and can calculate SHAP values exactly or with high accuracy and efficiency, taking into account all possible paths. Due to the nature of tree models, Tree SHAP is particularly effective at capturing and quantifying interaction effects among features.

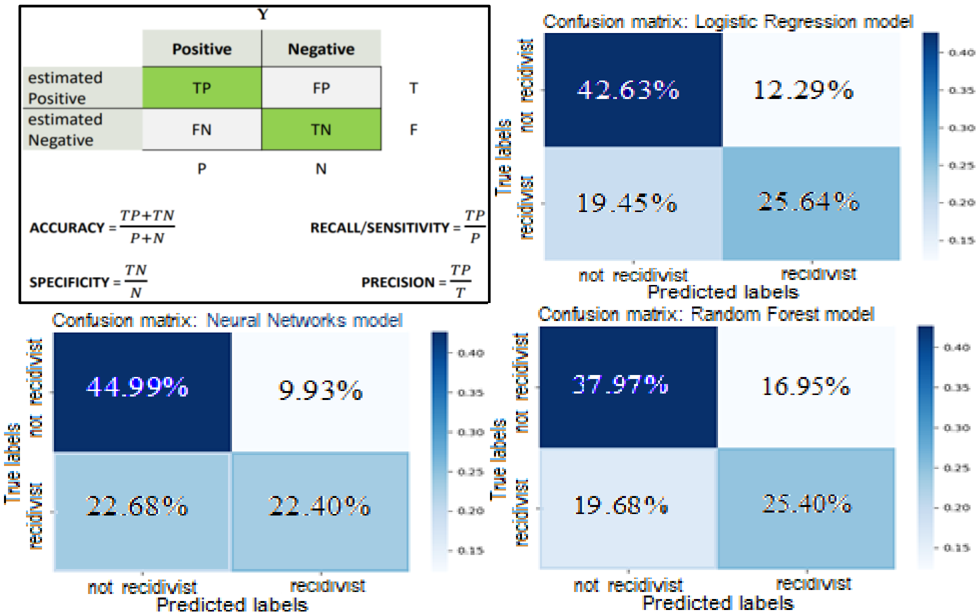
Figure 1 – *Feature importance based on the mean absolute SHAP values of the Random Forest model for the COMPAS dataset.*



The confusion matrix (Figure 2) shows consistent patterns across models, except for Random Forest. It has a higher proportion of FPs (16.95%) and a lower TNs (37.97%), indicating a greater risk of overestimate the recidivism (false positive errors).

In complex scenarios characterized by weak feature-target relationships or nearly indistinguishable classes, an AUC of 0.68 (as achieved by the Random Forest model) may still be acceptable. Further analysis is required: evaluation of feature selection, hyperparameters optimization and classifier behaviour across specific subsets.

Figure 2 – *Confusion matrix: comparison between actual (rows) and predicted values (columns). Logistic Regression, Neural Network and Random Forest models; COMPAS dataset.*



We grouped similar units, selecting a nine-cluster solution based on the peak Silhouette Score (k = 9, 0.57), a value considered reasonable (Figure 3).

Figure 3 – *Silhouette score and number of clusters for COMPAS dataset.*

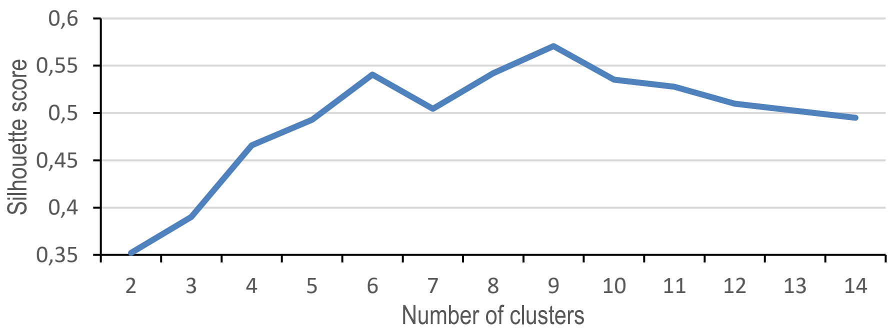
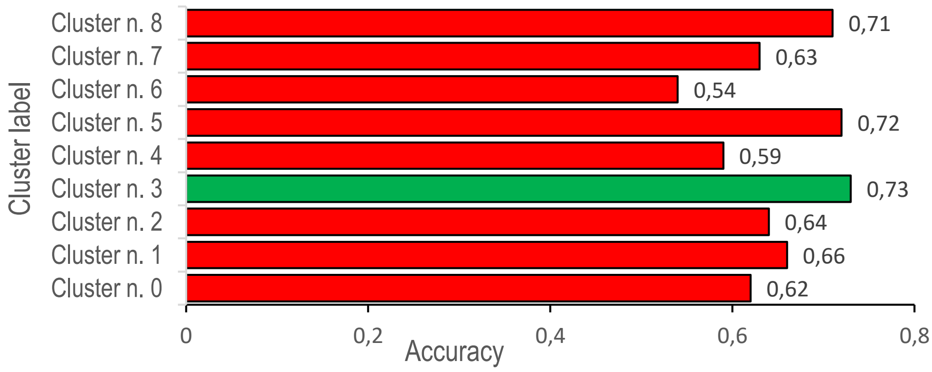


Figure 4 – *Accuracy of the Random Forest model calculated for each cluster identified using the K-Means method (k = 9) on the COMPAS dataset.*



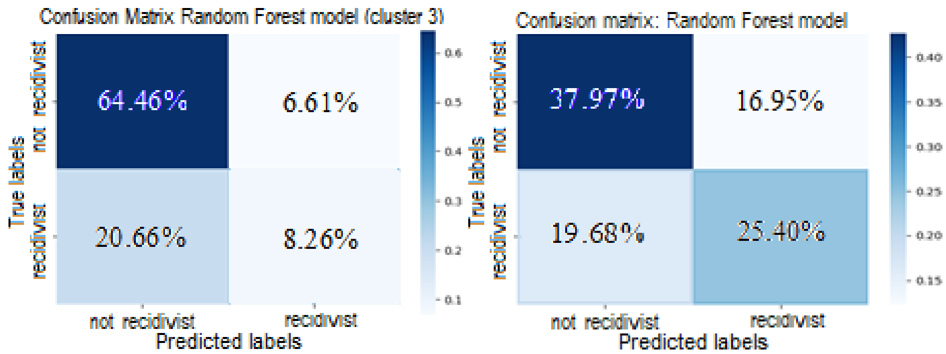
We compared K-means and K-prototypes for clustering numerical and binary features from the COMPAS dataset. The results showed no significant differences in classification accuracy across clusters – an essential focus of this study – nor in overall clustering quality (k = 8, silhouette score = 0.60). We selected K-means, which produced the widest accuracy range, and applied a Random Forest classifier to the resulting clusters (Figure 4). Cluster centroids and sizes were computed. Accuracy ranged from 0.54 (cluster 6) to 0.73 (cluster 3). Comparing the global metrics with those of cluster 3 with the highest accuracy, patterns emerged that helped identify factors potentially enhancing classification performance (Table 2).

Table 2 – *Classification report, main metrics for each class, cluster n.3 (left side) and whole dataset COMPAS (right side) based on the Random Forest model.*

	Random Forest (cluster n.3)					**Random Forest (whole dataset)**				
Random Forest	Preci-sion	Re-call	F1-score	Sup-port	Balan-ce	Preci-sion	Re-call	F1-score	Sup-port	Balan-ce
negative	0.76	0.91	0.83	86	0.71	0.66	0.69	0.67	1,189	0.55
positive	0.56	0.29	0.38	35	0.29	0.60	0.56	0.58	976	0.45
Accuracy			**0.73**	121				**0.63**	2,165	
Macro avg	0.66	0.60	0.60	121		0.63	0.63	0.63	2,165	
Weighted avg	0.70	0.73	0.70	121		0.63	0.63	0.63	2,165	

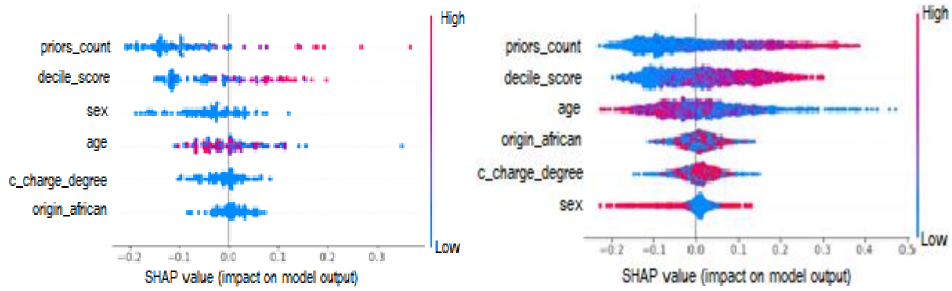
The confusion matrix for cluster 3 versus the whole dataset (Figure 5) indicates improved recall for the negative class but decreased recall for the positive class, reflecting class imbalance (>70% non-recidivists). Consequently, FPs decrease (6.61 vs. 16.95), TNs increase (64.46 vs. 37.97), while TPs decline (8.26 vs. 25.40).

Figure 5 – *Random Forest model's confusion matrix. Comparison between actual (rows) and predicted values (columns). Cluster n.3 (left) and the COMPAS dataset (right).*



SHAP values in the most accurate clusters identified key features influencing model behaviour. In cluster 3, gender – specifically female – was more important than age, while origin was not relevant; this cluster mainly include Caucasian women with minor offenses and low risk scores (Figure 6).

Figure 6 – *Random Forest model's SHAP values for cluster no. 3 (left) and COMPAS dataset (right).*



Cluster 8 (not shown), with similar profiles but of African origin, also demonstrated high accuracy (0.71). Conversely, cluster 6 (not shown): Caucasian men with serious offenses and high-risk scores, had the lowest accuracy (0.54), likely due to recidivism imbalance. Importantly, *ceteris paribus*, individuals of African origin (cluster 0) outperformed their Caucasian counterparts: accuracy increased from 0.54 to 0.62 (not shown).

4. Conclusions and future perspective

This study examines the predictive potential of ML models in assessing criminal recidivism risk, using the COMPAS dataset as a case study. Three supervised classification models – logistic regression, neural networks, and random forest – were tested and evaluated in terms of accuracy, interpretability, and subgroup variability. Although the overall predictive performance was moderate (maximum accuracy 0.68), logistic regression and neural networks outperformed random forest, particularly with respect to the area under the ROC curve (AUC) and F1-score.

This study shows that clustering reveals substantial accuracy heterogeneity in the COMPAS data: K-means++ identified stable subgroups with distinct performance levels, a structure confirmed by a Ward hierarchical solution and supported by slightly higher silhouette scores. SHAP values clarified why accuracy varies, with high-performing clusters showing consistent feature contributions and low-performing ones affected by heterogeneity and class imbalance. These results show that overall accuracy can hide substantial performance differences across subgroups. Cluster-based evaluation and interpretability methods therefore provide a more reliable and transparent assessment of model behaviour in criminal-risk prediction.

From a legal perspective, the study highlights the divergence between civil and common law systems: while tools like COMPAS are used in the U.S., In the Italian legal system, predictive assessments of individual risk are restricted by procedural norms (Article 220 of the Italian Code of Criminal Procedure), effectively limiting the use of AI tools to the post-sentencing phase. The EU AI Act (Regulation 2024/1689) classifies such tools as high-risk, subjecting them to strict requirements of transparency, reliability, and protection of fundamental rights

Despite their current limitations, ML models could serve a complementary but important role in public policy and correctional planning. Their use should not be confined to making decisions about alternative sanctions or sentencing. Instead, they may be strategically employed to identify individuals at higher risk of reoffending, thereby enabling targeted interventions and the efficient allocation of social support services. We suggest adjusting the classification threshold; in this context, prioritizing a slightly higher false positive rate may be ethically acceptable if it serves to initiate rehabilitative or generally supportive interventions for the individual – rather than punitive measures.

In conclusion, while machine learning holds promise for enhancing our understanding of recidivism patterns and informing data-driven policies, its application in the criminal justice domain must remain grounded in legal safeguards, methodological transparency, and a strong commitment to human rights. Future research should further explore the integration of richer contextual variables, investigate algorithmic fairness across demographic groups.

Acknowledgements

Deep gratitude is expressed to Prof. Donatella Firmani, for her exceptional guidance, unwavering support, and valuable feedback. We also express our gratitude to the RIEDS reviewers for their thoughtful and valuable feedback.

References

ANDREWS, D.A., BONTA, J. 2010. *The Psychology of Criminal Conduct*. New York: Routledge.
ANGWIN J., LARSON J, MATTU S, KIRCHNER L, 2016. *Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks*. ProPublica, May 23.
BARATTA A. 1998. *Criminologia critica e critica del diritto penale.* Napoli: Ed. Scientifiche italiane.
BERK, R., BURNETT, R., NELSON, C., SHEARER, C., BERGMAN, M. 2018. Fairness and Transparency in Risk Assessments, *Criminology & Public Policy*, Vol. 17, No. 4, pp. 857-872.
BURGESS E.W. 1928. Factors Determining Success or Failure on Parole. In *Illinois Committee on Indeterminate-Sentence Law and Parole*, Springfield: State Board of Parole, pp. 221–234.
CABRERA Á.A., EPPERSON W., HOHMAN F., KAHNG M., MORGENSTERN J., CHAU D.H. 2019. FAIRVIS: Visual Analytics for Discovering Intersectional Bias in Machine Learning. In *2019 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 46–56.
CITRON D.K., PASQUALE F. 2014. The Scored Society: Due Process for Automated Predictions, *Washington Law Review*, Vol. 89, No. 1, pp. 1–33.
FLORIDI L., COWLS J., BELLI L., TURILLI M., TADDEO M. 2018. AI4People – An Ethical Framework for a Good AI Society, Vol. 28, No. 4, pp. 689–707. https://doi.org/10.1007/s11023-018-9482-5
GLUECK S., GLUECK E. 1950. *Unraveling Juvenile Delinquency*. Cambridge: Harvard University Press.

LAPPI-SEPPÄLÄ T, 2003. Techniques in enhancing community based alternatives to incarceratio. In *Resource Material Series*, No. 61. Tokyo UNAFEI, pp. 61-87.
LUNDBERG S.M., LEE S.-I. 2017. A Unified Approach to Interpreting Model Predictions. In GUYON I. *et al.* (Eds.), *Advances in Neural Information Processing Systems,* No. 30, Curran Associates, Inc., pp. 4765–4774.
PROPUBLICA, 2016. Compas-Dataset https://github.com/propublica/Compas-analysis.
RUDIN C., WANG C., COKER B. 2020. The Age of Secrecy and Unfairness in Recidivism Prediction, *Harvard Data Science Review*, Vol. 2, No. 1 (31 March 2020). https://hdsr.mitpress.mit.edu/pub/7z10o269
SPACE II, 2023. Annual penal statistics: recidivism rates. Council of Europe.
UNODC, 2018. Manual on the measurement of juvenile justice indicators. United Nations Office on Drugs and Crime.
VENKATARAMAN N. 2025. *Quantitative Criminology Handbook*. (e-book) Educohack Press.
WANG C., HAN B., PATEL B., RUDIN C. 2023. In pursuit of interpretable, fair and accurate machine learning for criminal recidivism prediction, *Journal of Quantitative Criminology*, Vol. 39, No. 2.
WARD J.H., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, Vol. 58 No. 301, pp. 236-244.
WEATHERBURN D., FITZGERALD J., HUA J. 2003. Reducing Recidivism: Does Imprisonment Have Any Effect? *Crime and Justice Bulletin,* No. 58, NSW Bureau of Crime Statistics and Research, Sydney.

Flavia TAGLIAFIERRO, Sapienza University, flavia.tagliafierro@uniroma.it
Claudio CATERINO, Istat, caterino@istat.it

SURVEYING HOMELESSNESS IN ITALY: FRAMEWORK AND IN-DEPTH QUESTIONNAIRE OF THE NEW ISTAT NATIONAL SURVEY[1]

Eugenia De Rosa, Francesca Scambia

Abstract. Homelessness is an extreme form of social exclusion and a dynamic process difficult to measure. These are some of the methodological challenges addressed within the recent Istat national project on homelessness (2024-2026). In order to get information on the ‘visible’ homeless population at a certain time in a specific area, a Point in Time count of sheltered and unsheltered people experiencing homelessness is going to be conducted in 14 Italian Centers of the Metropolitan areas. In addition, a sample CAWI survey is planned shortly following the count on the same target population to collect more information on the profiles and condition of homeless people. Non-professional interviewers specifically trained, many of them volunteering on services for the homeless together with other citizens, will be in charge of data collection.
The aim of this paper is to discuss the main conceptual and operational issues related to the development of the in-depth questionnaire considering various aspects. Among them the definition and eligibility of the target population, balancing information on the main profiles and investigating reasons and dynamics of homelessness, use of services, denied rights and crucial needs.

1. Introduction

Combating poverty and social exclusion is one of the specific social policy goals of the EU and its Member States. The first-ever EU Anti-Poverty Strategy was announced in the 2024-2029 Political Guidelines with the aim ‘to help people to get access to the essential protection and services they need, along with addressing the root causes of poverty’ (2024)[2]. According to the EU Action Plan of the European Pillar of Social Rights by 2030 the EU should reduce the number of people at risk of poverty or social exclusion by at least 15 million. The AROPE rate[3]is the main indicator to monitor this target.

---

[1] This article is the joint work of the authors, however paragraphs 1, 3, 4.1, 4.2 and 4.3 are written by Eugenia De Rosa, paragraphs 2, 3.1, 4.4 and 5 are written by Francesca Scambia.
[2] On 18 July 2024, European Commission President Ursula von der Leyen, who was elected for a second mandate, presented to the European Parliament her Political Guidelines for the next European Commission 2024-2029.
[3] At risk of poverty or social exclusion (AROPE) corresponds to the sum of persons who are either at risk of poverty, or severely materially and socially deprived or living in a household with a very low

While progress in this regard is being achieved by a few European countries, incidence of poverty remains high and access to adequate and affordable housing is not a fundamental right for all people: in 2023 people at risk of poverty or social exclusion were more affected, with 8.5% of people aged 16 or older in the EU experiencing housing difficulties in the past, compared with 3.9% among those not at risk. But homelessness is not simply a result of a lack of affordable housing. The most common cause was reasons linked to family or relationships (30.0%). 26.5% of people in the EU reported that they had overcome housing difficulties by finding a job (Eurostat, 2024).

However, since a relevant portion of the roof/homeless have no civil registration, they usually escape the Population and Housing Census surveys as well as Household Surveys. For this, and other reasons, people experiencing homelessness are rarely included in the official statistics on poverty, falling in the area of the hard-to-reach populations. Definitional and methodological issues hind the collection of comparable and reliable data. People experiencing homelessness is a dynamic population with different forms and ‘paths through homelessness’ (transitional or short-term, episodic and chronic). They may be more or less ‘invisible’ and difficult to reach (e.g., undocumented migrants, EU migrants with no claim to social benefits, people with prolonged stay in institutions, or people dwelling with friends or families). Though it can be a reality in small municipalities, homelessness is especially a urban phenomenon across Europe.

Data varies not only depending on the definition of homeless but also whether what is measured is the stock, the flow, the prevalence, or the incidence of homelessness. Point-in-time (PIT) street or service-based counts and/or point in time surveys, presenting a ‘snapshot’ of homelessness at a single time and place, is one of the methods most often used to measure urban homelessness. Other methods are longitudinal studies or use of administrative data of social-service institutions based on information on service users (Braga *et al*., 2024: Busch-Geertsema *et al*., 2014; EC and Develtere, 2022; Geyer *et al*., 2021; OECD, 2024; Schnell and Musil, 2024).

The lack of a comprehensive and harmonise measurement framework, both at EU and national level, does not allow for adequate monitoring of homelessness. Tackling extreme poverty remains mainly the responsibility of EU countries' governments acting in complex national social settings.

2. Homelessness national strategy and data

The effects of the pandemic by Covid 19, which deeply affected people experiencing homelessness, gave a boost to policies aimed at combating extreme

work intensity - (quasi-)jobless households. People are included only once even if they are in more than one of the situations mentioned above.

poverty in Italy. The possibility of investing to eradicate homelessness was also based on the funds provided by the PNRR and the Poverty fund. The Ministry of Labour drafted the "Piano nazionale contro la povertà 2021-2027" aiming at providing more stable solutions for people experiencing homelessness. Among the main issues included in the Plan: emergency response, access to registered residence with related support services, allocation of funding for Housing Led and Housing First, postal address and Centers to combat poverty.

Another reference framework is provided by the "Piano Nazionale degli interventi e dei servizi sociali 2021-2023". On the specific issue of extreme marginalization, the "Linee di Indirizzo per il contrasto alla grave emarginazione adulta in Italia" by the Ministry of Labour and Welfare published in 2023 provided a deep analysis and updating on this phenomenon after the previous Guidelines published in 2015. They also offer some suggestions on how to develop services addressed to people experiencing homelessness. The effort in providing guidelines and responses to extreme poverty and homelessness requires a deeper knowledge of dimension and features that characterise this group of people.

Indirect sampling and a Point in Time approach with a street count are to date the two main approaches used by the Italian National Statistical Institute to carry out surveys on people in extreme poverty. The Istat national surveys on soup kitchens and overnight shelters' services (2011 and 2014) provided a first picture of homelessness. The survey estimated the number of people who sleep in the street (or in overnight shelters) using the indirect sampling detecting the homeless population, or rather a relevant part of it, in the places where the people go and receive the services they need (centers where soup kitchens and overnight shelters' services are provided). However, a large group of people experiencing homelessness do not reach-out services, and they keep therefore outside the observation field. The first Point in Time experience was carried out by Istat more recently with Roma Capitale in 2023-2024. The aim was to count the number of homeless people on the streets and in overnight shelters, and to provide a picture of their profiles (De Rosa *et al*., 2025a, 2025b). Another (partial) source of data is the Istat Permanent Population Census, as the result of the revision of population registers carried out by the municipalities. They cover people who are registered in the population register at a virtual address. Though interesting, however these data do not provide a picture of homelessness for they include also people who are housed but use a virtual address for other specific need (e.g., nomads, circus workers). At the same time data on undocumented homeless people are excluded and not counted by means of registers.

3. The new Istat National Project on homelessness (2024-2026)

In order to take into account the complexity of homelessness, for the first time Istat is going to undertake a project on this topic within the framework of the

National Census (Chieppa *et al.*, 2025). The project, which is planned to go in the field by the beginning of 2026, covers 14 Centres of Metropolitan areas.

Istat led an Inter-institutional Technical Scientific Committee to manage the project. In addition to Istat representatives, the Committee includes academics, experts appointed by 14 Italian metropolitan city municipalities and Fio.PSD (Italian Federation of Organisations for Homeless People). The latter is involved in the project as external body in charge of the organization, logistics and recruitment of interviewers.

The Istat national project intends to address homelessness in Italy by undertaking a comprehensive initiative. It aims to quantify the extension of homelessness by means of a PIT count of both sheltered and unsheltered individuals in major Italian cities. Additionally, the project plans to gain deeper insights by conducting a sample survey of the same population following the initial count. In both steps information will be collected by non-professional interviewers and grassroots organizations. A short electronic questionnaire (a counting sheet) was developed for the first step and an in-depth electronic questionnaire for the second step.

City counting is important not only to know the extent of the phenomena but also to mobilise local actors and local public opinion. Beyond simply counting, it is essential to examine the profiles of individuals experiencing homelessness and delve into the underlying causes and reasons for their situations. This comprehensive approach is vital for accurately assessing the scope of the social issue and for developing more sophisticated and diverse policies to address the numerous challenges faced by the homeless population.

Municipalities and organizations involved have direct knowledge of the phenomenon. They are interested in contributing to have data on the number and characteristics of the homeless people living on their territory. The comparability of data among municipalities is also relevant for policy makers.

3.1. The target population of the new survey

The homeless condition includes a range of different situations of accommodation or lack of it, that varies from country to country and also among the Italian municipalities. To this aim FEANTSA (European Federation of National Organisations working with the Homeless) drafted the Ethos classification (European Typology of Homelessness and Housing Exclusion) which conceptualises different conditions of housing exclusion providing also further descriptions for each category. FEANTSA also drafted the Ethos light version of the classification which is mainly thought for statistical purposes. Though very helpful these classifications, however they do not solve all the definition needs.

For the national survey Istat broadly identified the target population in Ethos light 1 (“People Living in the streets or public spaces without a shelter that can be defined as living quarters") and Ethos light 2 ("People with no place of usual residence who move frequently between various types of accommodation") with some specific choices. The first group includes unsheltered people who live in tents or cars, but excludes people living in caravans, mobile houses or shanties. As for the second group that includes sheltered homeless people, a definition of shelters was required because homeless people are hosted in a variety of accommodation services, which provide different services not only among different municipalities, but also within the same one. To this purpose Istat included in the target exclusively the low-threshold emergency and temporary overnight accommodation facilities. The low-threshold is intended as no specific requirements for access, but the condition of being roofless and an informal relationship between guests and staff; in addition access does not require participation in a therapeutic or rehabilitation programme, only compliance with the rules of cohabitation. Shelters or refuges for women who are victims of violence, as well as specific accommodation programmes for migrants are excluded from the target. On the other hand, both private and public-financed shelters for homeless are included whatever time span they cover (h9-h24).

4. Investigating homelessness: conceptual framework and in-depth questionnaire

4.1. Development and pre-testing of the questionnaire

The main aim of the in-depth survey, to be conducted among sheltered and unsheltered people experiencing homelessness, is to provide a portrait of people who are homeless in the main cities in Italy and investigating what drives people into homelessness and the problems they encounter in their daily lives. To this end, an electronic structured questionnaire has been designed with predominantly closed questions and some differentiated questions for the two main target populations (sheltered and unsheltered people). A paper version has also been prepared to deal with emergency situations.

Specifically trained non-professional interviewers will conduct data collection among a highly vulnerable target population. These and other aspects have been taken into account in designing the questionnaire and ad hoc solutions have been adopted.

Firstly, the questionnaire cannot be very long, with wording that is simple and friendly. Homeless people are a heterogeneous population (also in terms of country of origin and spoken language) not always visible; some individuals may experience mental health problems, alcoholism or other addictions.. Therefore, some difficulties may arise in defining the eligibility and involving homeless people in the research. At the same time greater attention has been paid to how to introduce the survey in suitable manner for such hard-to-reach population, and deal with privacy issues and sensitive questions. Finally, technical aspects are related to the platform used (Lime Survey) and the survey device that will be used (interviewers' personal cellular phones or tablet). All the above mentioned aspects have an impact on the quality of data.

A first draft of the questionnaire was discussed with experts (De Benedetti Foundation, Fio.PSD and researcher of the EU project "European Homelessness Count" led by the University of Leuven and, for Italy, by the University of Catania). Harmonization, at a European level, and attention to national specificities are key points.

Pre-testing interviews with eight homeless people living in Rome in public space or external space and people living in the overnight shelters (with different profiles by age, sex, citizenship…) were conducted. The main aspects considered were how to approach such a hard-to-reach population (e.g. question wording, language, technique, how respondents understand terms and reference period, test modalities and detect sensitive questions).

Pre-testing revealed the willingness of homeless people to tell their story but at the same time the sensitivity of certain themes and traumatic events. It also highlighted differences between Italians and foreigners and the difficulty for some people to sustain a long interview, so a more concise version of the questionnaire was drafted for specific situations.

Another aspect that emerged was that the interviewed persons do not always follow a logical order in their responses and do not consistently adhere to the questionnaire's structure; they sometimes report confused or unreliable answers. For this reason, mayor attention was paid to including both the direct collection of information from the homeless person and assessments and observations made by the interviewer, which concern, on one hand, the situation and the homeless, and on the other hand, the interview and the quality of the information gathered during the interview. For certain characteristics, both perception-based and interviewer assessment-based data are collected.

Finally, the core section of the interview aimed at reconstructing the housing and life history of the homeless person appeared very challenging. This led to redefining this part of the questionnaire and to include a section, at the end of the interview, in which interviewer and observer reconstruct the main stages of the story.

4.2. The final structure of the in-depth questionnaire

The in-depth questionnaire is structured into three parts: the first part is filled out exclusively by the interviewer with no interaction with the person encountered. It must be completed for each contact, even if the person later refuses the interview or is not in the condition to do it. A specific question distinguished the street from night shelter detection. In addition, this part detects some information useful in defining the eligibility. The second part of the questionnaire (interview) required interaction with the person. It is aimed at defining eligibility and collecting in depth information on the person and their story. The third and final part is filled in by the interviewer alone on the basis of personal assessment regarding aspects that concern both the condition of the person met and the progression of the interview, also to assess the quality of the information collected.

Eligibility and consent are two fundamental steps from which certain paths derive. A common eligibility requirement for street detection and night shelter facilities is to be 18 and over, and this is firstly assessed by observation. In the night shelters, the only fact of being a guest in the facility is considered evidence that the person is part of the target population. For the survey on people living in public space or external space, on the basis of objective signs (such as being placed in a certain place with blankets, or being on the move carrying bulky objects, etc.), the team assessed whether to consider the person encountered as homeless of legal age and start filling out the first part of questionnaire. Three questions in the interview complete the set of information to test eligibility. They are: being/not being already interviewed in the same evening as part of the same research, where the person is going to sleep that night and if they have slept in the street in the last seven days. The interview ends for those who answer they are going to sleep in a house (their own or a friends/relatives/third parties' house), in a caravan, or camper van, or prefer not to answer, as they are not included in the target population.

With regards to informed consent, the interviewer, after briefly presenting the aim of the survey, collects the explicit consent of a person to participate in the short interview (and ask some questions to assess their eligibility). In the event of a denial, the interviewer moves to the third part of the questionnaire where some additional information is collected including the reasons for the refusal. The same information based on the interviewer' assessment has to be indicated in case of interruption of the interview.

4.3. Main themes and indicators

The in-depth questionnaire has different objectives:

i) identifying the main profiles of homeless people and comparing the two main target populations of the study,
ii) investigating the housing trajectories and the complex interaction of causes of homelessness. Individual factors (low education, lack of job skills, mental health issues, substance use, family violence or instability, relationship breakups), interact with life events and structural factors such as barriers to education, unfavourable housing and labour market or discrimination based

on access to accommodation and goods and services, lack of support for immigrants and refugees, aging out of foster care and leaving prison;

iii) knowing obstacles encountered in their daily life (in meeting primary needs, when accessing services,), social relations, and needs of a group of people who often face stigma and prejudice ‘because housing status is perceived as somewhat under an individual’s control, whereby the homeless are often considered to be responsible for their lack of adequate housing’ (Johnstone *et al.,* 2025).

The final version of the questionnaire, individual and anonymous, is unique for street and night shelter detection. In addition to a consistent set of common core variables some differentiated paths and questions have been provided. The different sections of the questionnaire respond to the need to capture the dimensions considered conceptually relevant for the objectives of the research. The main sections are:

a) socio-demographic information with some specific questions for migrant people. The aim is to know the emerging groups of homeless;

b) homeless condition and housing trajectories. It is aimed to detect, as a proxy, chronic (for more than one year), intermittent (in and out of homelessness repeatedly) and transitional homelessness (once or twice for a relatively short period of time after a major life change or catastrophic event), as well as the main patterns and causes of homelessness. This section also includes the dimension of harassment and violence, by means of indicators of self-reported experiences of being robbed, disturbed, or assaulted, and the relationship with law enforcement bodies. At the end of section b, the interviewer is asked to assess whether the interviewed person is in a condition to continue a detailed interview. If not, the interview continues with only a selected set of core variables covering the other dimensions of the questionnaire;

c) health condition and the use of health services. The health status is detected as well as the use of health services (distinguishing between public and private sectors) and the main obstacles to the access. Very specific situations are investigated (e.g., hospitalization, attending a doctor, or a mental health centre) as well as having the health card (STP/ENI) that guarantee a series of basic health services for foreigners;

d) services for the homeless, social services and documents. This section focuses on the use of services dedicated to homeless people, such as the distribution of food parcels, clothing, showers and/or personal hygiene services, canteens, day shelters and other social services (e.g., municipality services). Additionally it identifies the main obstacles faced in accessing these services (e.g., ‘I don't know them’, ‘There are no services nearby’, ‘I have difficulty getting there/reaching them’, ‘There are too many people’, ‘I don't know where to leave my parcels’, ‘I was poorly received’). A specific question is aimed at addressing a very extreme situation (‘In the last year, have you ever had trouble in finding enough food?’). Another relevant dimension included is whether the person is a registered resident in an Italian municipality and the reasons why they are not registered. This helps indirectly examine the ability to access various rights. In addition, the possession of not-expired documents and the use of a series of services, such as bank account, postepay card, address for mail collection, phone number, internet connection, are also explored for their role in facilitating the exercise of certain rights.

e) work and sources of income (as pensions, public benefits, money from family members, and income received from employment) and use of job centers or participating in a training course or a job placement project;

f) family and social relations including romantic relationship, having children, friends, contact with family members, reference persons in situations of need, etc…..

g) typical day. This final section aims at understanding whether the person is engaged in activities that make them feel good, going beyond primary needs, such as reading, listening to music, playing cards, traveling. An open field to describe how the person usually spends their day is also provided, with the aim to understand how people experiencing homelessness construct their self-identity that goes beyond the stigmatised one.

4.4. Observation and evaluation parts

This part, as experienced in other surveys on the homeless people, is crucial to provide proxy information which enable to have a more complete picture on the situation and profile of the interviewed person. At the beginning, before interacting with the person, the interviewer is required to write some information and identification details concerning the area and time of the interview. In addition to that they write information on the location where the encounter takes place and to this aim some items are provided. They have to write weather an arranged place to sleep is prepared or not. In this same section they also choose between the interview to sheltered and unsheltered people.

After the interview is concluded or if the person does not complete the entire questionnaire, a designated section for the interviewer was created to offer an assessment and remarks on the interview. The first question is: Was the interview interrupted before the questionnaire was completed?

Considering the survey takes place in the street for a large part and involves a hard-to-reach population, the possibility it is interrupted is quite high. This situation can happen due to reasons related to the person' difficulties in answering (language, questions' understanding, physical limits), the context (other disturbing people arrive). Therefore, the second question for the interviewer is: Did the person demonstrate to understand Italian language? After that the language in which the interview was conducted is registered. Two questions follow to collect the interviewer’s observation weather the person was altered or not (i.e., under the effects of alcohol, drugs, etc.) there is also a possibility of answering that the interviewer is not able to assess alteration. The following question is: Do you believe that the person you met had mental health issues? Then a question on health problems (excluding mental health issues): Do you think the person you met has one or more health problems? If the interviewer believes that the person has health problems, the question is: Does the health problem limit the person's movement/mobility? Then questions focus on the obstacles the person met in answering questions: In understanding questions, in remembering events, In placing events in time, In the high number of questions, In the language in which the interview was conducted (poor knowledge), Due to the person's difficulty in speaking (e.g., dental problems), Other difficulties (Specify). Finally, the interviewer is asked to report if and which of the recorded answers provided by the interviewed person appeared not to be true. A final blank field is left for remarks by the interviewer, in the introduction of the blank field some examples are provided to help the interviewer in filling out the field with useful information (make a note, for example, if the person is with the partner, with fellow countrymen, accompanied by a pet, etc.).

This closing section of the questionnaire to be filled-out by the interviewer alone completes the information and provides a clearer context in which the interview took place. It provides a possibility of using also those interviews which were not completed by means of the reasons why. The aim is also to detect more general information that cannot be directly asked to people experiencing homelessness.

5. Concluding remarks: suggestions for training

This type of survey is quite peculiar because it targets a‘hard-to- reach’ population and is conducted by y non-professional interviewers. These two factors are crucial to consider when planning a successful training to achieve the desired outcomes

People experiencing homelessness are very different among themselves, but they share a condition of vulnerability and an experience of distress. For instance, the story of women experiencing homelessness is often characterised by violence or by the separation from their children due to different causes. This condition is to be communicated during training, emphasising the need of approaching people with a special care and attention especially on sensitive topics. Also, the way of speaking has to be very clear, in order to be understood, considering also the different mother tongues. The approach should be very calm and ready to listen also to topics that may appear not directly relevant to the purpose of the interview. At the same time, it is important not to miss information that may emerge in the conversation. To this aim the role of the observer is very relevant. Interviewers are to be trained on how to open a conversation and ask all the questions, but they should be aware that the sequence of the answers may not be perfectly followed, though the information maybe anyway provided by the homeless person while speaking of another topic. Therefore, interviewers should deeply know the questionnaire in order to have in advance a clear knowledge of the information to be collected.

On the other side the interviewers are chosen among associations and organisations used to deal with extreme marginalization. This characteristic may facilitate the approach and the interaction. However, during that night/evening they are not called to carry out their usual activity with the homeless people. This means that training should focus on the specific task as interviewers they are required to perform.

Due to the environmental and personal conditions of the respondent, especially when on the street, it may happen that the interview is suddenly interrupted, in these cases the interviewers have to be well trained on what to do to close the interview saving all the already collected data.

Training plays a crucial role and should consider all these aspects, and also use all the lessons learned during pre-test of the questionnaire considering the specific feature of this survey.

References

BRAGA M., CORNO L., MONTI, P. 2025. *Homelessness: Data, Prevalence and Features*. Milano: EGEA Spa - Bocconi University Press.

BUSCH-GEERTSEMA, V., BENJAMINSEN, L., HRAST, M., AND PLEACE, N. 2014. *Extent and Profile of Homelessness in European Member States – A Statistical Update*. Brussels: FEANTSA.

CHIEPPA A., DE ROSA E., DI LEO F., INGLESE F., SCAMBIA F. 2025. Disegno dell’indagine sulle persone senza dimora nei 14 comuni centro delle aree metropolitane. *Obiettivi, metodologia, strumenti di rilevazione, governance e modello organizzativo, Report interno*. Roma: Istat.

DE ROSA, E., IOANNONI, V., AND F. SCAMBIA. 2025A. Surveying homelessness in Rome, the way to the first Notte della Solidarietà, *RIEDS*, LXXIX-4 pp.43-54.

DE ROSA, E., DI LEO, AND F. SCAMBIA. 2025B. Homelessness in Rome 2024: The PIT count. *European Journal of Homelessness. Research Note*. Vol. 19, No. 1. pp. 205-223.

EC (EUROPEAN COMMISSION: DIRECTORATE-GENERAL FOR EMPLOYMENT, SOCIAL AFFAIRS AND INCLUSION) AND DEVELTERE,

P. 2022. *Data Collection Systems and Homelessness in the EU – An Overview*. Brussels: Publications Office of the European Union. Available at: https://data.europa.eu/doi/10.2767/989826

EUROSTAT. 2024. *Living conditions in Europe - housing and renting difficulties. Statistics Explained*. Available at: Living conditions in Europe - housing and renting difficulties - Statistics Explained - Eurostat

GEYER, L., SIMMONS, C., AND ZÓLYOMI, E. 2021. *Homelessness: Examining Data and Data Collection Methods*, European Centre Reports, Policy Brief 2021/22. Vienna: European Centre for Social Welfare Policy and Research.

MINISTERO DEL LAVORO E DELLE POLITICHE SOCIALI. 2021. Piano nazionale degli interventi e dei servizi sociali. Available at: https://www.lavoro.gov.it/priorita/Documents/Piano-Nazionale-degli-Interventi-e-dei-Servizi-Sociali-2021-2023.pdf

MINISTERO DEL LAVORO E DELLE POLITICHE SOCIALI. 2023. Linee di indirizzo per il contrasto alla grave emarginazione adulta. Available at: https://poninclusione.lavoro.gov.it/sites/default/files/2023-02/Linee-di-indirizzo-per-il-contrasto-alla-grave-emarginazione-adulta.pdf

OECD. 2024. *OECD Toolkit to Combat Homelessness*. Paris: OECD Publishing.

SCHNELL P., MUSIL R. 2024. Strategies for Measuring Homelessness in a Federally Organised State – The Case of Austria, *European Journal of Homelessness,* Vol. 18, No. 2, pp.121-148.

Eugenia DE ROSA, Istat, eugenia.derosa@istat.it
Francesca SCAMBIA, Istat, scambia@istat.it

A SURVEY-BASED IMPACT EVALUATION OF NRRP ON ITALIAN MUNICIPALITIES

Clio Ciaschini, Barbara Ermini, Luca Salvati, Gianluigi Salvucci

Abstract. This contribution presents results from the project *A Survey-Based Impact Evaluation of NRRP on Italian Municipalities*, which aims to assess the effects of the National Recovery and Resilience Plan (NRRP) on local public administration, territorial development, and citizen well-being. The research focuses on municipalities, as a significant share of NRRP investment lines directly involves their participation. Municipalities act as implementing authorities for many interventions funded by the NRRP, both in terms of territorial initiatives and the modernization of public administration. Data were collected through a national survey targeting Italian municipalities designated as NRRP implementing entities. The questionnaire included both subjective assessments and quantitative data about expected changes with and without NRRP support. The paper illustrates results as to territorial sustainability by highlighting the evidence emerging across geographical macro-areas. Findings aim to inform policymakers on the territorial effectiveness of the NRRP and to suggest improvements for the design and implementation of future policy instruments fostering inclusive growth.

1. Introduction

COVID-19 pandemic significantly worsened Italy's already fragile economic and social landscape, underlining its structural vulnerabilities as well as its sluggish productivity growth. Between 1999 and 2019, Italy’s GDP increased by only 7.9%, markedly less than Germany’s 30.2%, France’s 32.4%, and Spain’s 43.6%. Social issues have also intensified, as poverty rose from 3.3% in 2005 to 7.7% in 2019 and reached 9.4% in 2020, disproportionately impacting youth and women (Presidenza del Consiglio dei Ministri, 2025). Persistent regional disparities remain, notably between the more developed Northern part of the country and the less developed Southern regions, posing significant challenges to the catching up process of the latter to the former ones. Moreover, environmental and geomorphological issues brought by earthquakes, droughts, and floods further threaten the country's resilience.

In response to these challenges, Italy’s National Plan of Recovery and Resilience (NRRP), a key component of the European NextGenerationEU (NGEU) Program, designs reforms and investments to boost competitiveness, promote green and digital transitions, inclusivity, and foster regional cohesion, allocating around 40% of resources to the South. The plan identifies six priority areas: digital innovation, ecological transition, sustainable infrastructure, education and research, social inclusion, and healthcare, with a particular focus on gender equality and youth employment (Italia Domani, 2025). NRRP has been approved in July 2021 and relies on approximately 194.4 billion euro (including loans and grants), further supplemented by a national fund of 30.6 billion euro. It originally established 190 interventions—132 investments and 58 reforms—recently including additional missions like RePower EU. A significant amount (about 36%) of the resources are managed at regional and local levels, including municipalities and metropolitan cities, in charge of designing reforms and projects supporting the territorial recovery (IFEL, 2024; ANCI, 2023; Sacchi and Rubino, 2023).

The project entitled "A survey-based Impact Evaluation of NRRP on Italian municipalities" aims at monitoring the action of local authorities to assess the effectiveness of their policy designs and NRRP investments to modernize administration, reduce inequalities, improve infrastructure, and enhance citizens' quality of life. It also attempts to identify potential challenges and delays faced by municipalities in project implementation. Due to the unavailability of real outcome data, the study relies on survey-based opinions to evaluate potential impacts, following an evidence-based approach enabling more effective evaluations once actual results can be measured.

The main objective of the work is that of quantifying the extent at which NRRP has affected the development paths of Italian municipalities. In particular, this paper aims to assess whether the plan has fostered territorial development and sustainability, taking into account the specific critical issues of each municipality, including those related to geographical factors.

2. The construction of database

The database relies on the answers of 376 respondents to an original survey addressed to the entire population of Italian municipalities (nearly 7,900 units in 2024). Since participation was voluntary, and low response rates among local governments are well documented— a compelling issue in the era of widespread online surveys (Krause *et al.,* 2024)—we extended participation to all municipalities rather than relying solely on the initially planned stratified sample (Enticott, 2003; Ermini *et al.,* forthcoming). The elementary statistical unit is the individual municipality. The questionnaires focus on NRRP projects activated by municipalities in their role as implementing entities. Municipalities act as implementing authorities for many interventions funded by the NRRP, both in terms of territorial initiatives and the modernization of public administration.[1] As such,

[1] The questionnaire is available at: https://nrrpsurvey.econ.univpm.it

they are key actors in the implementation process and play a pivotal role in shaping the plan's local effectiveness. Each municipality was invited to participate to the survey by designating one or more officers directly responsible for NRRP-related functions—typically mayors, deputy mayors, administrative managers, or heads of technical offices. These actors constitute the institutional figures best positioned to report on the design, management, and local implementation of NRRP interventions (Moore *et al*., 2017). Data collection has been developed through online interviews using a CAWI (Computer-Aided Web Interview) system, further complemented with telephone interviews, reminders, and individual email follow-ups (Moore *et al.*, 2017), carried out by a TelePerformance's operational consultancy, specialized in data collection and analysis.

Table 1 – *The NRRP's Missions and Components.*

Mission	Id	Components
M1	M1C1	Digitalization, innovation, and Security of the Public Administration
M1	M1C2	Digitalization, innovation, and competitiveness in the production system
M1	M1C3	Tourism and culture 4.0
M2	M2C1	Green Firms and Circular Economy
M2	M2C2	Energy Transition and Sustainable Local Mobility
M2	M2C3	Energy Efficiency and Building Requalification
M2	M2C4	Protection and Enhancement of the Territory and Water Resources
M3	M3C1	High Speed and Mantainance of the Road Network
M3	M3C1	Intermodality and integrated logistics
M4	M4C1	Enhancement of Teaching and the Right to Study
M4	M4C2	From research to firms
M5	M5C1	Labour Policies
M5	M5C2	Social infrastructures, families, communities and the third sector
M5	M5C3	Special territorial cohesion interventions
M6	M6C1	Proximity assistance and telemedicine
M6	M6C2	Innovation, research and digitalisation of healthcare
M7	M7C1	RePower EU

Source: Italia Domani (2025).

The survey reports both qualitative and quantitative information on municipal activities connected to digitalization, territorial resource enhancement, and the promotion of human well-being. Specifically, it investigates aspects connected to digital public services and territorial investments—including waste separation, renewable energy, and energy efficiency —as well as tourism, culture, and social inclusion indicators like nursery and school places, green spaces, and parks. A detailed description of the missions in the NRRP has been provided in Table 1. Data, including opinion and quantities evaluation, are collected on the current and expected levels of these indicators for different key timeframes, including the ongoing year of implementation (2023) and projections for the end of the NRRP (2026).

Table 2 – *Respondents by macro-area (Panel A) and by population class (Panel B).*

		Panel A	
Macroarea	Respondents	Councils in the Sample (%)	Councils in Italy (%)
North - West	115	30.58	37.87%
North - East	78	20.74	17.57%
South	70	18.61	22.58%
Center	67	17.81	12.26%
Islands	46	12.23	9.73%
Total	376	100	100.00%
		Panel B	
Population - class	Respondents	Councils in the Sample (%)	Councils in Italy (%)
0-3000	187	49.73	56%
3000-5000	38	10.11	13%
5000-10000	52	13.83	15%
10000-20000	46	12.23	9%
over 20000	53	14.10	5%
Total	376	100	100%

Source: Our elaboration on Italian Municipalities NRRP Survey. Statistics for Italy are referred to 2024.

3. Description of respondents

Table 2 describes the distribution of respondents classified according to the corresponding macro-area and the size of population. Overall, the sample is reasonably aligned with the national municipal structure, even if respondents from the central macro-area and from larger municipalities are slightly more represented. In particular, looking at the macro – area, most of respondents, (30.5%), are located in the North – Western part of the country, while the number of respondents in the other continental areas are almost equal (around 18% of the total respondents). The lowest share of respondents, about 12%, belongs to the insular areas. The right side of the table classifies the respondents according to the size of its population. In the sample, a prevalence of respondents is included in the municipalities within the population class 0 – 3000, i.e., small municipalities, while, the other population classes include, on average, around 12% of respondents.

Almost all units of the sample (372 su 376), except for 4 north – western municipalities, apply to the NRRP at least with one project, mostly as exclusive implementing entity, (73% - 267 municipalities), with a minor share as executing entity or partner or in both roles (24%). Geographically, exclusive implementing entities are mostly located in North – Western areas, even if significant also in the South, North – East and Central areas, remarking a homogeneous distribution throughout the national territory. Municipalities acting as implementing and executing entities seems to be more common in the Northwestern, Northeastern and Central areas, thus suggesting a more articulated structure of partnerships. With reference to the outcomes of the calls, 46% of municipalities (174) declares that all presented projects have been approved and funded, while 50% (188 municipalities) reported a mixed outcome, resulting not all approved projects. Only the 1% (4 municipalities) reports received the approval of none of the projects. As highlighted in Table 3, most municipalities with approved projects, acted as exclusive implementing entity (about 71%), while the roles of mixed entity have been covered by 24%. The participation with exclusive roles of executing and partner is very low (lower than 2%). At a territorial level, the role of implementing subject dominates all areas, exhibiting higher values in the North – West and Southern areas, while the presence of mixed roles is more frequent in North – Western, North - Eastern and Central zones, delineating a wider variety of planning governance.

Table 3 – *Role of municipalities within the NRRP projects.*

Macroarea	Exclusively implementing entity	Exclusively executing entity/partner	Both implementing and executing entity/partner	n.d.	Total macroarea
North -West	87	1	21	6	115
North-East	51	2	23	2	78
Center	43	3	21	0	67
South	55	1	13	1	70
Islands	31	0	14	1	46
Italy	267	7	92	10	376

Source: Our elaboration on Italian Municipalities NRRP Survey.

4. A general assessment of territorial development and sustainability expectation over years 2021-2026.

This section illustrates outcomes of survey questions about the expected impact over the period 2021-2026 of the NRRP on territorial development and sustainability, which is analyzed along three dimensions: sustainable energy, energy efficiency, and renewable capacity. Respondents were also asked what development they would expect in the relevant area in the absence of NRRP funding. The answers reflect two distinct perspectives: a counterfactual scenario — that is, the expected outcomes in the absence of funding — as reported by those who received NRRP funds; and the actual expected outcomes as reported by those who did not receive any funds and thus carried out potential projects using their own resources. Overall, the evidence highlights the relevance of the role of the NRRP in contributing positively to all three areas, although with significant territorial disparities.

Table 4 – *Perceptions of respondents on Sustainable Energy: Changes with PNRR.*

Macro Area	No funds	Much worsened	Slightly worsened	Unchanged	Slightly improved	Much improved	n.d.	Tot
North -West	35	0	1	7	35	20	17	115
North-East	27	0	0	8	17	17	9	78
Center	17	0	0	7	20	10	13	67
South	20	0	0	6	21	18	5	70
Islands	19	0	0	2	10	9	6	46
Italy	118	0	1	30	103	74	50	376

Source: Our elaboration on Italian Municipalities NRRP Survey.

With reference to energy sustainability of municipalities, Table 4 shows that about 47% of municipalities expects an improvement, among which 20% expect a high improvement, while only 0.3% fears a consistent worsening. Perceptions of significant improvements are more spread in the northwestern and southern areas, especially in the insular zones. Nevertheless, a significant share (32%) of municipalities has not yet received specific dedicated funds.

More pessimistic perceptions arise looking at the framework without NRRP[2]. Within this context, about 33% municipalities expect unchanged conditions, while 7% expects a sharp worsening and the 10% a slight worsening. Municipalities located in the southern part of the country report worsened expectations. Among them, 17% expect a significant worsening. Opposite, the northwestern and northeastern areas are more optimistic, even if worrying about likely negative scenarios without the financial support. In general, municipalities report a moderate optimism with reference to benefits connected to NRRP funds on energy sustainability, thinking that, without them, the changings will be contained further implying a high risk of worse conditions.

[2] From this point onward, detailed data on expectations without NRRP funding are available from the authors upon request, due to space constraints.

Table 5 – *Perceptions of respondents on Energy Savings: Changes with PNRR.*

Macro Area	No funds	Much worsened	Slightly worsened	Unchanged	Slightly improved	Much improved	n.d.	Tot
North -West	35	0	0	14	28	20	18	115
North-East	26	0	0	6	20	16	10	78
Center	16	0	0	4	24	10	13	67
South	19	0	0	5	20	21	5	70
Islands	18	0	0	2	13	6	7	46
Italy	114	0	0	31	105	73	53	376

Source: Our elaboration on Italian Municipalities NRRP Survey.

Shifting to opinions on energy savings, NRRP funds reveals as a significant opportunity of improvement, as reported in Table 5. Namely, about 19% of respondents expects relevant benefits, while almost 28% predicts a modest improvement (low). Only 8% of municipalities report no changes while none of the respondents expect negative outcomes. At the territorial level, the northwestern macro area shows more positive expectations, with 48 municipalities expecting improvements, while the islands are more cautions, with only 19 optimistic municipalities. About 31% municipalities (114), most of them located in the islands and in northwest, have not yet received NRRP funds, therefore not yet benefitting from any intervention promoting energy savings. Without NRRP funds, interventions connected to energy savings have not been effectively implemented, thus suggesting an overall negative stagnant situation. Specifically, most municipalities expect a stable evolution, while about 20% fear a general worsening, among which 8% expect a sharp worsening. Only 6% expect a significant improvement, and the 18% expects a slight improvement. Southern regions show the highest concerns of worsening, with about 17% of municipalities fearing a significant deterioration of energy saving trends. In summary, without NRRP funds, the overall perception is that the energetic saving would remain stable or worse, with low expectations of significant improvement because of not access to resources or less effective policies.

Answers related to the added operative capacity for the renewable energies, showed in Table 6, depict a critical framework with reference to the development perspectives in Italian municipalities. Specifically, NRRP is expected to positively impact the area, but with rather contained expectations. For the period 2021-2026, most municipalities (more than 21%) expect a limited improvement of renewable capacity, while, only a minor share, 16%, foresees relevant enhancements. A significant share of respondents, i.e., 12% perceives a stable situation. The perception of a critical framework, with a less percentage of municipalities expecting significant improvements, is predominant in the Islands. Nevertheless, despite these regional peculiarities, the general framework remains positive. Remarkably, the first column of the Table also highlights that, for projects in this area, about 36% of municipalities, has not received NRRP funds, and have been excluded from the opportunities to increase the renewable capacity.

Table 6 – *Perceptions of respondents on Renewable Capacity: Changes with PNRR.*

Macro Area	No funds	Much worsened	Slightly worsened	Unchanged	Slightly improved	Much Improved	n.d.	Tot
North -West	41	0	0	16	29	11	18	115
North-East	30	0	0	8	15	14	11	78
Center	21	0	1	8	15	8	14	67
South	21	0	0	9	18	17	5	70
Islands	20	0	0	4	4	11	7	46
Italy	133	0	1	45	81	61	55	376

Source: Our elaboration on Italian Municipalities NRRP Survey.

Responses regarding expected changes in the absence of direct NRRP funding highlight a generally conservative outlook. Most municipalities (34%) foresee a stable trajectory, with renewable energy capacity remaining unchanged over time, although a considerable share anticipates a deterioration. Specifically, around 7% of respondents expect a significant decline, while 9% anticipate only a marginal decrease. Expectations of improvement are more cautious: only 5% of municipalities foresee a substantial increase in renewable capacity, while approximately 18% expect a modest improvement. Regional patterns are notable: in the Northwest, about 37% of municipalities expect no significant change, with relatively few negative expectations. In contrast, municipalities in the South and Islands display a more pessimistic outlook, with a larger proportion anticipating either no change or a deterioration. Overall, in the absence of NRRP support, the outlook appears more pessimistic: most municipalities expect limited or no growth, and in some cases even a decline in installed renewable capacity.

As seen in previous project areas, while the NRRP may foster some local improvements in renewable energy, expectations for significant growth remain limited. A lack of dedicated funding and persistent regional disparities contribute to a stagnant outlook unless stronger local policy action is taken.

5. Intermediate implementation of NRRP, in year 2023.

This section analyzes the outcomes of NRRP implementation as reported by the surveyed municipalities, focusing on territorial sustainability through specific projects aimed at promoting energy transition. As a relevant example, we illustrate in Table 7 outcomes as to the renewable energy production. The table presented refer to the year 2023 — two years after the launch of the NRRP in year 2021 — and illustrate results from an ongoing implementation phase. For 2023, it is recorded the actual value indicated by municipalities and the expected value in the absence of NRRP funding. As indicated above, the latter reflects the counterfactual scenario reported by those who received NRRP funds and the actual outcomes as reported by those who did not receive any funds and thus carried the activities using their own resources (therefore, in this case, the observed and the expected value without NRRP are the same).

Table 7 – *MWh through renewable resources (2021, 2023 actual and expected without NRRP – abs. and perc. variation).*

Macroarea	MWh through renewable resources (2021) – Abs. Values (1)	MWh through renewable resources (2023) – Actual MWh Abs. Values (2)	MWh through renewable resources (2023) – Expected value without NRRP Abs. Values (3)	MWh through renewable resources - 2021-2023 actual - Var: ((2)-(1))/(1)	MWh through renewable resources - 2021-2023 without NRRP - Var: ((3)-(1))/(1)
Centro	28561.91	53576.42	53576.22	0.876	0.876
Isole	756.43	663.95	658.53	-0.122	-0.129
Nord - Est	7966.46	8008.76	7734.81	0.005	-0.029
Nord - Ovest	18239.54	21965.23	21901.52	0.204	0.201
Sud	567.68	652.20	449.16	0.149	-0.209
Italy	56092.02	84866.57	84320.24	0.513	0.503

Source: Our elaboration on Italian Municipalities NRRP Survey.

The comparative analysis of renewable energy production across Italian regions between 2021 and 2023—based on observed values (including NRRP funds) and expected values without NRRP funds—offers a comprehensive overview of the dynamics driving territorial sustainability, as sustained by NRRP resources. At the national level, the total volume of renewable energy produced increased by approximately 51.3% with NRRP funding while it would grow only 50.3% in the absence of such funding. This modest difference suggests that, overall, the Italian renewable energy system exhibited significant growth during the period. However, in the absence of NRRP support, it would have grown at a slightly reduced pace. Therefore, the presence of targeted national funding appears to have contributed positively to the overall expansion.

Moreover, the aggregate figures conceal relevant regional disparities that highlight differentiated levels of structural capacity, institutional readiness, and dependence on external policy stimuli. When broken down by macro-area, these variations become particularly evident. Central Italy recorded the most substantial increase in renewable energy production between 2021 and 2023, with an identical growth rate of above 87.6% in both scenarios reflecting a mature and resilient regional ecosystem, likely due to a combination of pre-existing infrastructure, favorable regulatory environments, and active engagement in energy communities. In the North-West, growth was significant yet more moderate, ranging about 20.4%, both with NRRP support and without. This confirms the presence of a solid base for development, in which NRRP funds have acted as an enabling but non-essential factor. The North-East presents a less dynamic picture, with an almost negligible increase (+0.5%) under NRRP and a slight decline (–2.9%) in the absence of funds. This highlights a stagnating pattern and suggests that while public support may help maintain production levels, structural improvements or local mobilization mechanisms are still lacking. The presence of NRRP interventions helped contain a potential deterioration, indicating a beneficial alignment between policy priorities and local needs. This suggests that the NRRP has played a constructive role in supporting territorial resilience. The situation in Southern Italy is particularly telling: the area experienced a modest increase (+14.9%) with NRRP support but a substantial contraction (–20.9%) in its absence. This wide gap indicates a strong dependency on external funding to activate or sustain renewable energy production. The South appears to lack the autonomous capacity to support the energy transition without targeted investment. Finally, the Islands display a negative trend in both scenarios (–12.2% with NRRP; –12.9% without), suggesting the persistence of systemic obstacles—such as infrastructural deficits, administrative inefficiencies, or limited project uptake—that hinder renewable energy development despite favorable natural conditions. In conclusion, while national-level indicators point to strong overall growth, the disaggregated data underscore the critical role of place-based disparities. The NRRP has contributed to supporting weaker territories toward a more balanced and inclusive energy transition, especially in the South, yet its influence remains limited in structurally robust regions.

6. Conclusion

The paper illustrates outcomes from a survey directed to Italian municipalities on NRRP impact. The answers highlight a generic optimism within municipalities, with reference to the impact of NRRP funds on territorial development and sustainability. Most of the territories expect significant improvements, especially in the northwestern and southern areas, while the insular regions highlight a more cautious and less optimistic perception. Nevertheless, it is worth - mentioning that a relevant share of municipalities, 32%, has not yet received specifically dedicated funds for multiple areas, among which energy and territorial management, thus limiting the potential expected benefits. The perception of an improvement without the direct intervention of funds seems to be rather contained, with many territories expecting unchanging or worse situations. In particular, expectations related to renewable energies and energetic saving are generally moderate, indicating that, without dedicated NRRP resources, progress could be contained or insufficient. Nevertheless, the data suggest that financial support has contributed to strengthening local development efforts, reinforcing the view that dedicated investment is essential to accelerate the transition toward more sustainable and efficient territorial systems. While some encouraging signals have already emerged at this interim stage of NRRP implementation, the actual impact of project plans will need to be assessed at the end of the programming period. Continuous monitoring remains crucial—not only to evaluate the overall effectiveness of fund allocation, but also to assess territorial outcomes, particularly in light of one of the NRRP's key goals: reducing regional disparities and closing the longstanding development gap between Southern Italy and the rest of the country.

Acknowledgements

Research project "An impact assessment based on surveys of the NRRP on Italian municipalities" - PRIN 2022 PNRR (Project code P2022RR82F, CUP I53D23007340001) financed by the European Union - NextGenerationEU within the National Recovery and Resilience Plan - Mission 4 "Education and Research" - Component 2 "From research to enterprise" - Investment 1.1 "National Research Programme and Projects of National Interest (PRIN)".

References

ANCI. 2023. Il PNRR nelle città e nei comuni: Ruolo, progetti e strategie nella fase di attuazione del Piano Nazionale di Ripresa e Resilienza. Roma: ANCI.

ENTICOTT G. 2003. Researching Local Government Using Electronic Surveys. *Local Government Studies*, Vol. 29, No. 2, pp. 52–67.

ERMINI B., FIORILLO F., MATTIOZZI S., MARIANI F., SALVATI L. On the use of a Pre-Analysis Plan: an Application to an Observational Study in Social Sciences. *Forthcoming in: Economia Pubblica. The Italian Journal of Public Economics & Law.*

KRAUSE R.M., FATEMI S.M., NGUYEN LONG L.A., ARNOLD G., HOFMEYER S.L. 2024. What is the future of survey-based data collection for local government research? Trends, strategies, and recommendations. *Urban Affairs Review*, Vol. 60, No.3, pp. 1094-1115.

IFEL. 2024. *Lo stato di attuazione del PNRR e il ruolo dei comuni*. Roma: Fondazione IFEL

ITALIA DOMANI. 2025. *Il Piano Nazionale di Ripresa e Resilienza. #NextGenerationItalia*. Consultato il 4 maggio 2025, https://italiadomani.gov.it

MOORE L. V., CARLSON S. A., ONUFRAK S., CARROLL D. D., GALUSKA D. 2017. Development and implementation of a local government survey to measure community supports for healthy eating and active living. *Preventive medicine reports*, Vol. 6, pp. 74-79.

PRESIDENZA DEL CONSIGLIO DEI MINISTRI. 2025. Rapporto sul fenomeno della povertà in Italia. Roma: Governo.it.

SACCHI A., RUBINO P. 2023. Il disegno e l’attuazione del PNRR: investimenti, dimensione territoriale e percorsi di sviluppo.*Argomenti*, Vol. 24, No. 1, pp. 1-34.

Clio CIASCHINI, Università Politecnica delle Marche, c.ciaschini@univpm.it

Barbara ERMINI, Università Politecnica delle Marche, b.ermini@univpm.it

Luca SALVATI, Università degli Studi di Roma La Sapienza, luca.salvati@uniroma1.it

Gianluigi SALVUCCI, Istat, salvucci@istat.it

FEELING RESPONSIBLE, FEELING BETTER? THE CLIMATE-HAPPINESS LINK ACROSS EUROPEAN UNION

Vincenzo Marinello, Rossella Maria Pia Di Rocco, Alessandro Indelicato

Abstract. Climate change is one of the most pressing environmental issues, and often cause citizzens' anxiety. One such pressing concern is how citizens' sense of personal responsibility for climate change affects their overall life satisfaction. The current study examines this relationship using data from Round 11 (2023) of the European Social Survey (ESS). We examine the extent to which feeling personally responsible for mitigating climate change is associated with life satisfaction, controlling for a number of socio-demographic and attitudinal controls. These include age, gender, political orientation, country of residence, migrant status, whether one believes that climate change exists, how concerned one is about protecting the environment, how close one feels to Europe and one's country of origin, and trust in the European Parliament. The results suggest that individuals who identify more strongly with the task of tackling climate change, report more trust in European institutions, and reside in northern European Member States enjoy greater life satisfaction. Conversely, individuals - mainly from some Eastern and Southern European countries - who either express low identification with Europe or are apathetic about identifying with it are less satisfied.

1. Introduction

Climate change and environmental degradation represent some of the most serious challenges facing the planet today. The effects of climate change are becoming a reality for an increasing share of the population, so there is a parallel in the increase of people experiencing environmental discomfort, commonly known as "eco-anxiety" or "climate anxiety" (Ojala *et al.*, 2021). Passmore *et al.*, (2023) define the latter as "persistent feelings of concern, anxiety, dread or doom regarding environmental degradation and the impacts and implications of climate change on our planet as a whole". Therefore, the prevalence of such feelings and the multiple ecological challenges facing the population pose a threat to one's level of life satisfaction (LS). The scientific literature describes LS as a general assessment of one's attitude and feelings about one's life at a given time and is an important indicator of one's well-being (Diener, 1984). Most of the investigations conducted have focused more on the influence of climate change on LS. Rising temperatures for those living in areas with a harsh climate lead to an improvement in LS, reducing it in warm climates (Maddison and Rehdanz, 2011). On the other hand, an australian study showed that heat-related stress has no effect on either LS or happiness (Zander *et al.,* 2019). Drought and the threat of drought have been shown to have a negative impact on LS, especially among the poorer segments of the population (Berlemann and Eurich, 2022). In addition, extreme weather events, such as floods and hurricanes, have been shown to impair the LS of people in affected areas beyond the immediate impact on well-being, with prolonged effects over time (Calvo *et al.,* 2015; Fernandez *et al.,* 2019; Sekulova and Van den Bergh, 2016). Growing levels of anxiety about climate change are now evident worldwide, with women, young people and indigenous communities being most affected (Burke *et al.,* 2018; Coffey *et al.,* 2021; Petheram *et al.,* 2010). In response to eco-anxiety, defence mechanisms are activated, prompting subjects to action, resulting in a sense of personal responsibility (Innocenti *et al.,* 2023). The latter can have an adaptive effect, prompting concrete actions to mitigate and diminish the effects of climate change through personal behaviour and political support (Bouman *et al.,* 2020). Overall, feeling responsible on a personal level can reflect care for oneself, the community and society at large, accompanied by a drive to build a better world (Lio *et al.,* 2023). Pro-environmental behaviour changes consumption patterns by directing individuals towards relatively low-impact alternatives (e.g. buying an electric car instead of petrol) or towards reducing overall consumption (C. Chen *et al.,* 2024; W. Chen and Xia, 2020). However, there is more research focus on climate anxiety and the resulting sense of individual responsibility, but a gap remains in that there is no study or empirical evidence in the current academic literature linking LS to a sense of responsibility in climate change mitigation. The main objective of this study is to fill the gap regarding the aforementioned relationship by answering several research questions that, in turn, aim to explore the underlying dynamics of this relationship. First, it investigates whether there are actual differences in levels of LS and sense of personal responsibility for climate among the EU member states included in the study. On the other hand, the extent to which feeling personally responsible for climate change mitigation is associated with the level of LS is investigated, controlling for a number of socio-demographic and attitudinal variables. The article is structured as follows: section 2 shows the dataset used; the methodology for the analysis is outlined in section 3; section 4 reports the results obtained, while the last section presents a focus on the discussions and conclusions, also highlighting the limitations of the study.

2. Data

The data for this study are extracted from Round 11 (2023) of the European Social Survey (ESS) in order to analyse the level of LS and sense of responsibility on the implementation of measures to reduce climate change. The ESS is, without a doubt, a valuable source of scales measuring the environmental attitudes of European citizens. At the same time, it provides solid information on levels of LS, demonstrating the importance and robustness of the data collected in order to carry out comparative analyses on a continental scale and study connections (Ferreira *et al.,* 2013; Kácha *et al.*, 2022).

The research focuses on 11 EU member countries, providing a robust framework for examining the interaction between LS and European citizens' sense of responsibility towards climate change, taking into account various socio-economic and attitudinal factors. The choice to include only EU member states may be motivated by the fact that they share a common environmental policy, thus facilitating the interpretation of our study results. Furthermore, a cautious methodological approach was chosen in order to ensure a balance between representativeness of the sample and empirical practicality.

The countries included in the analysis are Austria, Germany, Spain, Finland, France, Greece, Hungary, Italy, Poland, Portugal and Sweden. The socio-economic variables taken into account include gender, age, political orientation, income, migrant status, attachment to the country, perceptions on the causes of climate change, care for the environment, confidence in walking in the dark and trust in the European Parliament. The descriptive statistics of the sample are given in Table 1 and show the number of respondents and the percentage of each category included in the study, revealing significant heterogeneity. There are over 21,000 respondents distributed among the 11 European countries. Italy, Greece, Germany and Austria account for almost half of the sample. In terms of age, the most representative group is made up of those aged over 65 (26.9%), while the least representative are those aged under 25 (10.6%). In terms of gender, there is a slight predominance of women (53.8%). Concerning the economic situation, a large majority of the respondents (48.6%) state that they manage to meet their daily expenses, only 3.5% are in great difficulty. Regarding personal responsibility in the fight against climate change, the sample is characterised by citizens who feel highly (35.8%) or moderately (34.5%) responsible, against only 5% who feel they have no responsibility at all. Moreover, 40% of respondents attach great importance to caring for the environment. More than 90% are indigenous, and at the same time, almost half of the sample surveyed are strongly attached to their country.

Table 1 – *Descriptive statistics.*

Country	n	%	CC Cause	n	%
Austria	2354	10.8%	Entirely by natural processes	364	1.7%
Germany	2420	11.1%	Mainly by natural processes	1360	6.3%
Spain	1844	8.5%	About equally by natural processes and human activity	8952	41.2%
Finland	1563	7.2%	Mainly by human activity	8893	40.9%
France	1771	8.1%	Entirely by human activity	2013	9.3%
Greece	2757	12.7%	I don't think climate change is happening	155	0.7%
Hungary	2118	9.7%	CC Responsability	n	%
Italy	2865	13.2%	No responsibility	1086	5.0%
Poland	1442	6.6%	Low responsibility	2026	9.3%
Portugal	1373	6.3%	Moderate responsibility	7501	34.5%
Sweden	1230	5.7%	High responsibility	7781	35.8%
EP Trust	n	%	Full responsibility	3343	15.4%
No trust	2939	13.5%	Gender	n	%
Low trust	3925	18.1%	Male	10034	46.2%
Moderate trust	9860	45.4%	Female	11703	53.8%
High trust	4183	19.2%	Age	n	%
Full trust	830	3.8%	under 25	2303	10.6%
Political orientation	n	%	26-35	2708	12.5%
Left	1115	5.1%	36-45	3311	15.2%
Centre Left	3297	15.2%	46-55	3685	17.0%
Centre	12392	57.0%	56-65	3888	17.9%
Centre Right	3649	16.8%	over 65	5842	26.9%
Right	1284	5.9%	Income	n	%
Safe Dark	n	%	Living comfortably on present income	6581	30.3%
Very safe	6366	29.3%	Coping on present income	10565	48.6%
Safe	10999	50.6%	Difficult on present income	3829	17.6%
Unsafe	3606	16.6%	Very difficult on present income	762	3.5%
Very unsafe	766	3.5%	Born	n	%
Attachment country	n	%	Born in country	19913	91.6%
No attachment	243	1.1%	Foreign Born	1816	8.4%
Low attachment	572	2.6%	Environment Care	n	%
Moderate attachment	3029	13.9%	Very much like me	6703	30.8%
High attachment	7068	32.5%	Like me	8694	40.0%
Strong attachment	10825	49.8%	Somewhat like me	4181	19.2%
			A little like me	1677	7.7%
			Not like me	388	1.8%
			Not like me at all	94	0.4%

EP, European Parliament; CC, climate change.
Source:own elaboration.

3. Methodology

An ordered logit regression model is a statistical technique used to analyse ordinal data, where the dependent variable (Y) has more than two ordered categories. In the present study, the dependent variable was life satisfaction (LS) on an 11-point rating scale (0–10). The ordered logit model predicts the probability of an individual belonging to a higher LS category based on a set of explanatory variables (Alemi et al., 2019). The model can be expressed algebraically as follows:

$$P(Y \leq j \mid X) = \frac{exp(\tau_j - X\beta)}{1 + exp(\tau_j - X\beta)}, \; j = 1, \ldots, J-1$$

where $P(Y \leq j \mid X)$ is the cumulative probability of measuring life satisfaction at or below category j for the given explanatory variable X, and where the cut-points (or thresholds) to be estimated are represented by τ_j and the regression coefficients on the independent variables are represented by β. The logit distribution is chosen because it has a simple functional form and is easily interpretable in terms of odds ratios.

The independent explanatory variables are sense of personal obligation to combat climate change, gender, age, political views, income, migrant status and sense of belonging to one's nation. Other variables include perceived explanations for climate change, concern for the environment, perceived safety when walking alone at night and trust in the European Parliament.

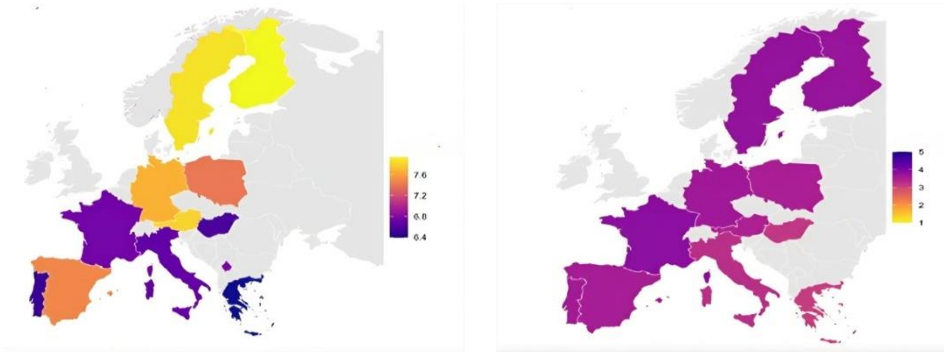
As with many studies using ordered logit models, dummy coding was employed to introduce categorical variables into the model, with one category serving as the reference point (West et al., 1996). The reference categories were chosen to be middle or neutral points to enable comparison between groups (e.g. female gender, centre political orientation and moderate institutional trust).

This specification enables us to make maximum use of the life satisfaction scale range without losing information introduced by dichotomising the dependent variable. Lastly, the ordered logit model provides a richer explanation of how climate responsibility and related socio-economic and attitudinal factors cause shifts in the probability distribution between life satisfaction levels.

4. Results

This section presents the results obtained from the analysis. First of all, choroplethic maps are shown, which are a survey tool that allows us to graphically represent the spatial distribution of the variables under study according to the responses obtained from the last ESS survey. This type of map makes it possible to identify trends, patterns and perceptions that vary according to geographical area. Figure 1(a) shows the values of the average perception of LS on a colour scale, highlighting significant geographical differences. First of all, it outlines a relevant aspect: Northern European countries are characterised by a higher perception in terms of LS, with an average of more than 7.6. Central European nations, such as Germany and Austria, together with Poland and Spain, are in an intermediate position regarding the perception of LS. In these nations, the average values observed vary between 7 and 7.6. However, France presents itself as an exception: although it is seen as one of the largest economies in Western Europe, it is characterised by a lower level of perceived satisfaction, showing similar values typical of southern countries. In addition, Figure 1(a) shows that citizens in southern European countries, such as Italy, Portugal and Greece, tend to report significantly lower levels of satisfaction than the European average. Figure 1(b) illustrates, on the other hand, the average perceived level of personal responsibility in mitigating climate change. The results indicate a rather uniform and generally high sense of responsibility among the countries analysed, with the average being between 4.5 and 3.5 in most cases. In the Nordic countries, such as Sweden and Finland, citizens seem to have a higher level of personal responsibility. The same trend is recorded in France. Also in the central-western and southern European countries, such as Germany, Austria, Poland, Spain and Portugal, inhabitants tend to be more inclined to perceive that sense of responsibility for the climate. Significantly lower than average values are observed, however, in Italy, Greece and Hungary in Italy, Greece and Hungary.

Figure 1 - *Cross-country variation in perceptions within the European Union.*



(a) Perceived LS

(b) Perceived personal responsibility for climate change mitigation.

Source: own elaboration.

Furthermore, the study investigates the possible relationship between the different levels of life satisfaction (LS) and the sense of personal responsibility towards climate change, controlling for socio-demographic and attitudinal factors. Table 2 shows the results of the estimation of an ordered logit regression model, focusing on statistically significant independent variables.

Table 2 – *Ordered Logit Regression Model Results.*

Variable	Item	Estimate	Std. Error	z-value	p-value
Country	ES	-0.32	0.06	-5.67	0.00
	FR	-1.05	0.06	-18.06	0.00
	GR	-0.93	0.05	-17.86	0.00
	HU	-1.06	0.06	-19.16	0.00
	IT	-0.95	0.05	-19.04	0.00
	PL	-0.44	0.06	-7.07	0.00
	PT	-1.04	0.06	-16.89	0.00
	SE	-0.22	0.06	-3.45	0.00
EU Trust	No trust	-0.18	0.04	-4.60	0.00
	Low Trust	-0.18	0.03	-5.14	0.00
	High Trust	0.18	0.03	5.50	0.00
	Full trust	0.64	0.07	9.42	0.00
Political orientation	centre left	-0.25	0.04	-7.20	0.00
	centre right	0.18	0.03	5.44	0.00
	right	0.68	0.06	12.09	0.00
Safe Dark	very safe	0.34	0.03	11.63	0.00
	unsafe	-0.21	0.04	-6.01	0.00
	very unsafe	-0.44	0.07	-6.26	0.00
Attachment country	Low attachment	-0.34	0.08	-4.15	0.00
	High attachment	0.36	0.04	9.27	0.00
	Strong attachment	0.78	0.04	20.08	0.00
CC Cause	clim. ch. 1	-0.50	0.10	-4.85	0.00
	clim. ch. 4	-0.08	0.03	-3.07	0.00
	clim. ch. 5	-0.16	0.05	-3.59	0.00
	NO clim. ch.	-0.43	0.15	-2.91	0.00
CC Responsability	No responsibility	0.28	0.06	4.38	0.00
	High responsibility	0.21	0.03	6.99	0.00
	Full responsibility	0.39	0.04	9.31	0.00
Gender	male	-0.11	0.03	-4.23	0.00
Age	46-55	-0.14	0.04	-3.40	0.00
	56-65	-0.22	0.04	-5.19	0.00
	over 65	-0.26	0.04	-6.68	0.00
Income	income 1	0.66	0.03	21.89	0.00
	income 3	-0.71	0.04	-20.14	0.00
	income 4	-1.27	0.07	-18.07	0.00
Environment Care	Very much like me	0.26	0.04	7.11	0.00
	Like me	0.17	0.03	5.03	0.00

Source: own elaboration.

Feeling strongly responsible for mitigating climate change is a significant predictor of the perception of having a fulfilling life. Another equally relevant predictor is trust in European institutions, particularly the European Parliament: high levels of trust are positively associated with the likelihood of perceiving one's life as highly satisfying. Conversely, citizens who report low levels of trust show an inverse relationship with life satisfaction.

A further key factor is the degree of attachment to one's country of origin: those who have strong ties to it report higher levels of satisfaction. A similar trend can be observed among those who adhere to centre-right and especially conservative political positions: in fact, they tend to report higher levels of satisfaction than those who identify with progressive positions. Compared to men and older individuals, women and young people report a more satisfying perception of life.

The coefficients relating to the perception of the causes of climate change emerge as relevant, being negative and significant. The belief that climate change is attributable to natural factors or human activities, or denying it altogether, tends to reduce the level of life satisfaction. Finally, with regard to the relationship between life satisfaction and the importance attributed to caring for the environment, the analysis highlights a positive association: people who say they pay close attention to the environment are more likely to perceive a highly satisfactory standard of living.

5. Discussions and conclusion

The analysis of the choroplethic maps indicates that respondents show a rather high average level of personal responsibility for climate change, but with variations of considerable interest. The Nordic countries, together with France, show more pronounced levels of responsibility, while the southern and eastern European nations appear to lag behind on this front. Already (García-López and Allué, 2012)in their study, point out that Finland and Sweden are the countries with a greater "positive responsivity" to climate change, indicating a public and proactive perception compared to southern Europe, where there is a negative responsivity, which can be reflected in a more passive attitude of citizens. Focusing on Italy, the "Lancet Countdown" report argues that the national response has been partial, the involvement of citizens in combating climate change remains fragmented and unsystematic. Low institutional trust, low climate literacy and a cultural model less oriented to the collective good contribute to this (Alfano *et al.*, 2023). On the other hand, among Spanish citizens, the sense of responsibility for climate change mitigation is remarkable. Despite institutional inertia and regulatory inadequacy, citizenship's responsibility for a paradigm shift is growing, and in the wake of international examples such as Urgenda in the Netherlands, Spain also sees the growing role of civil society in taking legal action against the state for climate failures and invoking the rights of future generations (Mateo, 2019). Also in France, the high sense of personal responsibility spills over into actions against the federal government and a well-known case is represented by "Grande Synthe", the city that petitioned the high court to take further action against climate change (Torre-Schaub, 2023).

The result of the ordered logit regression model's estimation of the relationship between LS and sense of personal responsibility in climate mitigation is aligned with previous studies: Martin *et al.,* (2024) suggest that a sense of responsibility may play a protective role, while according to Arslan and Wong, (2022) personal responsibility contributes to providing a sense of meaning and purpose in the face of uncertainty that may affect the LS. Baldwin *et al.,* 2023) believe that the link between LS and responsibility could be influenced by the belief that actions can make a difference (self-efficacy). Furthermore, our results showed that high levels of trust in the LS are positively associated with the likelihood of perceiving one's life as highly satisfying. Clench-Aas and Holte, (2021) found that trust in institutions provides a sense of stability, order and predictability, especially in times of crisis or adverse conditions trust acts as a buffer, according to the buffer hypothesis, cushioning the negative impact on life perception, as well as the resulting satisfaction. On the other hand, Deák *et al.,* (2024) argue that those with high institutional trust are more likely to be climate proponents, contributing to civic mobilisation and engagement in climate change mitigation.

This study contributes to the scarce existing literature on mapping personal climate responsibility and analysing this relationship. First, a large and representative sample of the countries included in the analysis is used in order to provide a solid empirical basis for generalising the results. Furthermore, the estimation of a regression model allows the association between LS and responsibility to be captured. The results of this study can be used to improve, and in some cases, readjust educational campaigns and programmes in order to contribute to a sense of personal accountability especially in those countries where there is less perception. This implies that climate communication campaigns should emphasise the effectiveness of individual actions rather than simply focusing on the urgency of the problem: in fact, a greater sense of responsibility, translated into concrete behaviour, tends to lead to a general increase in LS. However, some limitations are present in this study. Firstly, this analysis refers to a single year, namely 2023; therefore, future research should extend the investigation to multiple rounds of the ESS, allowing for longitudinal studies, generating more robust conclusions in causal terms and assessing the stability of results over time. At the methodological level, more sophisticated methodologies could be used to improve the robustness of the results regarding perceptions on the sense of responsibility, such as Fuzzy logic, which allows composite indices to be created by combining various socio-demographic and attitudinal factors in order to better understand the “grey and nuanced areas” between responsibility and LS. In addition, future research could explore the non-linearity of the relationship between LS and responsibility by differentiating various forms of climate engagement, such as activism, sustainable consumption and everyday behaviour. Finally, it would be useful to investigate the role of supranational institutions, such as the EU, in shaping the sense of responsibility in climate dynamics.

References

ALEMI F., CIRCELLA G., MOKHTARIAN P., HANDY, S. 2019. What drives the use of ridehailing in California? Ordered probit models of the usage frequency of Uber and Lyft. *Transportation Research Part C: Emerging Technologies*, Vol. 102, pp. 233-248.

ALFANO R., DE’DONATO F., VINEIS P., ROMANELLO M. 2023. Gli indicatori di Lancet Countdown per l’Italia: monitorare i cambiamenti climatici e la salute. *Epidemiologia e Prevenzione*, Vol.47, No.3, pp. 6–21.

ARSLAN G., WONG P. T. P. 2022. Measuring personal and social responsibility: an existential positive psychology approach. *Journal of happiness and health*, Vol. 2, No.1, pp.1–11.

BALDWIN C., PICKERING G., DALE G. 2023. Knowledge and self-efficacy of youth to take action on climate change. *Environmental Education Research*, Vol. 29, No.11, pp.1597–1616.

BERLEMANN M., EURICH M. 2022. Does drought risk depress expected well-being? *Applied Economics Letters*,Vol. 29, No.13, pp. 1229–1233.

BOUMAN T., VERSCHOOR M., ALBERS C. J., BÖHM G., FISHER S. D., POORTINGA W., WHITMARSH L., STEG L. 2020. When worry about climate change leads to climate action: How values, worry and personal responsibility relate to various climate actions. *Global Environmental Change*, Vol. 62, 102061.

BURKE S. E. L., SANSON A. V, VAN HOORN J. 2018. The psychological effects of climate change on children. *Current Psychiatry Reports*, Vol.20, pp. 1–8.

CALVO R., ARCAYA M., BAUM C. F., LOWE S. R., WATERS M. C. 2015. Happily ever after? Pre-and-post disaster determinants of happiness among survivors of Hurricane Katrina. *Journal of Happiness Studies*, Vol.16, pp. 427–442.

CHEN C., LI D., QIAN J., LI, Z. 2024. The Impact of Green Purchase Intention on Compensatory Consumption: The Regulatory Role of Pro-Environmental Behavior. *Sustainability (2071-1050)*, Vol. 16, No.18.

CHEN,W., XIA N. 2020. The expression of social pro-environmental preference influence on green consumption behavior. *IOP Conference Series: Earth and Environmental Science*, Vol.552, No. 1, 012019.

CLENCH-AAS J., HOLTE A. 2021. Political trust influences the relationship between income and life satisfaction in Europe: differential associations with trust at national, community, and individual level. *Frontiers in Public Health*, Vol. 9, 629118.

COFFEY Y., BHULLAR N., DURKIN J., ISLAM M. S., USHER K. 2021. Understanding eco-anxiety: A systematic scoping review of current literature and identified knowledge gaps. *The Journal of Climate Change and Health*, Vol. 3, 100047.

DEÁK Z., VAZNONIENĖ G., WOJEWÓDZKA-WIEWIÓRSKA A. 2024. Understanding the attitudes towards climate change in the context of trust in selected countries of Central and Eastern Europe. *Serbian Journal of Management*, Vol. 19, No. 2, pp. 449–470.

DIENER E. 1984. Subjective well-being. *Psychological Bulletin*, Vol. 95, No.3, p. 542.

FERNANDEZ C. J., STOECKL N., WELTERS R. 2019. The cost of doing nothing in the face of climate change: a case study, using the life satisfaction approach to value the tangible and intangible costs of flooding in the Philippines. *Climate and Development*, Vol. 11, No.9, pp. 825–838.

FERREIRA S., AKAY A., BRERETON F., CUÑADO J., MARTINSSON P., MORO M., NINGAL T. F. 2013. Life satisfaction and air quality in Europe. *Ecological Economics*, Vol. 88, pp.1–10.

GARCÍA-LÓPEZ J. M., ALLUÉ C. 2012. A phytoclimatic-based indicator for assessing the inherent responsitivity of the European forests to climate change. *Ecological Indicators*,Vol. 18, pp. 73–81.

INNOCENTI M., PERILLI A., SANTARELLI G., CARLUCCIO N., ZJALIC D., ACQUADRO MARAN D., CIABINI L., CADEDDU C. 2023. How does climate change worry influence the relationship between climate change anxiety and eco-paralysis? A moderation study. *Climate*, Vol. 11, No.9, p. 190.

KÁCHA O., VINTR J., BRICK C. 2022. Four Europes: Climate change beliefs and attitudes predict behavior and policy preferences using a latent class analysis on 23 countries. *Journal of Environmental Psychology*, Vol. 81, 101815.

LIO S., DEGENG I. N. S., HAMBALI I., HITIPEUW I. 2023. Self-awareness and personal responsibility among seminary students. *Innovations in Education and Teaching International*,Vol. 60, No. 4, pp. 544–554.

MADDISON D., REHDANZ K. 2011. The impact of climate on life satisfaction. *Ecological Economics*, Vol. 70, No. 12, pp. 2437–2445.

MARTIN G., ROSWELL T., COSMA A. 2024. Exploring the relationships between worry about climate change, belief about personal responsibility, and mental wellbeing among adolescents and young adults. *Wellbeing, Space and Society*, Vol. 6, 100198.

MATEO B. S. 2019. Responsabilidad pública, vulnerabilidad y litigios climáticos. *Revista Aragonesa de Administración Pública*, Vol. 54, pp. 57–140.

OJALA M., CUNSOLO A., OGUNBODE C. A., MIDDLETON J. 2021. Anxiety, worry, and grief in a time of environmental and climate crisis: A narrative review. *Annual Review of Environment and Resources*, Vol. 46, No.1, pp. 35–58.

PASSMORE H.-A., LUTZ P. K., HOWELL A. J. 2023. Eco-anxiety: A cascade of fundamental existential anxieties. *Journal of Constructivist Psychology*, Vol. 36, No. 2, pp.138–153.

PETHERAM L., ZANDER K. K., CAMPBELL B. M., HIGH C., STACEY N. 2010. 'Strange changes': Indigenous perspectives of climate change and adaptation in NE Arnhem Land (Australia). *Global Environmental Change*, Vol. 20, No.4, pp. 681–692.

SEKULOVA F., VAN DEN BERGH J. C. J. M. 2016. Floods and happiness: Empirical evidence from Bulgaria. *Ecological Economics*, Vol. 126, pp. 51–57.TORRE-SCHAUB M. 2023. *Affaire Grande-Synthe: doit-on s' inquiéter pour l'avenir de la justice climatique en France?*

WEST S. G., AIKEN L. S., KRULL J. L. 1996. Experimental Personality designs: analyzing categorical by continuous variable interactions. *Journal of personality*, Vol. 64, No. 1, pp. 1-48.

ZANDER K. K., MOSS S., GARNETT S. T. 2019. Climate change–related heat stress and subjective well-being in Australia. *Weather, Climate, and Society*, Vol. 11, No. 3, pp. 505–520.

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

Rossella Maria Pia DI ROCCO, Università degli Studi di Enna Kore, rossellamariapia.dirocco@unikorestudent.it

Alessandro INDELICATO, Universidad de Las Palmas de Gran Canaria, alessandro.indelicato@ulpgc.es

INNOVATION PATHWAYS IN THE TRAINING OF THE PERMANENT CENSUS INTERVIEWERS AND STAFF (2018-2024)[1]

Nunzia Balì, Gabriella D'Ambrosio, Maria Teresa Fiori, Alessandra Lugli, Nadia Mirante

Abstract. A crucial element for the survey's success is, undoubtedly, the participation of the interviewers in a dedicated training course before starting the fieldwork phase. In this context, the Italian National Institute of Statistics (Istat) has now developed a specific know-how aiming to improve it year after year, also based on the feedback provided by teachers and students. Over time, the training courses have shifted from being entirely in-person to a blended format, introducing new IT support tools such as Moodle.

The purpose of this research is to outline the evolution of the training models for the interviewers of the Permanent Census of Population and Housing from the first year of this new modality (2018) to today (2024). This evolution is the result of the comments and suggestions provided by the interviewers and staff at the end of the training course through the compilation of a satisfaction questionnaire.

The analysis employed non-parametric statistical techniques, including the Kruskal-Wallis test and Dunn's *post hoc* comparisons, to evaluate trends across years.

The results obtained highlighted a positive trend in the appreciation by the learners regarding the usability of the platform, the availability of materials and the support provided by teachers. The data collection process and analysis allowed us to constantly refine the training path, implementing targeted changes to respond to emerging needs, thus confirming the success of the model. The evolution of the system emphasized how the circular approach, based on feedback, continues to guarantee more effective learning and increasing satisfaction of the learners.

1. Introduction

The training of interviewers before the start of fieldwork is undoubtedly a crucial factor in the success of the survey as it represents the transition from the planning phase, in which the research design was developed, to the data collection phase (Istat, 2019). Indeed, the interviewers play a pivotal role in contacting the survey units and motivating them to participate in the survey. Moreover, conducting an interview

[1] The authors share contents and views expressed in this paper. However, N. Balì drew up the sections 1, G. D'Ambrosio the section 2, A. Lugli the section 3, N. Mirante the sections 4 and 5 and M.T. Fiori the section 6.

requires a series of competences, not limited to guiding the interview, that include managing appointments, asking questions exactly as worded in the questionnaire to avoid the interviewers’ effect (West *et al.*, 2013), accurately recording the answers, and ensuring a comfortable interview environment (Gubrium *et al.*, 2012). So, recognizing the key role of interviewers means paying particular attention to the definition of training processes for both surveyors and field staff, essential to ensure that all interviewers acquire the competencies required to perform their tasks in accordance with the high-quality standards established for national statistical surveys (Istat, 2012).

Over time, both training courses and supporting materials have been significantly shaped by methodological and technological advancements. These developments have led to increasingly differentiated training designs, tailored to the specific objectives of each survey and the corresponding data collection techniques (Istat, 2019). Specifically in the case of the Permanent Census of Population and Housing, introduced in 2018 to replace the traditional Census carried out every ten years, a range of new organizational aspects, data collection methods, and technological tools have been implemented (Mirante *et al.*, 2025). In this context, the transition to the new Census model necessitated the redefinition of the training strategy, for example by introducing training models for interviewers based on e-learning or blended approaches, aimed at designing a permanent training model in the light of a circular approach (Istat, 2024).

In recent years, several national statistical agencies and academic institutions have experimented with e-learning or blended training programs for interviewers involved in large-scale surveys and censuses. For example, the U.S. Census Bureau has implemented self-study and web-based modules as part of its interviewer training since the 2010 Census, complemented by classroom and on-the-job sessions (Goerman *et al.*, 2019). Similarly, Statistics Canada employs a blended approach to interviewer training for its Continuous Household Surveys, combining virtual learning environments with monitored field practice and continuous feedback via its Computer-Assisted Personal Interviewing (CAPI) system. The Brazilian Institute of Geography and Statistics (IBGE) has also integrated digital resources into its training programs for enumerators during national census operations, supported by cognitive testing studies to enhance data reliability. In Europe, organizations such as Statistics Netherlands (CBS) emphasize structured training and supervision processes, increasingly supported by e-learning components, as key factors for maintaining data quality in household and demographic surveys (Kockelkoren, 2011). At the international level, the United Nations Statistics Division has promoted e-learning platforms for capacity building in official statistics, including modules on survey methodology and data collection practices. Although these initiatives share common objectives with the Italian National Institute of Statistics (Istat), few published studies have systematically evaluated the effectiveness of e-learning specifically for interviewer training in the context of continuous or permanent census operations (Sun *et al.*, 2024). Therefore, the present study contributes to the existing literature by providing an empirical evaluation of e-learning training outcomes for interviewers in the Italian Permanent Census of Population, offering a novel perspective on digital training within official statistics.

So, following these premises, the objective of this research is to explore learners' assessments of the Italian Permanent Census of Population and Housing self e-learning training over the years from 2018 to 2024.

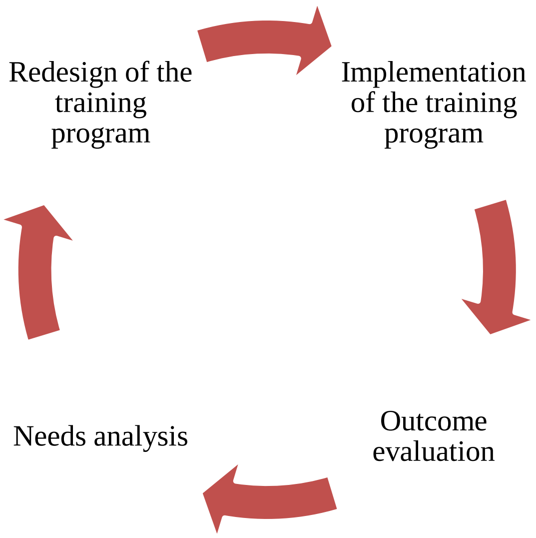
2. From linear to circular: a new training approach for the Permanent Census of Population and Housing

As mentioned in the introduction of this work, following the definition of the new model of the Permanent Census of Population and Housing, also the training models for interviewers and staff have been modified (Castagna, 2022). In detail, the training approach shifted from linear to circular, which aim is to support learners in their professional activities – specifically during the data collection phase – by providing agile, accessible, and practical solutions to real-world problems within a framework of continuous improvement. For this reason, the circular approach is structured in four phases:

- *Needs analysis*: this phase involves identifying educational needs and learner requirements through tools such as questionnaires.
- *Training program design*: this phase emphasizes collaboration among all stakeholders involved in the training design process, ensuring that the educational intervention aligns coherently with the identified needs.
- *Implementation of the training*: this phase represents the moment when the actual training activities take place, actively involving both trainers and trainees.
- *Outcome evaluation*: this final phase aims to assess whether the training objectives have been achieved, while also identifying and addressing any residual learning gaps that may have emerged during the course.

This scheme is summarized in the following Figure (Figure 1).

Figure 1 – *Circular approach implemented for the training course for the interviewers and staff of Italian Permanent Census of Population and Housing 2018-2024.*



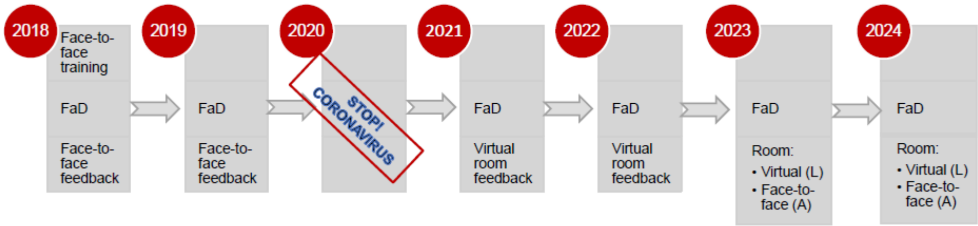
Consequently, in line with a continuous improvement perspective, the training model is periodically revised based on the feedback collected through course evaluation anonymous questionnaires, which learners are invited to complete upon conclusion of the training courses.

3. The transition to a blended learning model

Following this framework and in compliance with the needs of the interviewers and field staff, training models moved from being entirely face-to-face to a blended format, incorporating new IT support tools such as Moodle LMS[2] (Learning Management System) for self-e-learning (the so-called "FaD" – Formazione a Distanza accessible through the link https://formazionereti.istat.it/). The onset of the COVID-19 pandemic further accelerated the demand for greater flexibility and adaptability in training methodologies (Figure 2).

[2] Modular Object-Oriented Dynamic Learning Environment (Moodle) is a computer system for creating and sharing educational materials online.

Figure 2 – *Training models in the Italian Permanent Census of Population and Housing 2018-2024.*



Overall, the adoption of the platform has facilitated the shift from a traditional training model to a continuous training framework. This approach is supported by regularly updated instructional materials, which remain accessible to trainees throughout the entire duration of the survey, enabling autonomous learning.

4. Methodology

As outlined before, the training programs for the Permanent Census of Population and Housing comprised two distinct phases: a face-to-face component and an e-learning part (FaD). The present article is exclusively focused on the self-e-learning training element.

From 2018 to 2024 (except for 2020), the FaD courses were delivered via the self-learning platform Moodle. This platform provided learners with access to a variety of materials, which were organized into different sections. These included learning modules, intermediate self-assessment tests, a final learning test, additional materials, and a course evaluation anonymous questionnaire.

Learners who successfully completed the FaD were invited to fill in the evaluation questionnaire. In order to give the trainees more freedom, completing the questionnaire was anonymous and not mandatory.

The questionnaires were designed to gather feedback on the perceived quality of the FaD. We were particularly interested in receiving evaluations on the following areas: platform usability, material quality, teacher support, overall course satisfaction and any additional suggestions. Except for the final domain, all domains were measured using a Likert-type scale. All questionnaire items were mandatory, resulting in a complete dataset with no missing values.

At the conclusion of each training year, the collected questionnaires were analyzed with the objective of identifying areas for enhancement in subsequent years. Following a review of the suggestions, we implemented those that were feasible.

The present study analyzes data collected from questionnaires completed by learners between 2018 and 2024, by using appropriate statistical tests. For each domain, a trend evaluation was conducted over the years.

Non-parametric tests were applied due to the ordinal nature of the data (Likert scale). Specifically, the Kruskal-Wallis test was used to assess differences across years, as it does not assume normality and is suitable for comparing multiple independent groups. It provides a global test of whether differences exist among independent groups. Dunn's *post hoc* test was employed to identify pairwise differences following significant Kruskal-Wallis results. This choice was preferred over ordinal regression because the primary aim was to detect distributional differences rather than model predictors of satisfaction. To strengthen methodological transparency, the main Kruskal-Wallis and Dunn's results are reported in the results section. The data were also visualized in graphical form (line graph).

5. Results

Between 2018 and 2024, 100,429 learners completed FaD training, with an average questionnaire response rate of 21.34% (N=21,433).

The socio-demographic characteristics remained stable over the years. Overall, the majority of respondents were woman (65.18%), the median age was 46 years (range 18-80 years) and 52.14% of the sample had received a high school education. Almost half of the sample lived in northern Italy (45.02%), while 38.60% lived in the south and 16.38% in the centre.

No statistically significant differences were found across years in terms of professional role (interviewer, supervisor, coordinator, staff member or back-office operator), prior Istat experience (having or not having participated in other Istat surveys), or Census expert status (having or not having previously worked on the Census). Also, no statistically significant differences were observed when the data were stratified by role or Census status (expert or non-expert). Therefore, we reported the overall results for each questionnaire domain.

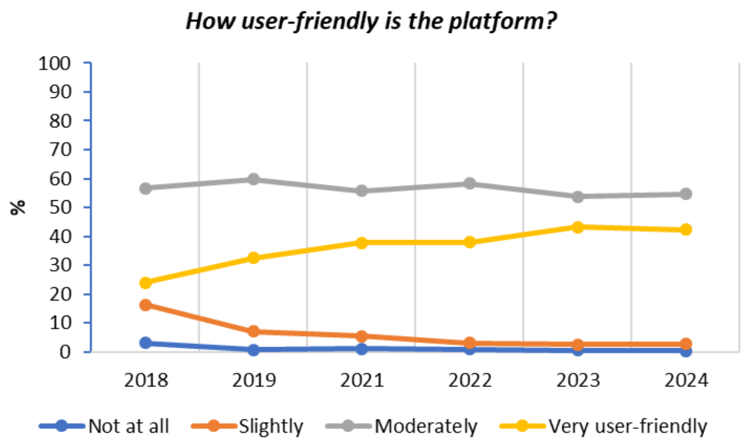
Similar percentages were observed for each domain of the questionnaire, indicating a significative positive trend over the years (Table 1). This is a sign that the improvements made to the FaD courses were in line with the learners' requests.

Table 1 – *Kruskal-Wallis Test Results by Domain.*

Domain	H statistic	df	p-value	Significant (p<0.05)?
Platform usability	721.29	5	0.0001	Yes
Course usefulness	244.62	5	0.0001	Yes
Topics to add	342.10	5	0.0001	Yes
Support from teachers	51.50	5	0.0001	Yes
Overall satisfaction	462.42	5	0.0001	Yes

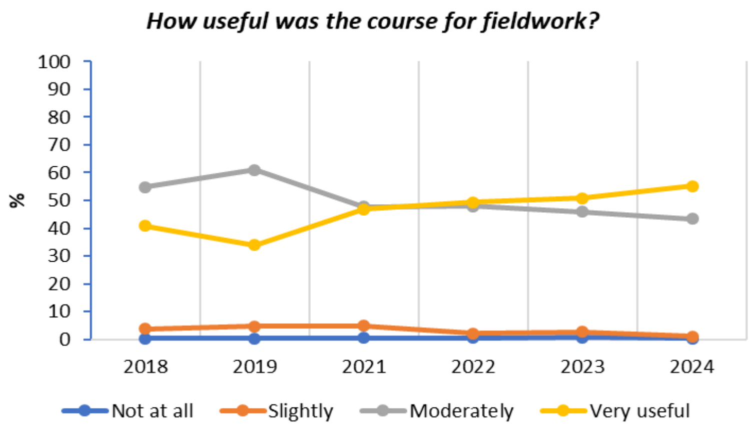
As shown in Figure 3, the majority of respondents defined the platform as either ‘very’ or ‘moderately’ user-friendly. These percentages increased statistically over the years (Kruskal-Wallis test, $p < 0.001$). Dunn’s test revealed that all pairwise comparisons between years were statistically significant ($p < 0.001$), except for the comparison between 2021 and 2022 ($p = 0.046$) and between 2023 and 2024 ($p = 0.270$), which were not significant. Notably, the proportion of respondents who rated it as ‘very user-friendly’ or ‘moderately user-friendly’ significantly increased from 80.56% in 2018 to 96.96% in 2024.

Figure 3 – *Results to the question “How user-friendly is the platform?” by years.*



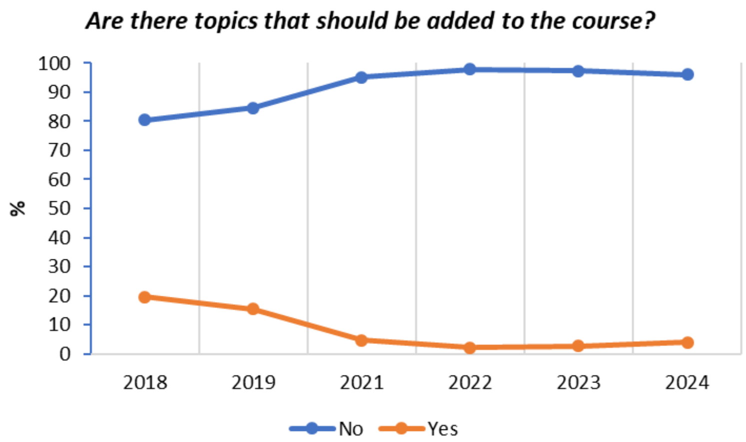
In answer to the question “*How useful was the course for fieldwork?*” (Figure 4), 40.87% of the 2018 sample answered ‘very useful’. This percentage rose to 55.13% in 2024 (Kruskal-Wallis test, $p < 0.0001$). Dunn’s test indicated that all pairwise comparisons between years were statistically significant ($p < 0.001$), except for the comparison between 2021 and 2022 ($p = 0.179$).

Figure 4 – *Results to the question "How useful was the course for fieldwork?" by years.*



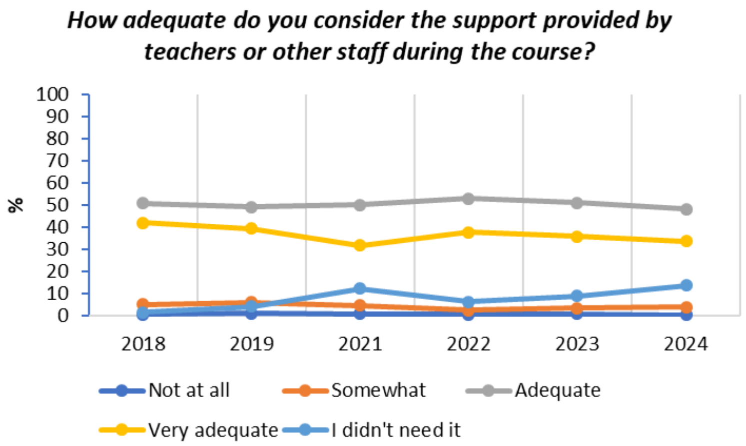
As shown in Figure 5, while in 2018 as many as one-fifth of the respondents (19.49%) believed that 'There were topics to be added to the FaD' this percentage dropped significantly over the years to 3.98% in the year 2024 (Kruskal-Wallis test, $p < 0.0001$). Dunn's *post hoc* test showed that most pairwise comparisons were statistically significant ($p < 0.001$), except for 2021 *vs.* 2024 ($p = 0.171$), 2022 *vs.* 2023 ($p = 0.178$), and 2023 *vs.* 2024 ($p = 0.0586$).

Figure 5 – *Results to the question "Are there topics that should be added to the course?" by years.*



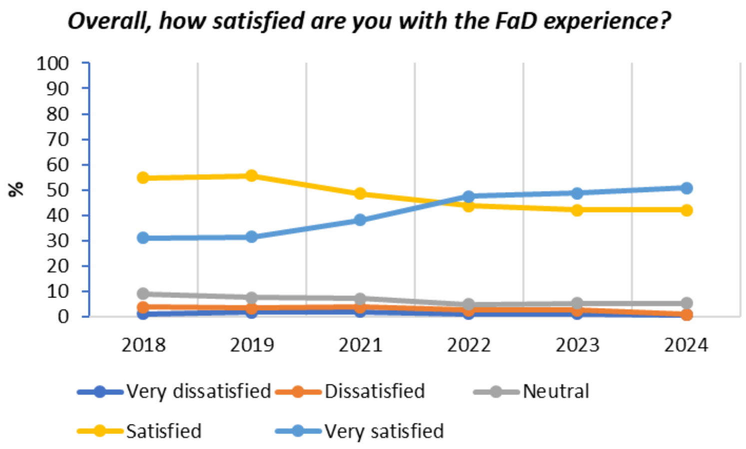
Regarding the support received from teachers or other staff during FaD (Figure 6), it is interesting to note that the proportion of respondents who said they did not need support has increased over the years (from 1.50% in 2018 to 13.65% in 2024; Kruskal-Wallis test, $p < 0.0001$). Dunn's test demonstrated that most pairwise comparisons were statistically significant ($p < 0.001$), except for 2018 *vs.* 2019 ($p = 0.343$), 2021 *vs.* 2022 ($p = 0.139$), 2021 *vs.* 2023 ($p = 0.441$), and 2022 *vs.* 2023 ($p = 0.170$). This is undoubtedly a sign that learners are becoming familiar with the IT platform.

Figure 6 – *Results to the question "How adequate do you consider the support provided by teachers or other staff during the course?" by years.*



There was also a statistically significant increase in overall satisfaction with FaD (Figure 7). The proportion of users who said they were 'very satisfied' increased from 31.14% in 2018 to 50.85% in 2024 (Kruskal-Wallis test, $p < 0.0001$). Dunn's test showed that most pairwise comparisons were statistically significant ($p < 0.001$), except for 2018 *vs.* 2019 ($p = 0.235$) and 2022 *vs.* 2023 ($p = 0.214$).

Figure 7 – *Results to the question "Overall, how satisfied are you with the FaD experience?" by years.*



6. Discussion and conclusions

The transition to the new Permanent Census of Population and Housing model replacing the traditional one required a new planning of the training strategy. Moreover, an unpredictable condition such as the Covid-19 pandemic emergency, led to an urgent rethinking of the training model to continue to let the trainees achieve satisfying levels of competencies, necessary to perform the tasks required for national statistical surveys.

The need for redefinition of the strategy year after year led to the adoption of a circular approach, starting with the annual evaluation of the previous outcome through satisfaction questionnaires, investigation of the emerging needs through the analysis of the answers to both closed and open-ended questions with suggestions or negative comments provided by the interviewers. This evaluation led to a redesign and new implementation of the training program in order to support trainees in their professional performance.

So, the present study, focused on the analysis of the effectiveness of the training models for both interviewers and staff, allows us to highlight that from the first year of this new approach (2018) to today (2024), there was a favorable evaluation of various aspects of the self-e-learning model and a positive trend regarding platform usability, availability of useful materials and teachers support. This approach also led the system to become more user-friendly, flexible and familiar, therefore ready to adapt to the training needs for other national statistical surveys.

Several potential biases should be considered when interpreting the findings of this study. First, data were collected through self-reported questionnaires, which may introduce social desirability bias or response bias. Second, the relatively low response rate could limit the representativeness of the sample and increase the risk of non-response bias. These factors may affect the generalizability of the results. Nevertheless, the consistency of trends observed across multiple years provides some reassurance regarding the robustness of the findings. Future studies should explore strategies to increase participation and assess whether improvements in satisfaction translate into better field performance.

In summary, the following results were obtained: positive growth in satisfaction rate of training courses was observed over time; the need for support to increase familiarity with the platform was limited; complaints about missing content decreased and alignment between training needs and delivery improved.

References

CASTAGNA M. 2022. *L'analisi delle esigenze: dal fabbisogno all'intervento formativo*. Milano: FrancoAngeli.

GOERMAN P., MEYERS M., GARCÍA TREJO Y.A. 2019. *Development of Bilingual Interviewer Training at the U.S. Census Bureau*. Presentation at Comparative Survey Design & Implementation (CSDI) 2019 International Workshop, Warsaw, Poland, 18-20 March.

GUBRIUM J.F., HOLSTEIN J.A., MARVASTI A.B., MCKINNEY K.D. 2012. *The SAGE Handbook of Interview Research: The Complexity of the Craft*. Thousand Oaks, California, United States: SAGE.

ISTAT. 2012. *Linee guida per la qualità dei processi statistici*. Rome. On https://www.istat.it/it/files/2010/09/Linee-Guida-Qualit%C3%A0-_v.1.1_IT.pdf

ISTAT. 2019. *L'approccio trasversale alla formazione delle reti di rilevazione.* Rome. On https://www.istat.it/it/files//2020/05/approccio-trasversale-alla-formazione-delle-reti-di-rilevazione-2020.pdf

ISTAT. 2024. *L'approccio circolare alla formazione delle reti di rilevazione del primo ciclo del Censimento permanente della popolazione e delle abitazioni 2018-2021*. Rome. On https://www.istat.it/wp-content/uploads/2024/10/Formazione-Censimento-Permanente-2018-2021-Ebook.pdf

KOCKELKOREN S. 2011. *Statistical Methods: Organisation of fieldwork/Interviewer training*. The Hague/Heerlen: Statistics Netherlands.

MIRANTE M., BERNARDINI A., BALÌ N. 2025. Training model and interviewer performance in the permanent population Census. *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXXIX, No.2, pp. 257-66.

SUN H., SMITH R., JOHNSON K. 2024. Comparing in-person, video-conference, and self-administered interviewer training: Effects on performance and data quality. *Journal of Survey Statistics and Methodology*, Vol. 12, No. 3, pp. 455-472.

WEST B., KREUTER F., JAENICHEN U. 2013. Interviewer effects in face-to-face surveys: a function of sampling, measurement error or nonresponse? *Journal of Official Statistics*, Vol. 29, No.2, pp. 277-97.

Nunzia BALÌ, Istat, bali@istat.it
Gabriella D'AMBROSIO, Istat, gabriella.dambrosio@istat.it
Maria Teresa FIORI, Istat, mtfiori@istat.it
Alessandra LUGLI, Istat, lugli@istat.it
Nadia MIRANTE, Istat, nadia.mirante@istat.it

OPTIMISM AND PARTICIPATION IN BREAST CANCER SCREENING: EVIDENCE FROM THE UNITED STATES

Silvana Robone, Michela Meregaglia,
Rostand Arland Yebetchou Tchounkeu

Abstract. This study investigates the influence of the personality trait of optimism on participation in breast cancer screening in the US. Breast cancer is the most common cancer among women in the US, accounting for approximately 30% of all new cancer diagnoses annually (American Cancer Society, 2025). Beyond sociodemographic factors, recent literature highlights the significant role of psychological factors, norms, and beliefs in the decision to participate in breast cancer screening. However, the impact of optimistic beliefs on screening uptake has not yet been explored. This study aims to address this gap.
We analyse a sample of approximately 4,500 women aged 50 and older from the US Health and Retirement Study (HRS) spanning 2006 to 2020. A dynamic probit panel data model with random effects is estimated, employing Mundlak's (1978) approach to account for correlated individual effects. Our findings indicate that mammography uptake exhibits strong state dependence. Furthermore, optimism negatively influences mammography uptake among younger women (under 61 years) but positively affects the uptake among women aged 70 and over. These results can be interpreted through the lens of Prospect Theory (Kahneman and Tversky, 1979; Rothman and Salovey, 1997) and Socioemotional Selectivity Theory (Carstensen, 1995). Our findings suggest that policymakers should consider age-specific and psychologically tailored messaging strategies to enhance breast cancer screening adherence across diverse population groups.

1. Introduction

In our study we investigate the influence of the personality trait of optimism on the participation in breast cancer screening programs (through the uptake of mammography) in the US. We investigate this research hypothesis by exploiting data from the US Health and Retirement Study (HRS) in 2006-2020.

Breast cancer is the most common female cancer in the United States (US) accounting for about 30% of all new cancer diagnoses each year and representing the second leading cause of death from cancer in women, behind lung cancer. Mortality rates have been decreasing steadily since 1989, with an overall decline of 44% through 2022, thanks to earlier detection and advancements in treatments. However, in recent years, incidence rates have increased by 1% on an annual basis, due to rising risk factors such as overweight (American Cancer Society, 2025).

Breast cancer screening can help detect breast cancer before there are signs or symptoms of the disease, when it is easier to treat it. Mammography using x-ray is the most common screening test for breast cancer. Magnetic resonance imaging may be used to screen women who have an increased risk of developing cancer (CDC, 2024).

Until April 2024, the U.S. Preventive Services Task Force recommended that women aged 50 to 74 years received a mammography every 2 years. The age target was later expanded to include 40-49-year-olds. In 2023, 79.8% of women aged 50-74 years complied with these recommendations. Despite the screening uptake was lower for low-income (71.3%) and low-educated women (69.9%), the Health People Target 2030 of 80.3% has been almost reached (NCI, 2025). Therefore, the US qualifies as a virtuous country for breast cancer screening, and lessons can be drawn from its experience. In particular, identifying (and addressing) factors that influence women's participation in screening may increase screening uptakes elsewhere.

Apart from sociodemographic characteristics, recent literature has shown that psychological factors, norms and beliefs can play an important role in the choice to take up cancer screening. For instance, the study by Prowse *et al*. (2024) shows that the fear of the unknown regarding a possible diagnosis of cancer or abnormal test results, and a general lack of knowledge around screening programmes are relevant factors that negatively affect the participation in cancer screening (including breast) in high-income countries. Similarly, a systematic review (Tavakoli *et al*., 2024) focussed on breast cancer identifies personal health beliefs, knowledge, perceptions, cultural factors, cues to action, motivation, and self-efficacy, among the factors influencing screening practices in women worldwide.

Among the possible psychological traits, the role played by the Big Five personality traits (i.e., neuroticism, extraversion, openness to experience, agreeableness, conscientiousness) in the choice to participate in cancer prevention programs has been extensively investigated in the literature (Bahat, 2021, Le Clainche *et al.*, 2024, Hajek *et al*., 2020, Niedzwiedz *et al.*, 2019). Previous studies investigating screening practices in countries such as France, Germany, Ireland, Israel, and the UK show clear evidence of a positive association between conscientiousness and extraversion and participation in breast cancer screening, while they report mixed and inconclusive findings about the other three personality traits (i.e., agreeableness, neuroticism, and openness to experience). These findings are confirmed by Aschwanden *et al.* (2019) when exploiting HRS data on older adults in the US. They show that higher conscientiousness and higher extraversion are associated with a higher likelihood of undergoing mammography.

Optimism and pessimism, which are defined as the tendencies to expect positive or negative outcomes in life (Carver *et al.*, 2010), are other psychological traits which have been considered in the health economics literature as determinants of health behaviours. With regard to the participation in screening programs, the role played by optimistic beliefs have been investigated by few studies. Oster *et al.*, (2013), who consider the genetic testing for Huntington disease, a hereditary disease with limited life expectancy, show that taking the test is rare, and that individuals who express optimistic beliefs about their health are more likely to be untested. Moyer *et al.* (2008), using a sample of pregnant women in Ghana, find that optimism is negatively correlated with HIV knowledge and positively correlated with having never been tested before pregnancy. However, as far as we know, no study has investigated the effects of optimism on the participation in cancer screening programs before. We aim at filling this gap in the literature by investigating the extent to which the personality trait of optimism influences the participation in breast cancer screening (through the uptake of mammography) in the US.

We investigate our research hypothesis by using data from the HRS in 2006-2020. The empirical analysis exploits information about a sample of women living in the US and aged 50 and older. We estimate a dynamic probit panel data model with random effects, using the Mundlak correction to account for correlated individual effects (Mundlak, 1978; Rice and Robone, 2022). We also estimate heterogenous effects by age. When considering the full sample, our estimates suggest a positive, although not statistically significant, effect of optimism on mammography uptakes. However, when we stratify our sample by age groups, we find that for women aged between 50 and 61 *optimism* has a negative and statistically significant influence on the likelihood of undergoing breast cancer screening, while for women aged 70 and above *optimism* has a positive and statistically significant influence. Additionally, our estimates show that participation in breast cancer screening is path-dependent.

Our study provides an original contribution to the literature on the determinants of screening since it is the first investigating the role played by optimism in choosing to undergo cancer screening. We focus in particular on breast cancer screening (i.e., mammography) and exploit longitudinal information from the HRS. The use of a dynamic panel probit model is another original feature of our study.

2. Data and Estimation Strategy

We exploit eight waves of the HRS collected in the years 2006-2020. The HRS is a longitudinal panel survey comprising a large representative sample of ageing adults (older than 50 years) in the US using biennial interviews. The HRS was launched in 1992 by the Institute for Social Research (University of Michigan). In 2004, a new Psychosocial and Lifestyle Questionnaire was introduced, which is also referred to as the "Leave Behind Questionnaire" because it was left for participants to self-complete after their in-person interview. Since 2006, the “Leave Behind” module has been completed by participants every 4 years, as it has been left to a random 50% of the full sample in each wave (Clarke, 2008).

Our study considers women in couple aged 50 and over and the final sample is made by 4,466 observations. Our dependent variable is *mammography*, a dummy variable equal to 1 if the woman has undergone breast cancer screening (through mammography) in the previous year, and 0 otherwise. Our main independent variable is *optimism*, an index created by averaging the scores obtained using a 6-point Likert scale (from 1 - strongly disagree to 6 - strongly agree) across three different items.[1]

Table 1 - *Descriptive statistics of the main variables.*

Variable	Obs.	Percent	Mean	Std.dev.	Min	Max
Mammography	4466	80%			0	1
Optimism	4466		4.62	1.08	1	6
Religiosity	4466		5.04	1.40	1	6
Locus of control	4466		5.01	1.07	1	6
Age (year)	4466		64.78	10.04	32	93
High_education	4466	57%			0	1
Income (1000 $)	4466		17.77	38.80	0	765
Retired	4466	47%			0	1
Black_other	4466	13%			0	1
Drinking	4466	60%			0	1
Smoking	4466	8%			0	1
No children	4466	4%			0	1
Wealth_total (1000 $)	4466		581.72	1117.76	0	37100
SAH_poor/fair	4466	16%			0	1
SAH_good	4466	32%			0	1
SAH_Vgood_excellent	4466	52%			0	1
No_insurance	4466	28%			0	1
Neuroticism	4466		0.54	0.14	0.25	1
Extraversion	4466		0.81	0.14	0.30	1
Agreeableness	4466		0.91	0.10	0.25	1
Conscientiousness	4466		0.85	0.10	0.35	1
Openness	4466		0.74	0.14	0.25	1

The potential confounders include sociodemographic covariates (i.e., age, education, income, retired, black or other ethnic groups, no children, and log of total wealth), health behaviours (drinking, smoking), self-assessed health (“poor and fair health”, “good health” and “very good and excellent health”), having no health insurance, the Big Five personality traits (i.e., neuroticism, extraversion,

[1] Q19g (*I’m always optimistic about my future*), Q19h (*In uncertain times, I usually expect the best*) and Q19i (*Overall, I expect more good things to happen to me than bad*).

agreeableness, conscientiousness, and openness) and beliefs (religiosity and locus of control).

Table 1 shows the descriptive statistics of the main variables included in our econometric model. Women in the sample are on average almost 65 years old, 57% of them declare to have a high education level, 47% are retired, their mean income is about $17,770 and their mean total wealth is about $581,720. About 80% of the women in the sample undergo breast cancer screening, which is in line with US data (NCI, 2025), and their average level of optimism is 4.6 on a scale from 1 to 6. In the sample, 16% declares to be in poor or fair health, 32% in good health, and 52% in good or excellent health; 28% declares to have no health insurance.

Our baseline model is a panel probit model over the period 2006-2020 (corresponding to 8 waves), defined as follows:

$$Y_{it} = \beta\ Optimism_{it} + X'_{it}\gamma + \pi_t + \eta_i + \varepsilon_{it}, \tag{1}$$

where Y_{it} denote a binary outcome for the uptake of mammography for the i-th women at time t (i=1, …, N, and t=1, … ,T) . X'_{it} represent women characteristics; π_t is a fixed time effect; η_i is a women-specific random component and ε_{it} is an idiosyncratic time varying error term which is assumed to follow a standard normal distribution. To account for the possibility that the observed regressors are correlated with the unobserved individual effects η_i, we exploit the Mundlak (1978) approach. This approach allows to address this potential correlation by modelling individual-specific effects as a function of the means of the time-varying regressors.[2] Following Wooldridge (2005), the distribution of the individual effects is parameterized as:

$$\eta_i = \theta_0 + \theta_1 Y_{it} + \theta_2 \bar{X}_i + \mu_i, \tag{2}$$

where $\bar{X}_i$ is the average in the sample of the observations on the time-varying women characteristics and μ_i is assumed to be distributed N (0, σ_μ^2), independent of the regressors, the idiosyncratic error term ε_{it} and the initial conditions.

We further estimate a dynamic panel probit over the same time span, by including Y_{it-1} in our specification:

$$Y_{it} = \alpha Y_{it-1} + \rho Y_{i1} + \beta\ Optimism_{it} + X'_{it}\gamma + \pi_t + \eta_i + \varepsilon_{it}, \tag{3}$$

To correct for the initial conditions problem, we adopt the Wooldridge's (2005) approach and include Y_{i1}, the observed values of mammography uptake in the first

[2] The Mundlak (1978) approach is applied by including the within-individual mean of the time-varying regressors.as regressors in our specification.

wave, in the model's specification. Estimating a dynamic panel probit allows us to model *true state dependence*, where past outcomes (e.g., prior uptakes of mammography) directly influence current outcomes, beyond what is explained by observed and unobserved factors.[3] We adopt the Mundlak's (1978) approach even when estimating equation (3).

3. Results

Table 2 displays the baseline results of the estimates of equation (1), investigating the effects of optimism on the uptake of breast cancer screening. The results reveal that, when running our regression model using the entire sample, there is a positive but statistically insignificant relationship between optimism and undergoing a mammography test. However, when the sample is stratified by age groups, the findings vary considerably.[4] For women aged less than 70, the association between optimism and mammography uptake is negative and not statistically significant; however, for women aged 70 and above, the relationship becomes positive and strongly statistically significant (at 1% level).

Table 3 presents the main findings from a dynamic panel probit model based on equation (3). There appears to be a positive and strongly and significant relationship between having undergone a mammography test in the previous wave and the likelihood of undergoing the test in the current wave, particularly for women over 61. Therefore, the uptake of mammography seems to be strongly path-dependent. The results concerning *optimism* align with the baseline findings, although the (negative) coefficient for optimism for women under 61 becomes statistically significant at the 5% level, while the (positive) coefficient for women 70 and above reduces its level of statistical significance at the 5% level. Some covariates, such as *locus of control* and *ethnicity*, also show some positive association with mammography testing.

We also compute the average marginal effects from the fitted model on the probability of undertaking mammography (detailed results are available upon request). For women under 61, a one-point increase on the optimism scale [1, 6] reduces the probability of having the mammography by about 4%, while for women 70 and above it increases such probability by about 3%. These effects are small but not negligible.

[3] A dynamic panel probit models to study cancer screening behaviour has been already adopted, for instance, by Carney et al. (2013).
[4] The three groups are created based on the age variable distribution so that each groups contains approximately one third of the total sample.

Table 2 - *Probit model for mammography uptakes.*

	Full sample b/se	age < 61 b/se	61 <= age <= 69 b/se	age> 69 b/se
Optimism	0.033	-0.067	-0.039	0.155**
	(0.036)	(0.071)	(0.089)	(0.055)
Religiosity	0.049+	0.088+	0.099	-0.043
	(0.027)	(0.047)	(0.068)	(0.046)
Locus of control	0.098**	0.007	0.192+	0.144**
	(0.037)	(0.076)	(0.098)	(0.053)
Age (in 100 years)	-1.781***			
	(0.508)			
High_education	0.148+	0.259	0.254	-0.067
	(0.087)	(0.164)	(0.201)	(0.125)
Income (10k)	-0.013	-0.022	0.022	-0.072
	(0.016)	(0.021)	(0.056)	(0.103)
Retired	-0.044	0.048	-0.439	-0.100
	(0.125)	(0.311)	(0.285)	(0.240)
Black_other	0.272*	0.221	0.733*	0.329
	(0.120)	(0.196)	(0.343)	(0.218)
Drinking	0.272+	0.155	-0.451	0.685**
	(0.150)	(0.319)	(0.400)	(0.257)
Smoking	0.254	-0.442	0.309	0.715
	(0.268)	(0.579)	(0.646)	(0.567)
No children	0.080	0.244	-0.011	-0.082
	(0.200)	(0.335)	(0.493)	(0.340)
Log total wealth	0.130*	0.215+	0.233	0.054
	(0.060)	(0.111)	(0.169)	(0.115)
Sah_good	-0.035	0.280	-0.082	-0.272
	(0.121)	(0.250)	(0.320)	(0.197)
Sah_fair/poor	-0.155	0.691	-0.667	-0.569*
	(0.178)	(0.422)	(0.540)	(0.262)
No_insurance	0.044	0.050	0.208	0.369
	(0.143)	(0.266)	(0.280)	(0.919)
Neuroticism	0.572*	0.173	-0.328	1.102*
	(0.277)	(0.536)	(0.691)	(0.432)
Extraversion	0.387	1.469*	-0.260	-0.214
	(0.354)	(0.661)	(0.861)	(0.549)
Agreeableness	0.106	0.229	0.121	0.535
	(0.431)	(0.786)	-1.072	(0.684)
Conscientiousness	0.080	-0.576	1.536	-0.051
	(0.412)	(0.799)	-1.023	(0.622)
Openness	-0.099	0.042	-0.927	0.051
	(0.342)	(0.656)	(0.837)	(0.506)
constant	-1.088	-1.837+	-2.734+	-2.495**
	(0.670)	-1.114	-1.528	(0.942)
aic	4.081.230	1.653.772	1.014.089	1.529.378
bic	4.273.357	1.813.158	1.160.618	1.683.635
Number of observations	4,466	1,801	1,156	1,509

*Notes: *** p <.001, ** p < .01, * p < .05, + p < .1; standard errors in parentheses. The specification includes as controls also Mean_Income_10k, Mean_Lwealth_total Mean_Retired, Mean_Drinking Mean_Smoking, Mean_SAH_good, Mean_SAH_fair/poor, Mean_No_Insurance.*

Table 3 - *Dynamic probit model for mammography uptakes.*

	Full sample b/se	age < 61 b/se	61 <= age <= 69 b/se	age> 69 b/se
Optimism	0.012	-0.337*	-0.014	0.110*
	(0.043)	(0.168)	(0.079)	(0.053)
Religiosity	-0.024	-0.181+	0.036	-0.012
	(0.031)	(0.100)	(0.057)	(0.040)
Locus of control	0.120**	0.112	0.218*	0.100+
	(0.045)	(0.143)	(0.088)	(0.053)
Age (in 100 years)	-3.631***			
	(0.662)			
High_education	-0.027	-0.022	0.083	-0.123
	(0.087)	(0.253)	(0.154)	(0.108)
Income (10k)	-0.031	-0.063+	0.029	-0.097
	(0.020)	(0.038)	(0.040)	(0.090)
Retired	0.080	0.180	-0.343	0.292
	(0.155)	(0.548)	(0.242)	(0.234)
Black_other	0.347*	0.125	0.488+	0.444+
	(0.149)	(0.322)	(0.268)	(0.231)
Drinking	0.369*	0.237	-0.113	0.544*
	(0.179)	(0.503)	(0.323)	(0.235)
Smoking	0.119	-0.262	0.084	0.681
	(0.336)	(0.949)	(0.520)	(0.531)
No children	-0.417*	-0.610	-0.128	-0.268
	(0.202)	(0.469)	(0.340)	(0.350)
Log total wealth	0.110	0.365	0.095	0.058
	(0.075)	(0.226)	(0.136)	(0.108)
Sah_good	0.034	0.527	-0.068	-0.095
	(0.142)	(0.407)	(0.254)	(0.192)
Sah_fair/poor	-0.069	1.779*	-0.472	-0.334
	(0.208)	(0.832)	(0.416)	(0.254)
No_insurance	-0.079	-0.267	0.215	-0.951
	(0.183)	(0.453)	(0.241)	(0.845)
Neuroticism	0.265	-0.066	-0.064	0.698+
	(0.313)	(0.849)	(0.557)	(0.410)
Extraversion	0.512	0.620	1.225+	0.293
	(0.388)	-1.085	(0.696)	(0.516)
Agreeableness	0.278	0.489	0.325	0.321
	(0.476)	-1.219	(0.853)	(0.640)
Conscientiousness	0.247	0.520	0.244	0.307
	(0.492)	-1.470	(0.856)	(0.628)
Openness	-0.615	0.245	-2.229**	0.081
	(0.378)	-1.039	(0.723)	(0.474)
Mammography (t-1)	0.688***	0.555	0.857***	0.779***
	(0.184)	(0.692)	(0.241)	(0.179)
Mammography (t=1)	0.861***	1.694	0.521*	0.603**
	(0.243)	-1.084	(0.238)	(0.186)
constant	0.178	-0.100	-2.505*	-3.218***
	(0.809)	-1.748	-1.251	(0.959)
aic	1.639.424	346.508	476.820	877.415
bic	1.818.844	475.831	614.686	1.026.358
Number of observations	2,012	479	631	902

*Notes: *** p <.001, ** p < .01, * p < .05, + p < .1; standard errors in parentheses. The specification includes as controls also Mean_Income_10k, Mean_Lwealth_total Mean_Retired, Mean_Drinking Mean_Smoking, Mean_SAH_good, Mean_SAH_fair/poor, Mean_No_Insurance.*

4. Discussion

In our study we investigate the influence of the personality trait of optimism on the choice to participate in breast cancer screening programs and uptake a mammography. We investigate our research hypothesis by considering a sample of about 4,500 women aged 50 and older from the US HRS in 2006-2020.We estimate a dynamic panel probit model, adopting the Mundlak's (1978) approach to account for the possibility that the observed regressors are correlated with the unobserved individual effects. The uptake of mammography seems to be strongly path-dependent. Optimism appears to have negative influence on mammography uptakes for younger women (aged less than 61), while it appears to have a positive influence for older women, aged 70 and above.

There are different possible explanations for the negative influence of optimism on mammography uptake we find for younger women in our sample. First of all, optimistic individuals often exhibit a "positive illusion" or "optimism bias", meaning they believe they are less likely than others to experience negative events (Weinstein, 1980). This leads them to underestimate their personal risk of diseases like breast cancer. Moreover, they are less influenced by fear-based appeals (Aspinwall and Brunhart, 1996). This psychological profile may interact with the way mammography is promoted, which in the US is typically through a negative framing. According to Rothman and Salovey (1997), health behaviours can be divided into two broad categories: prevention behaviours (e.g., using sunscreen, exercising) and detection behaviours (e.g., mammography, HIV testing). On the basis of Prospect Theory (Kahneman and Tversky, 1979), they argue that detection behaviours are more effectively encouraged through loss-framed messages, which emphasize the risks of not performing the behaviour.[5] Optimistic individuals, however, tend to be less responsive to such negatively framed messages, as their general outlook leads them to downplay threatening information. As a result, standard screening campaigns, which often emphasize the potential loss associated with not getting screened, may fail to motivate the group of the optimists effectively.

The negative influence of optimism on mammography uptakes can also be explained in the light of the *information avoidance* mechanism, well-illustrated by Golman *et al.* (2017). The theoretical model of Brunnermeier and Parker (2005) shows that individuals can choose to hold optimistic beliefs, which are a source of anticipatory utility and thus improve immediate well-being, potentially at the risk of intensifying future disappointment. In this context, individuals might indulge in

[5] The predictions of Rothman and Salovey (1997) are confirmed, for example, by a study conducted in Sicily (Italy) by Bertoni *et al*. (2020). They show that loss-framed letters, which emphasize the potential negative consequences of not getting a mammography, stimulate higher attendance to breast cancer screening program.

information avoidance, because acquiring information can interfere with their ability to maintain unwarranted optimism. With regard to health, information avoidance might induce individuals at risk for health conditions to eschew medical tests. Evidence of this behaviour are provided, for instance, by Oster *et al*. (2013) when considering people at risk of Huntington's disease.

With regard to the positive influence of optimism on mammography uptake among women aged 70 and above, one possible explanation lies in age-related shifts in motivational orientation. According to the Socioemotional Selectivity Theory (Carstensen, 1995), older adults increasingly prioritize emotionally meaningful goals and health maintenance. Therefore, optimistic women in this age group may be more motivated to engage in preventive care to preserve their quality of life and autonomy. Moreover, research suggests that optimism enhances health behaviour when individuals perceive a high degree of personal control (Aspinwall and Taylor, 1992). Older women may see mammography as an empowering and manageable step toward maintaining their wellbeing. Finally, this effect may also partly reflect a survivor or selection bias, whereby healthy ageing women with a generally proactive health orientation, besides higher optimism, continue to engage with screening services.

From a policy perspective, we suggest policy makers to promote breast cancer screening initiatives which employ gain-framed, positively oriented messages to effectively reach younger women, who are less responsive to fear-based appeals. Differently, for older women, communications that emphasize empowerment and personal control are more advisable. In general, the adoption of an age and psychologically tailored messaging approach could enhance the screening adherence across diverse population groups.

Acknowledgements

We acknowledge financial support by the EU Next Generation, Mission 4 Component 1, PRIN Project 2022CZKPMY - Health and Wellbeing of the Elderly, (CUP J53D23004470006).

References

AMERICAN CANCER SOCIETY. 2025. Key Statistics for Breast Cancer. https://www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html (last accessed 30 June 2025).

ASCHWANDEN D., GEREND M.A. LUCHETTI M., STEPHAN Y., SUTIN A.R., TERRACCIANO A. 2019. Personality traits and preventive cancer screenings in the Health Retirement Study, *Preventive Medicine*, Vol. 126.

ASPINWALL L. G., BRUNHART S. M. 1996. Distinguishing optimism from denial: Optimistic beliefs predict attention to health threats. *Personality and Social Psychology Bulletin*, Vol. 22, No. 6, pp. 993–1003.

ASPINWALL, L. G., TAYLOR, S. E. 1992. Modeling cognitive adaptation: A longitudinal investigation of the impact of individual differences and coping on college adjustment and health. *Journal of Personality and Social Psychology*, Vol. 63, No. 6, pp. 989–1003.

BAHAT E. 2021. The Big Five personality traits and adherence to breast cancer early detection and prevention, *Personality and Individual Difference*s, Vol. 172.

BERTONI M., CORAZZINI L., ROBONE S. 2020. The good outcome of bad news. A Field Experiment on Formatting Breast Cancer Screening Invitation Letters, *American Journal of Health Economics*, Vol. 6, No. 3, p.p. 372-409.

BRUNNERMEIER M.K., PARKER J.A. 2005. Optimal Expectations. *American Economic Review*, Vol. 95, No. 4, p.p. 1092–118.

CARNEY P., O'NEILL S., O'NEILL C. 2013. Determinants of breast cancer screening uptake in women, evidence from the British Household Panel Survey, *Social Science & Medicine*, Vol. 82, pp. 108-114.

CARVER C.S., SCHEIER M.F., SEGERSTROM S.C., 2010. Optimism, *Clinical Psychology Review*, Vol. 30, No. 7, pp. 879-889.

CENTRE FOR DISEASE CONTROL (CDC). 2024. Screening for Breast Cancer.

CLARKE P., FISHER G., HOUSE J., SMITH J., WEIR D. 2008. https://hrs.isr.umich.edu/sites/default/files/biblio/HRS2006LBQscale.pdf.

CARSTENSEN L.L. 1995. Evidence for a life-span theory of socioemotional selectivity. *Current Directions in Psychological Science*, Vol. 4, No. 5, pp. 151–156.

GOLMAN R., HAGMANN D., LOEWENSTEIN G. 2017. Information avoidance, *Journal of Economic Literature*, Vol. 55, No. 1.

HAJEK A., KRETZLER B., KÖNIG H-H. 2020. Personality and the use of cancer screenings. A systematic review, *PLOS One*, Vol. 15, No. 12.

KAHNEMAN D., TVERSKY A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, Vol. 47, No. 2, pp. 263–291.

LE CLAINCHE C., MARSAUDON A., ROCHAIX L., HAON B., VERGNAUD J-C. 2024. Do Behavioral Characteristics Influence the Breast Cancer Diagnosis Delay? Evidence From French Retrospective Data, *Value Health*, Vol. 27, No. 10, pp. 1408-1416.

MOYER C.A., EKPO G., CALHOUN C.L., GREENE J., NAIK S., SIPPOLA E., *et al.* 2008. Quality of life, optimism/pessimism, and knowledge and attitudes toward

HIV Screening among pregnant women in Ghana, *Women's Health Issues*, Vol. 18, No. 4, pp. 301-309.

MUNDLAK Y. 1978. On the Pooling of Time Series and Cross Section Data. *Econometrica*, Vol. 46, No. 1, pp. 69-85.

NATIONAL CANCER INSTITUTE. 2025. Breast cancer screening.

NIEDZWIEDZ C.L., ROBB K.A., VITTAL KATIKIREDDI S., PELL J.P., SMITH D.J. 2019. Depressive symptoms, neuroticism, and participation in breast and cervical cancer screening: Cross-sectional and prospective evidence from UK Biobank, *Psychooncology*, Vol. 29, No. 2, pp. 381–388.

OSTER E., SHOULSON I., RAY DORSEY E. 2013. Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease, The American Economic Review, Vol. 103, No. 2, pp. 804-830.

PROWSE S.R., BRAZZELLI M., TREWEEK S. 2024. What factors influence the uptake of bowel, breast and cervical cancer screening? An overview of international research, *European Journal of Public Health*, Vol. 34, No. 4, pp. 818–825.

RICE N., ROBONE S. 2022. The effects of health shocks on risk preferences: Do personality traits matter? *Journal of Economic Behavior and Organization*, Vol. 204, pp. 356-371.

ROTHMAN A.J., SALOVEY P. 1997. Shaping perceptions to motivate healthy behavior: the role of message framing, *Psychological Bulletin*, Vol. 121, No. 1, pp. 3-19.

TAVAKOLI B., FEIZI A., ZAMANI-ALAVIJEH F., SHAHNAZI H. 2024. Factors influencing breast cancer screening practices among women worldwide: a systematic review of observational and qualitative studies, *BMC Women's Health*, Vol. 24, No. 268.

WEINSTEIN N. D. 1980. Unrealistic optimism about future life events, *Journal of Personality and Social Psychology*, Vol. 39, No. 5, pp. 806–820.

WOOLDRIDGE J.M. 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity, *Journal of Applied Econometrics*, Vol. 20, No. 1, pp. 39-54.

Silvana ROBONE, University of Eastern Piedmont, Dondena Università Bocconi, HEDG University of York, Catholic University, silvana.robone@uninsubria.it
Michela MEREGAGLIA, CERGAS Bocconi, michela.meregaglia@unibocconi.it
Rostand Arland YEBETCHOU TCHOUNKEU, University of Eastern Piedmont, rostand.yebetchou@uniupo.it

TOOLS OF BUSINESS INTELLIGENCE FOR MANAGING ISTAT INTERNAL PROCESSES[1]

Alessandra Dentini, Amanda Muratori, Claudio Massimiliano Clemente

Abstract. In this paper, the analysis of data extracted from the Italian National Institute of Statistics (Istat) administrative database, called Urbi Smart (a structured collection of data and information managed by an organization, accessible only by its employees and used for business purposes, meaning for internal use) is used to showcase the power of the Business Intelligence for increasing the efficiency of the management system. As part of Business Intelligence process, organizations collect data from internal IT systems and external sources, prepare it for analysis, run queries against the data and create data visualizations, Business Intelligence dashboards and reports to make the analytics results available to business users for operational decision-making and strategic planning.

The information on welfare benefits selected from Istat administrative database consists of a series of variables that include employee identification data (registration number, profile, personnel plan), administrative data relating to subsidy applications and accounting data (amount submitted, amount allocated).

The goal of the analysis is to support a rational and equitable distribution of the funds allocated for employee welfare benefits. By examining patterns in subsidy requests and allocations, the BI process enables evidence-based decisions that improve transparency, optimize resource use, and strengthen the overall management system.

1. Introduction

Business Intelligence is a set of processes, technologies and tools that analyse company data (both historical and current) to generate useful information and support decision-making. It is used to better understand the current and historical performance of the company in order to identify opportunities, optimise processes and predict future scenarios, enabling more informed and strategic decision-making.

Its benefits include greater operational efficiency, improved decision-making, the identification of new market opportunities and better control over performance in the public sector. (De Vivo *et al*, 2011).

[1] The paper is the results of the common work of the authors. In particular, A. Dentini has written Sections 4 and 5; A. Muratori has written Sections 1 and 3; C.M. Clemente has written Section 2

Business Intelligence has developed to include more processes and activities to enable performance improvement, including a statistical approach.

- Data Mining: use databases, statistics, and machine learning to uncover trends in large datasets.
- Reporting: sharing data analytics with stakeholders so they can draw conclusions and make decisions.
- Performance metrics and benchmarking: compare current performance data with historical data to monitor performance against goals. Typically, this is done using custom dashboards.
- Descriptive analytics: using preliminary data analytics to understand what happened.
- Query execution: querying data with specific questions, for which Business Intelligence extracts answers from datasets.
- Statistical analysis: starting from the results of the descriptive analysis, further exploration of the data using statistics, for example in relation to how and why a certain trend has occurred.
- Data visualization: transform data analysis into visual representations, such as graphs, charts and histograms, for easier data consumption.
- Visual analytics: exploration of data through visual representations to communicate information on the fly and follow the flow of analysis.
- Data preparation: compiling various data sources, identifying their dimensions and measurements, and preparing them for data analysis. (Sharda *et al*, 2015).

In the Public Administration (PA), internal administrative sources constitute an invaluable wealth of information. However, their utilization is strongly influenced by the specific nature of the public context, particularly by the principles of legality, transparency and sound administration that govern its actions, as well as by the presence of rigorous legislation concerning privacy and data access.

The welfare benefits information selected from administrative Istat database are made up of several variables which include employee identification data (registration number, profile, staffing plan), number and submission date of subsidy applications, the expenditure incurred and disbursed, ISEE certification availability (if attached) and, for school and university contributions, the reference to the children and the schools or institutions attended. The Equivalent Economic Situation Indicator (ISEE) is a tool that allows for the measurement of the economic condition of families in the Italian Republic.

The data analysis of welfare benefits covers the years 2005-2022, so that it is possible to assess the situation before, during and after the pandemic crisis.

The first phase in building the database involved extracting the activity reports from the Istat management information system, called Urbi Smart, structured by year and contribution.

The statistical analyses used have extracted from a very large dataset of unstructured administrative data the information necessary to design strategic company policies focused on cost savings and employee welfare.

The goal is to carry out a rational distribution of the fund allocated for benefits through the analysis of subsidy requests submitted by employees.

The paper describes one example of the use of Business Intelligence in Istat. Another example is the analysis of employee business travel (Dentini *et al.*, 2022; Dentini and Zeppieri, 2023), in which data purely used for administrative matters are used to obtain useful information for the management of activities. We are talking about a wealth of information to be exploited to increase the effectiveness of some internal processes of the Istat.

The paper is structured as follows. The second section includes the legislation related to corporate welfare and the analysis of the trend in subsidy applications over the years; the third section is focused on the results of the analysis conducted since the database construction; the fourth section is dedicated to the conclusions regarding the assessment of advantages and critical issues of welfare initiatives in the PA.

2. Welfare benefits in the Public Administration

Corporate welfare is the set of initiatives, goods and services made available by employers to improve the well-being of employees and their families, beyond regular remuneration. Welfare in PA, while sharing the basic objectives of corporate welfare in the private sector, presents specific characteristics linked to its public nature and the regulatory framework that governs it. Unlike the private sector, where article 51 of the TUIR (Consolidated Income Tax Act) and the National Collective Labor Agreements (CCNLs) serve as pillars, in the PA, welfare is strictly tied to the National Collective Labor Agreements (CCNLs) of the various sectors. These contracts regulate the possible forms of supplementary welfare.

Welfare services in the PA can be provided directly by the public body, offering services to its own employees (for example internal canteens, direct agreements with gyms or nurseries); or through collective agreements, where CCNLs and supplementary decentralized agreements provide for the activation of specific benefits; alternatively, through the National Institute for Social Security (INPS) - Public Employees Management, which manages a wide range of benefits and services for public employees and their families, often through competitive calls for applications.

Istat is a research institution that offers its employees a series of benefits, structured through a specific internal regulation that outlines their types, requirements and methods of provision.

The granting of social and welfare benefits for the personnel of public bodies is regulated by Presidential Decree n. 509 of 16 October 1979, which extends the aforementioned regulation to all public bodies in the Research Sector through Article 59. The CCNL 2016-2018 introduces the definition of "supplementary welfare": article 96 defines it as "a set of activities and services provided by public entities to their employees with the aim of improving their private and working life."

Each year a fund for social and welfare benefits is established, to which an overall allocation is dedicated. This allocation is determined by considering 1% of the sums recorded in the chapters of the adjusted budget for the reference year that fund the fundamental and ancillary economic treatment of personnel.

The allocation is affected by personnel movements (hirings, terminations, career progressions) that occurred in the previous year.

The establishment of general criteria for the activation of supplementary welfare plans and the rules for granting social and welfare benefits to personnel are entrusted to integrative collective bargaining between the Administration and trade unions. Istat, therefore, adopts each year (through the instrument of decentralized bargaining), these criteria for the distribution and allocation of the identified sums destined for these benefits.

Table 1 – *Welfare Fund allocation Distribution by type of benefit, 2020-2023.*

Type of benefit	2020	2021	2022	2023
School Study Support	105.000	98.157	124.249	166.797
School Scholarships	35.000	34.600	45.790	48.626
University Study Support	30.000	28.200	35.000	43.908
University Scholarships	21.000	20.000	16.000	33.525
Use of Public Transport	21.000	10.200	31.400	39.434
Nursery-Primary School*	107.628	/	96.350	27.507
Healthcare Subsidies	617.970	694.120	884.901	896.605
Summer Camps	16.600	27.400	36.000	35.427
Total Allocated Fund	**919.582**	**911.237**	**1.269.690**	**1.291.829**

Our elaborations on the Istat welfare benefits database
**The 2021 contribution for nurseries, kindergartens and primary schools was deferred to 2022 to align the annual periods with other school contributions.*
*******Due to graphic reasons, the years 2013/2019 can not be inserted in the table.*

Table 1 examines the distribution of funds in the period 2020-2023, showing significant growth since 2022. This growth was influenced, in addition to workforce population movements, by a new computation of the 1% of the sums allocated for staff remuneration.

The entire observed interval, from 2013 to 2023, highlights increases in 2014 and 2017, determined by the increase in staff. Conversely, the declines recorded between 2018 and 2021 correspond to staff terminations, without equivalent recruitment to compensate. School and university scholarships benefits have seen an improvement in the sums allocated to them starting from 2020 and, in particular, in 2023, due to political choices adopted by the Administration in order to reward meritocracy.

The use of public transport benefit was influenced by circulation limits due to the pandemic in the period 2020-2021, with progress in the subsequent two-year period that was not as incisive as in the first part of the entire observed interval. This is a result of the development of remote working organizational forms.

A particular mention is due for nursery - primary schools benefit: a decrease is observed between 2020 and 2022, due to a lower number of children of kindergarten and primary school age; however, in 2023, a change of regulation transferred the primary school benefit to school study support, resulting in the sum allocated here being much lower compared to past years.

Healthcare subsidies benefits attract the highest sum, due to the high number of requests received each year; for this reason, they have always been strengthened, with the exception of 2018, sometimes penalizing the amounts allocated to other benefits.

3. Data source and database building

The Human Resources Department of Istat has been using Urbi Smart as a management information system since 2016. The application system manages various areas of human resources, including the legal and accounting management of welfare benefits due to Istat employees who have applied for them.

To build a statistically informative database for the study of the distribution of the fund allocated for benefits, it is necessary to review the work process, divided into phases, starting from the rationalization of the information present in Urbi to select the variables to be analysed for statistical purposes.

The goal is to transform an administrative database, made up of multiple management variables, mainly used as a repository of administrative data, into a statistical database, capable of making the data classifiable and usable for statistical purposes.

The data analysis covers the years 2005-2022 and the welfare benefits analysed are the following: healthcare subsidies, use of public transport, school study support, school scholarships, university study support, university scholarships, summer camps, subsidy for nursery, kindergarten and primary school attendance.

The welfare benefits information selected from Urbi are made up of several variables which always certainly includes employee identification data (registration number, profile, staffing plan), number and submission date of subsidy applications and, finally, the expenditure incurred and disbursed.

After extracting the activity reports from Urbi information system, structured by year and ~~by~~ contribution, we created a new database for each contribution, consolidating the various annual data and then calculating the total amount disbursed per employee. In the case of school contributions (where the same employee submitted multiple applications per child), we proceeded to count the subtotals per child and then the overall totals per employee.

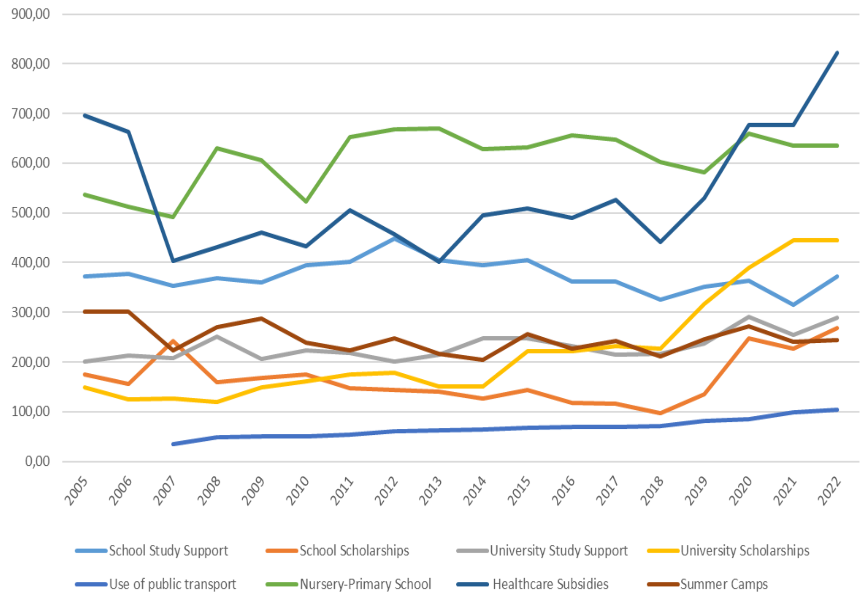
4. Analysis of the results

After building the database, we compute descriptive statistics in order to highlight particular trends. As already mentioned, the goal is to analyse the trends to make strategic decisions regarding the allocation of financial resources. We reported the extracted data for each benefit type (healthcare subsidies, school and university study support, school and university scholarships, nursery-primary school, use of public transport, summer camps) in pivot tables and then we created charts and tables for the analysis of the following measures:

- average amount disbursed per employee;
- number of applications per benefit;
- scatterplot matrix between welfare benefits and average age;
- distribution of school grades levels (first year of middle school and subsequent grades;
- school grant types (middle school degree, intermediate high school grades and high school degree);
- trends in summer stay types (daily and residential summer camps);
- transport modalities.

Figure 1 shows the average financial resources disbursed per applicant employee for each benefit. The trends are fairly constant but there are exceptions that are interesting to focus on. First of all, the healthcare subsidy which has a trend probably influenced by exogenous factors. Indeed, there are two significant moments in the time series: the economic crisis of 2007-2008 and the pandemic of 2020. As if the demand for benefits was a reflection of the times.

Figure 1 – *Average amount disbursed per employee.*

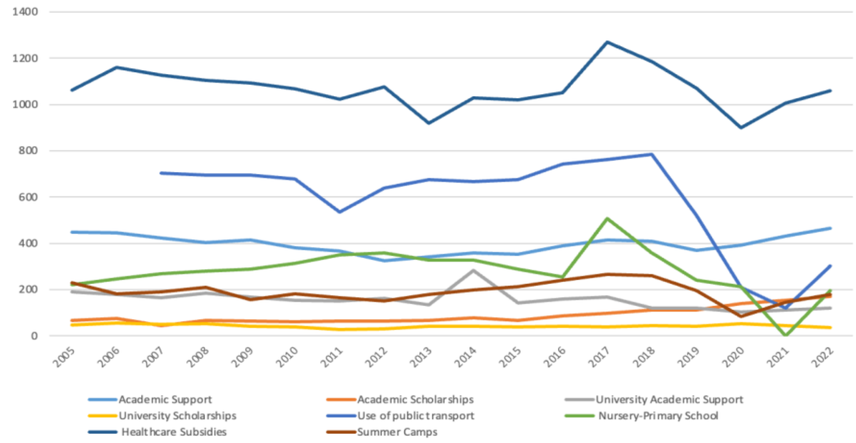


Our elaborations on the Istat welfare benefits database

We know that the average age of Istat employees is progressively increasing and this factor is having an impact on benefit claims. The demand for university scholarships is steadily increasing, especially after 2018; while school study support is decreasing. In fact, there are no young employees; those who are there have few children, and the children of older employees are already attending university. From this Business Intelligence analysis, it is easy to see what will happen soon and where it will be most useful to allocate financial resources.

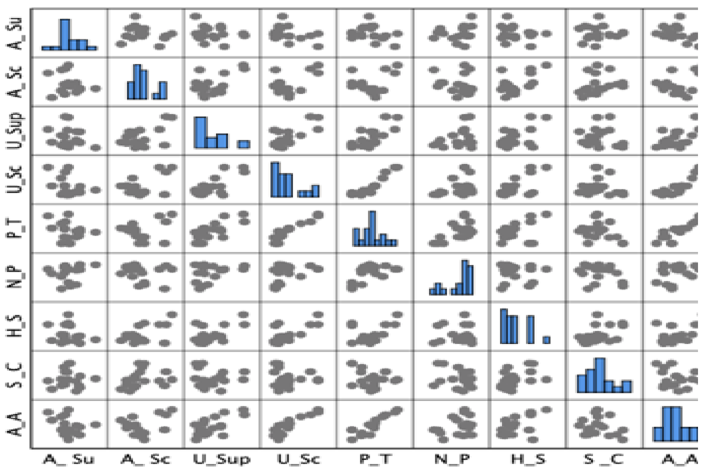
Figure 2 confirms the evidence of Figure 1 and shows very well that applications for subsidies for using public transport plummeted during the pandemic period. This shows that, with the extended use of remote working, the resources for this benefit could be reallocated to other items.

Figure 2 – *Number for application per benefit.*



Our elaborations on the Istat welfare benefits database

Figure 3 – *Scatterplot matrix between welfare benefits and average age.*



A_A	Average Age
A_Su	School Study Support
A_Sc	School Scholarships
U_Sup	University Study Support
U_Sc	University Scholarships
P_T	Use of public transport
N_P	Nursery-Primary School
H_S	Healthcare Subsidies
S_C	Summer Camps

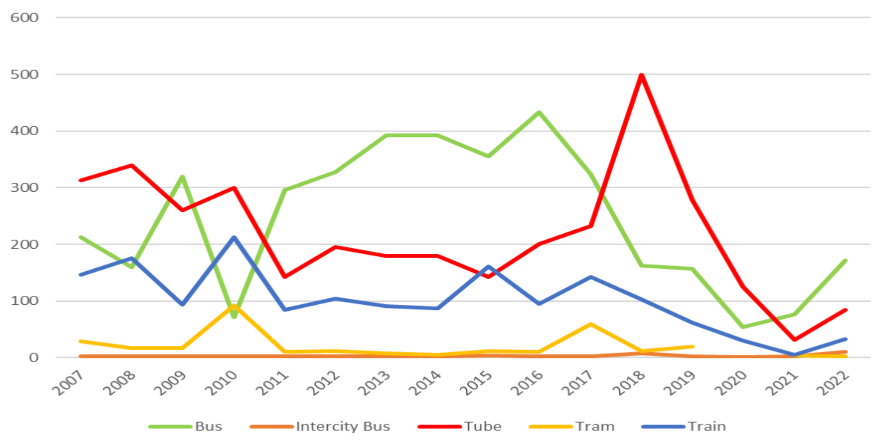
Our elaborations on the Istat welfare benefits database

Figure 3 presents the scatterplot matrix between welfare benefits and the average age of Istat employees. A scatterplot matrix displays pairwise scatterplots for multiple variables, with each scatterplot showing the relationship between two variables. To interpret, identify the row and column of a plot to find its variable axes (row variable for y-axis, column variable for x-axis). Diagonal plots show individual variable distributions, often as histograms. Average age is positively correlated with healthcare benefits, public transport usage and university benefits. In addition, mean age is negatively correlated with summer camps and primary and secondary school subsidies.

Therefore, the results confirm the theory that the rising average age is a parameter that influences the trend of welfare benefit applications and the allocation of welfare benefits in the coming years; unless public competitions are instituted to encourage the entry of younger staff who obviously need different support.

By analysing selected modalities in the Urbi database related to specific applications, it was possible to examine how the working conditions of the Institute's employees changed over the period 2005–2022.

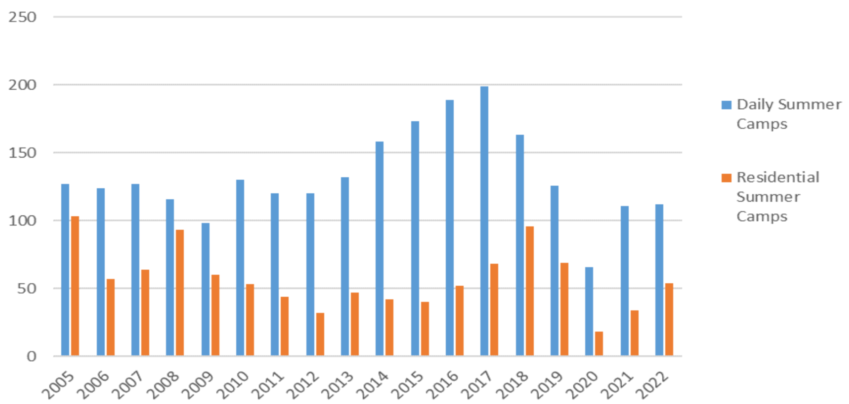
Figure 4 – *Transport modalities (2007-2022).*



Our elaborations on the Istat welfare benefits database

Figure 4 depicts the different means of transport used by Istat employees to travel between home and work. The welfare contribution for the use of public transport has been introduced in the benefits fund since 2007. In recent years, the most frequently used means of transport have been buses and tube. The analysis of the data allows interaction with the mobility manager in order to conclude agreements with the companies providing transport services. There is evidence of a substantial drop in applications in the Covid 19 period, a trend confirmed precisely by the limitation of travel.

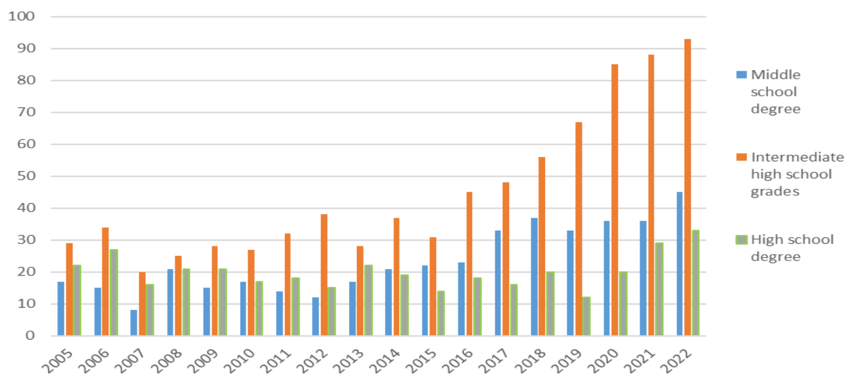
Figure 5 – *Trends in summer stay types (2005-2022)***.**



Our elaborations on the Istat welfare benefits database

Figure 5 analyses the trend of summer stay types (daily and residential summer camps); there is an increase in day-trips until 2017, then they begin to decrease, which became more pronounced during the pandemic.

Figure 6 – *School grant types (2005-2022).*



Our elaborations on the Istat welfare benefits database

In Figure 6, school grants are analysed. Scholarships are distributed at the end of the study cycle and for intermediate classes. Since 2016, applications for school grants for intermediate levels have increased and since 2020, grants for the last year of upper classes will also increase, which suggests that applications for university grants and scholarships will increase in the coming years.

5. Conclusion

The PA is today engaged in a renewal process that focuses on the internal efficiency of each individual administration, on greater transparency and on more accessible, flexible and timely services, for facilitating the relationship between the public service and citizen.

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.

Resource reallocation is an economic concept that refers to the reuse or redistribution of productive assets (such as capital, labour, time, technology) from one activity or sector to another in order to optimise efficiency and productivity.

The distribution of funds for the respective types of welfare benefits provided by Istat to its employees may be modified based on data analysis. In particular, with the rising average age of employees, applications for nursery school have dropped, while those for university support have risen. In such a case, a different allocation of economic resources would be appropriate, considering the changing needs of employees.

Welfare plays a crucial role in PA, as it contributes to greater personal and family well-being, reducing stress and improving work-life balance. Benefits, being often tax-free or contribution-free, allow employees to access services and goods with a higher net value.

A public authority that invests in the welfare of its employees shows a more modern, responsible and people-oriented image. Welfare is a concrete way to show care for employees, reinforcing a sense of belonging and gratitude. A competitive welfare offer can make PA more attractive to professionals, helping to retain internal skills (Yamamoto, 2011).

Despite clear advantages, welfare in PA has faced limitations in the past due to budgetary constraints and regulatory complexities. However, recent interpretations of the Court of Auditors on the expenditure ceiling are opening new perspectives, allowing administrations to invest more in these measures.

It is crucial that public administrations, also through supplementary bargaining, explore and implement welfare plans that respond effectively to the specific needs of their employees, making the most of the opportunities offered by the legislation and available funds.

In conclusion, the analyses developed show that the inexorable ageing of Istat employees combined with remote working must push towards a new allocation of welfare benefit funds. In the future, benefit resources need to be concentrated on higher levels of education (university) and medical expenses, rather than on schooling (from nursery to high school) and public transport. This scenario can only change if new young employees are recruited through public competitions.

References

DE VIVO M.C., POLZONETTI P., TAPANELLI P. 2011. Open Data, Business Intelligence e Governance nella Pubblica Amministrazione, *Informatica e diritto*, Vol. 20, No 1-2, pp. 239-262.

SHARDA R., DELEN D., TURBAN E. 2015. *Business intelligence and Analytics: Systems for decision support.* New Jersey: Pearson.

DENTINI A., MAZZIOTTA M., ZEPPIERI I. 2022. The Destinations of Research Activity in Istat: Statistical Methods of Business Intelligence for Enhancing Territorial Information, *Italian Journal of Economic, Demographic and Statistical Studies*, Vol. 77, No. 1, pp. 125-136.

YAMAMOTO H. 2011. The relationship between employee benefit management and employee retention. *The International Journal of Human Resource Management*, Vol. 22, No. 17, pp. 3550–3564.

Alessandra DENTINI, Istat, aldentin@istat.it
Amanda MURATORI, Istat, amanda.muratori@istat.it
Claudio Massimiliano CLEMENTE, Istat, mcclemente@istat.it

THE NEW ITALIAN CLASSIFICATION OF ECONOMIC ACTIVITIES ATECO 2025: A FOCUS ON INDIVIDUAL ENTREPRENEURS

Francesca Alonzi, Caterina Viviano

Abstract. During the recent review of the Italian classification of economic activities — that led to the release of Ateco 2025 — discussions with stakeholders highlighted concerns regarding the clarity of the classification when it comes to describing the activities of individual entrepreneurs. Individual entrepreneurs often find that the contents of Ateco 2025 do not reflect their activities because of the difficulty of separating the concept of 'what one is' (status) from 'what one does'. Drawing on descriptive examples based on user inquiries and registry data, this study confirms that the application of the Ateco classification to the economic activities of individual entrepreneurs is significantly affected by coding errors, which largely depend on the user's level of knowledge and expertise in applying the classification. It emerged that, in such cases, ordinary users — including the entrepreneurs themselves, accountants, and labour consultants — often confuse the statistical concept of 'occupation' with that of 'economic activity'. This situation undermines the quality of statistics and requires action from the custodians of the classification to enhance its clarity.

1. Introduction

In Italy, there are approximately 4.7 million active enterprises, most of which are small and medium-sized. Of the over 4.5 million enterprises with fewer than nine employees, 67% are individual entrepreneurs.

Economic activity is the primary criterion for classifying economic units, including individual entrepreneurs, in business and economic statistics. The NACE statistical classification is used across Europe, covering more than eighty statistical domains. At the national level, statistics by economic activity must be produced using NACE or a national variant. The Italian classification of economic activities is called Ateco, an acronym for *attività economiche* (economic activities in Italian).

In the Italian statistical business register (SBR), NACE/Ateco codes are assigned to enterprises using both statistical and administrative sources. While larger enterprises are manually profiled by statistical experts, codes for individual entrepreneurs are typically generated automatically from administrative data. Indeed, the Ateco classification is also used by major Italian administrative bodies, such as the Chambers of Commerce and the Tax Agency, for non-statistical purposes. Consequently, incorrect classification of individual entrepreneurs according to the Ateco system can significantly affect the quality of official statistics.

This research presents a discussion on the challenges that individual entrepreneurs face in correctly applying the classification of economic activities. Ordinary users — including individual entrepreneurs, accountants, and labour consultants — often misinterpret the classification due to the difficulty of distinguishing between the concept of ‘what one is’ (status) and ‘what one does’, which can also affect the quality of official economic statistics.

The reason is twofold. First, the terminology used in the classification tends to reflect the activities of larger enterprises, making it less intuitive for individual entrepreneurs. Second, individual entrepreneurs sometimes confuse the classification of economic activities with a separate statistical system that categorizes occupations, further complicating correct interpretation.

To date, and to the best of our knowledge, no official mapping exists between the Italian classification of economic activities and that of occupations, nor are there sufficiently comprehensive technical documents and studies on how individual entrepreneurs should be treated within economic activity classifications. This study represents a first attempt to address the issue and to lay the groundwork for practical guidelines. The analysis seeks not only to identify the key challenges but also to provide insights to support improvements in statistical systems and administrative frameworks.

This study is relevant for scholars and practitioners involved in official statistics and economic classifications, with relevance for international practices and standards. Indeed, similar interpretation issues concerning the activities of individual entrepreneurs may arise in other European countries that use the NACE classification or a derived version. This study also provides guidance for public administrations that classify entities for administrative purposes, primarily for fiscal reasons. The topic also engages with broader debates in labour economics, especially concerning non-standard forms of employment and self-employment, highlighting their implications for labour market analysis and policy.

This research work is presented into five sections. A brief introduction is provided in the first section while the second section is devoted to the background framework that is especially focused on the difference between a statistical classification of economic activities and that of occupations. All data and methods used in this study are presented in the third section. The main results are described in the fourth section while the conclusions and some future developments are provided in the last section.

2. Background

An individual entrepreneur (also known as sole proprietorship) is an enterprise owned exclusively by a single natural person, with no legal distinction between the owner and the business entity. The entrepreneur exercises his activity without having created a distinct legal person. Like all companies, every Italian individual entrepreneur is registered in the Italian SBR and is assigned a five-digit Ateco code that identifies the main economic activity.

2.1. An economic activity or an occupation?

From a statistical perspective, an economic activity is not an occupation. Economic activities are represented through statistical classifications of economic activities — e.g. the international ISIC, the European NACE and the Italian Ateco — while occupations are classified using statistical classifications of occupations — e.g. the international ISCO, the Italian classification of occupations (Istat, 2024).

An economic activity takes place when inputs to a production process — such as natural resources, equipment, labour, manufacturing techniques, information networks or intermediary products — are combined to produce specific goods or services. Thus, an economic activity is characterised by inputs of resources and production processes that generate outputs (goods or services). More specifically, the criteria used to define and delineate NACE/Ateco categories can be summarised as follows: the inputs of goods, services and production factors; the production process and technique; the characteristics of outputs; the intended use of outputs; the statistical purposes of the classification; and the availability of data.

By contrast, a job consists of the tasks and duties performed, or expected to be performed by one person. Occupation refers to the kind of work performed in a job and represents a set of jobs whose main tasks and duties share a high degree of similarity. Skill denotes the ability to perform the tasks and duties associated with a given job. Within ISCO classification and its Italian version, occupations are grouped according to two dimensions of skill: skill level and skill specialisation.

2.2. The new NACE Rev. 2.1 and Ateco 2025 classifications

According to the NACE introductory guidelines, units engaged in the same kind of economic activity are classified in the same NACE category. The NACE does not draw distinctions according to the kind of ownership of a production unit or its type of legal organisation or mode of operation, because such criteria do not relate to the characteristics of the activity itself.

Recently, a new classification rule concerning individual entrepreneurs has been introduced in the NACE Rev. 2.1 methodology (Eurostat, 2025; European Union, 2023). It states as follows.

The activities of individual entrepreneurs are classified according to the economic activity they carry out, in other words according to the goods or services they are producing, which is not necessarily identical with the economic activity of the unit they are working for. For example, the principal activity of an independent doctor working in a hospital must be classified in group 86.2 'Medical and dental practice activities', depending on the specialist area for which medical services are provided.

The above NACE guidelines are also valid for the Ateco classification representing the Italian version of the European NACE. Ateco 2025 (Alonzi and Viviano, 2025), directly derived from the NACE Rev. 2.1 classification, has replaced the previous version Ateco 2007 Update 2022 (also known as Ateco 2022) that instead had been directly derived from NACE Rev. 2 (Eurostat, 2008; European Union, 2006).

3. Research questions, data and methods

During the revision process that led to the release of the new Ateco 2025 classification, Istat collected several proposals aimed at improving the classification of individual entrepreneurs within the Ateco classification. For instance, national stakeholders called for a clearer recognition of economic activities run independently for example by herbalists, pharmacists, chefs, waiters, bus drivers, medical doctors, technical inspectors responsible for the periodic road-safety testing of vehicles, travel consultants, nurses, archivists, and others. This working arrangement has become increasingly prevalent across several sectors in recent decades; for instance, an increasing number of individuals prefer working on demand as independent pharmacists in various drugstores rather than being employed only by one of them.

Nevertheless, the terminology used within the Ateco classification is not always self-explanatory. Category labels often refer to the locations where activities are performed (e.g. 'activities of drugstores') or describe economic activities typically conducted by larger enterprises (e.g. influencer marketers may not identify with the definition of 'advertising agencies' which also encompasses agencies operated by individual entrepreneurs alongside larger firms).

Even if the contents of the new Ateco 2025 tried to solve the above gap, after the adoption of the new classification, users asked to Istat several clarification requests for the classification of individual entrepreneurs within the Ateco classification (e.g. brewers, mechanics, flight pilots, flight attendants, X-ray technologists, ironers). All the above examples represent the starting list of input data of our study that fostered the main research questions of this research work:

- Is the classification of activities run by individual entrepreneurs enough clear within the NACE/Ateco?
- If not, which is the impact on statistical data based on the NACE/Ateco?

To address the above questions, a descriptive and exploratory study drawing on user inputs and administrative evidence was conducted. This approach intended to describe complex issues in real-world contexts and to offer insights into contemporary, unfamiliar phenomena.

More specifically, the study examined a selection of classification issues concerning the activities of individual entrepreneurs, as raised by users. Each user inquiry was first analysed from a conceptual perspective, which entailed go through the structure and explanatory notes of the statistical classification of economic activities to assess their clarity and comprehensiveness. In parallel, both statistical and administrative sources were consulted. The statistical data included official information on the structure of enterprises by legal status and economic activity, while the administrative data, provided by the Chambers of Commerce, contained textual descriptions of the economic activity for each enterprise.

It should be noted that not all individual entrepreneurs are registered in the Italian Business Register maintained by the Chambers of Commerce; some are recorded only in fiscal registers. However, administrative sources provided by the Tax Agency do not include open-text descriptions of the economic activities carried out by enterprises. For this reason, although the Chambers of Commerce database may be considered a limitation of the study, it currently represents the most suitable source for achieving the objectives of this research.

4. Main results

The main results are presented using a user-centred perspective and a case-study approach. Given the specificities of each case study across different sections and industries, this paper focuses on a selected subset, specifically the activities of:

i. independent flight pilots and bus drivers (transportation industry);
ii. independent chefs and waiters (restaurant industry);
iii. independent medical doctors and x-ray technologists (health industry).

The presentation of each case study follows a consistent structure. Firstly, the original input request from national users is presented in order to contextualise the issue; secondly, the results of the conceptual and thematic NACE/Ateco analysis are provided; and thirdly, findings from the consultation of registry counts and textual entries are reported.

4.1. The activities of independent flight pilots and bus drivers

Original input request from national users concerning independent flight pilots:

Good morning, we would like your position about the correct classification for a freelance airline pilot, currently a pilot of airliners. He is currently classified in Ateco 2022 class 52.23 'Service activities incidental to air transportation'. Alternatively we would choose Ateco 2025 class 74.99 'All other professional, scientific and technical activities n.e.c.'.

Independent flight pilots are individuals working for airline companies or private jet charters companies without being paid employees of those companies. Thus, flight pilots are involved in the transport of passengers or freight by air and must be classified within NACE/Ateco division 51 'Air transport', specifically in group 51.1 if it concerns passenger air transport or in group 51.2 if it concerns freight air transport. On the contrary, the classes proposed by national users in the original input request encompassed activities like:

- operation of terminal facilities (e.g. airway terminals), airport and air traffic control activities, storage of aircraft, fire-fighting and fire-prevention services at airports;
- appraisal activities (for antiques, jewellery, etc.), bill auditing and freight rate information, weather forecasting activities, security consulting, agronomy and environmental consulting.

The above original input request confirms that there must exist individual entrepreneurs undertaking scheduled passenger air transport activities. In order words, there could be individual entrepreneurs classified in NACE/Ateco class 51.10. However, such an evidence is not supported by official statistical data where passenger air transport activities are broken down into two Ateco 2022 categories:

51.10.1 Scheduled air transport of passengers over regular routes
51.10.2 Charter flights for passengers and sightseeing flights

As shown in Table 1, official statistical data report no individual entrepreneurs classified under scheduled air transport of passengers over regular routes (Ateco 2022 category 51.10.1).

Table 1 – *Legal units classified in Ateco 2022 class 51.10 Passenger air transport by Ateco category and legal form (reference year 2023).*

Ateco 2022 category	No. of individual entrepreneurs	No. of other legal forms	Total
51.10.1	0	17	17
51.10.2	8	71	79
Total	**8**	**88**	**96**

These findings can be read in two ways: either the only existing freelance airline pilot is the one who contacted Istat for clarification, or freelance airline pilots are classified under a different Ateco 2022 code. To support the latter assumption, a text search was conducted among the activity descriptions of companies registered with the Chambers of Commerce. The results suggest that individual entrepreneurs involved in air transport are more commonly engaged in pilot training courses or in aerial filming and photography services using aircraft.

A similar issue arisen within the NACE/Ateco transportation industry is related to freelance bus drivers.

Original input request from national users concerning independent bus drivers:

Good morning, I am a bus driver working on a fee or contract bases. Despite having professional certifications as well as a vat number, my activity is not mentioned in the classification. Which NACE class would you suggest?

The NACE/Ateco group 49.3 ‘Other passenger land transport’ includes all land-based passenger transport activities other than rail transport; this group also includes land transport by motor bus. Freight transport by road is instead classified in NACE/Ateco group 49.4. Thus, the above groups are not reserved only to bus companies owning the buses and employing bus drivers but also to independent bus drivers.

4.2. The activities of independent chefs and waiters

Original input request from national users concerning independent chefs and waiters:

Good morning, are independent chefs and waiters working for a restaurant classified in Ateco 2025 class 96.99 ‘Other personal service activities n.e.c.’?

According to the new rule on individual entrepreneurs introduced in the methodology of the NACE Rev. 2.1 classification, both the activities of chefs and waiters working for one or even more restaurants should be classified according to the goods or services they are producing, and therefore fall within the new NACE/Ateco class 56.11 ‘Restaurant activities’. Importantly, even when chefs and waiters provide their services independently to households, their activities remain part of class 56.11 and can not be regarded as personal service activities (NACE/Ateco division 96). Such a positioning is further mentioned in NACE/Ateco new class 96.91, which covers the provision of a range of domestic personal service activities, such as cooking, washing, cleaning and ironing services, at home for households (these are combined activities); however, when only one activity is performed it is classified according to that specific activity.

In the field of restaurant activities, individual enterprises managing a restaurant account for 27% of all enterprises classified in NACE/Ateco division 56 'Food and beverage service activities' and for 45.4% of those in NACE/Ateco group 56.1 'Restaurants and mobile food service activities' according to the Ateco 2022 code.

This finding confirms the widespread presence of restaurants managed by individual entrepreneurs. By contrast, the phenomenon of freelance chefs working outside their own restaurants or households is less readily measurable, as some are classified under 'Event catering and other food service activities' (group 56.2) or under 'Other personal service activities n.e.c.' (former NACE/Ateco 96.09). In fact looking at the textual description of the economic activities registered in the Italian Business Register managed by the Chambers of Commerce there is a certain number of chefs who declared that they work independently at the customers' location (in Italian *a domicilio*). However, no other information about the type of customers (households or other enterprises, e.g. restaurants) is available.

On the one hand, if the service is provided for households, it had to be classified in Ateco division 96 according to a proper explanatory note included in the national classification Ateco 2022. According to the new Ateco 2025 classification, it should instead be recorded into NACE/Ateco division 56 in line with the new classification rule concerning individual entrepreneurs. On the other hand, if the service is provided for other enterprises, e.g. restaurants, it had to be classified in NACE/Ateco group 56.1; according to the new classification, it should be still classified in NACE/Ateco group 56.1.

To sum up, such a phenomenon (i.e. independent chefs working for one or more restaurants) is not measurable even if the original input request confirms that it exists.

4.3. The activities of independent medical doctors and X-ray technologists

Original input request from national users concerning independent medical doctors:

Good morning, what kind of units are those classified in Ateco 2022 class 86.10 as individual entrepreneurs?

The above request concerns hospital activities. The former NACE/Ateco class 86.10 in fact included short- or long-term hospital activities, i.e. medical, diagnostic and treatment activities. It covered general hospitals (e.g. community and regional hospitals, hospitals of non-profit organisations, university hospitals, military-base and prison hospitals) as well as specialised hospitals (e.g. mental health and substance abuse hospitals, hospitals for infectious diseases, maternity hospitals, specialised sanatoriums).

At national level, the former NACE class 86.10 ‘Hospital activities’ was broken down into four Ateco 2022 categories:

86.10.1 Activities of general hospitals (excluding university hospitals)
86.10.2 Activities of specialised hospitals
86.10.3 Activities of university hospitals
86.10.4 Provision of long-term hospital activities

Table 2 – *Legal units classified in Ateco 2022 class 86.10 Hospital activities by Ateco category and legal form (reference year 2023).*

Ateco 2022 category	No. of individual entrepreneurs	No. of other legal forms	Total
86.10.1	383	92	475
86.10.2	243	41	284
86.10.3	0	0	0
86.10.4	106	15	121
Total	**732**	**148**	**880**

From the point of view of the users providing the original input request, the former NACE/Ateco class 86.10 was supposed to classify only activities performed by hospitals legally recognised to treat patients. However, this criterion was less clearly defined in the previous classification, and a number of independent medical doctors working primarily for hospitals were also recorded under this category.

Fortunately, the new classification rule concerning individual entrepreneurs made clear that individual entrepreneurs are classified according to the economic activity they carry out, in other words according to the goods or services they are producing, which is not necessarily identical with the economic activity of the unit they are working for. Thus, the principal activity of an independent doctor working in a hospital must be classified in group 86.2 ‘Medical and dental practice activities’, depending on the specialist area for which medical services are provided, and not within hospital activities.

A similar issue arisen within the NACE/Ateco health industry concerns activities of independent x-ray technologists.

Original input request from national users concerning independent x-ray technologists:

Good morning, we need your support about the classification of independent x-ray technologists represented by our association.

Activities of independent x-ray technologists are classified in Ateco 2022 category 86.90.1 ‘Diagnostic imaging services and medical laboratory activities’ corresponding to Ateco 2025 category 86.91.0 where are included activities of medical laboratories providing analytic or diagnostic services, body fluid analysis or genetic testing, directly to outpatients with or without referral from healthcare practitioners. Specifically, activities of x-ray laboratories and other diagnostic imaging centres are classified here. Despite the wording used to describe the contents of this code (‘laboratories’, ‘centres’) also activities of independent x-ray technologists working for the same laboratories or centres, hospitals and other human health institutions are represented by the same Ateco code.

Table 3 – *Legal units classified in Ateco 2022 category 86.90.1 by legal form (reference year 2023).*

Legal form	No. of legal units
Individual entrepreneurs	4.989
Other	4.899
Total	**9.888**

Unfortunately, official statistical data (Table 3) are not meaningful in this case because among legal units operating as individual entrepreneurs there are both people running their own diagnostic laboratory/centre and people working as freelancers at diagnostic laboratories/centres of someone’s else.

Nevertheless, the fact that the input request was provided by the representing association is instead very significant; a further discussion with the association itself has shown that currently freelance x-ray technologists are not classified in a consistent way due to the wording used within the classification to describe the activities.

5. Conclusions and future developments

In Italy, the NACE/Ateco classification is applied to all types of enterprises, including individual entrepreneurs, who account for 67% of the total. From a statistical perspective, economic activity defined through the NACE/Ateco system constitutes the primary criterion for business and economic statistics. From a non-statistical perspective, the NACE/Ateco code is also employed for tax purposes and for the allocation of incentives and contributions to economic activities.

The quality of official statistics based on the NACE/Ateco classification depends on the accuracy of the code assigned to each enterprise, including individual entrepreneurs. Likewise, administrative decisions that rely on NACE/Ateco codes can be considered reliable only if enterprises are classified in a consistent and coherent manner.

However, this study confirms that the application of statistical classifications, such as those of economic activities, is strongly influenced by coding errors, which largely depend on the knowledge and expertise of the users applying them. The descriptive and exploratory analysis shows that such errors are particularly frequent among individual entrepreneurs, with direct implications for statistical production, as their records are typically not subject to manual verification by expert statisticians.

Nevertheless, the authors acknowledge that this research work represents only a first and preliminary step in addressing an important and persistent issue that requires further investigation. In particular, the methodology needs to be improved by identifying additional sources capable of providing textual descriptions of economic activities carried out by individual entrepreneurs. In this regard, the existence of suitable fiscal sources could be better explored. Moreover, consideration could be given to designing an ad hoc survey targeting individual entrepreneurs classified within the main catch-all NACE/Ateco classes (e.g. new classes 74.99, 82.99 or 96.99).

At the moment of writing, all the issues arisen by users are being discussed with the main administrative bodies, especially the Chambers of Commerce, with the aim of strengthening cooperation with institutions that make use of the same NACE/Ateco classification. In parallel, most classification issues derived from the case-based analysis are being dealt as index entries, thus as extensive explanatory notes of the new Ateco 2025 classification, that will be soon disseminated to users (enterprises including individual entrepreneurs themselves, administrative bodies, accountants and labour consultants, trade associations). The short-term objective of this strategy is to promote the correct interpretation and, consequently, the proper application of the classification. The ultimate and most challenging objective is to achieve a more consistent use of the NACE/Ateco codes at the administrative level, thereby enhancing the quality of official statistics based on administrative data.

References

ALONZI F., VIVIANO C. 2025. *La struttura (codici e titoli) della classificazione delle attività economiche ATECO 2025*.
https://www.istat.it/wp-content/uploads/2024/12/Nota-metodologica-1.pdf

EUROPEAN UNION. 2023. *Commission Delegated Regulation (EU) 2023/137 of 10 October 2022 amending Regulation (EC) No 1893/2006 of the European*

Parliament and of the Council establishing the statistical classification of economic activities NACE Revision 2. Official Journal of the European Union. http://data.europa.eu/eli/reg_del/2023/137/oj

EUROPEAN UNION. 2006. *Regulation (EC) No 1893/2006 of the European Parliament and of the Council of 20 December 2006 establishing the statistical classification of economic activities NACE Revision 2 and amending Council Regulation (EEC) No 3037/90 as well as certain EC Regulations on specific statistical domains Text with EEA relevance*. Official Journal of the European Union. http://data.europa.eu/eli/reg/2006/1893/oj

EUROSTAT. 2025. *NACE Rev. 2.1 Statistical classification of economic activities in the European Union.* 2025 edition.
https://ec.europa.eu/eurostat/documents/3859598/21633320/KS-GQ-24-007-EN-N.pdf#page=33

EUROSTAT. 2008. *NACE Rev. 2 Statistical classification of economic activities in the European Community*. Methodologies and working papers
dd5443f5-b886-40e4-920d-9df03590ff91

ISTAT. 2024. *Classificazione delle professioni. Revisione 2021.*
https://www.istat.it/wp-content/uploads/2024/01/FOCUS_PROFESSIONI_2021.pdf
https://www.istat.it/classificazione/classificazione-delle-professioni/

Francesca ALONZI, Istat - Italian National Institute of Statistics, francesca.alonzi@istat.it
Caterina VIVIANO, Istat - Italian National Institute of Statistics, caterina.viviano@istat.it

THE PERFORMANCE OF LOCAL ECONOMIES BEFORE AND AFTER THE COVID-19: AN ANALYSIS AT THE MUNICIPAL LEVEL[1]

Elisabetta Bilotta, Ilaria Straccamore

Abstract. The economic crisis triggered by the Covid-19 pandemic has had a profound impact on companies' performance and, consequently, on local economies. Against this backdrop, this paper aims to offer a comparative analysis of local economic performance before the pandemic (2015–2019), during its first year (2020), and during the economic recovery phase (2022). The analysis highlights the impact of the crisis on the structure and dynamics of the Italian production system at a territorial level. The analysis focuses on the manufacturing and personal service sectors and uses data from the "Territorial Frame SBS" statistical register produced annually by ISTAT. Firstly, the analysis involves constructing profiles at the municipal level. For each period, it takes into account the combination of measures that capture employment trends and value added dynamics with different degrees of intensity. Secondly, the analysis focuses on business productivity, applying statistical spatial analysis methods to identify spatial correlations at the municipal level post-pandemic (2022). The method used is LISA (Local Indicator Spatial Analysis), developed by L. Anselin.

1. Introduction

This paper presents a comparative analysis of the performance of Italian local economies before the pandemic (2015–2019), during the first year of its spread (2019–2020), and during the subsequent economic recovery phase (2020–2022). The aim is to evaluate the impact of the pandemic on the structure and dynamics of Italy's production system at a local level.

Apparent labour productivity is used as a key indicator of economic performance, reflecting both classical and contemporary theories on growth and territorial development. As highlighted by Porter (1990) and Becattini (1975; 1979; 1987), local environments shape firms' productivity and competitiveness, making territory an active economic resource.

---

[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 and 4 to Elisabetta Bilotta and paragraphs 2, 3 and 5 to Ilaria Straccamore.

The Covid-19 pandemic was a major shock to local economies, evolving rapidly from a health emergency into a global economic crisis. In Italy, its impact on the production system was examined by institutions such as ISTAT and the Bank of Italy, alongside academic contributions like Caddemi (2023), who highlighted firms' resilience strategies and the role of organisational flexibility. Most studies focused on regional or provincial levels, revealing uneven effects: the North was initially hit hardest, while the South experienced more persistent challenges due to structural vulnerabilities (ISTAT, 2022b; Bank of Italy, 2022). Reports by Confindustria and Cerved further underscored territorial disparities in both impact and recovery capacity.

In contrast, comprehensive assessments at the municipal level remain limited. It is precisely within this gap that the present contribution is situated. Building on the initial effort presented in the Istat 2023 Annual Report—which offered a first reading of the evolution of local economies before and during the pandemic—this study extends that framework, both by incorporating post-Covid dynamics and by deepening the analytical scope.

The contribution is organised into three sections. Paragraph 2 describes the official data and methods used to compile the database. Paragraph 3 provides detailed territorial information at a municipal level on the spatial dynamics that have characterised local economies in recent years in the manufacturing and personal services activities. Paragraph 4 examines business productivity through the application of spatial statistical analysis, with the aim of identifying clusters of municipalities that demonstrated elevated productivity levels in the aftermath of the Covid-19 pandemic. Paragraph 5 presents the concluding remarks, summarising the key findings of the analysis.

2. Official data source and method

This study considers data from the Territorial Structural Business Statistics Frame (T-SBS Register), which is produced annually by ISTAT and provides information on the main economic account variables of local enterprise units operating in the industrial and non-financial service sector within the national territory.

For our purposes, we considered T-SBS Register data for 2015, 2019, 2020 and 2022. These different sources are integrated at municipal level, so the municipality is the unit of analysis. The final database is a panel consisting of municipalities that existed in all the considered years (2015, 2019, 2020 and 2022). Integrating the different sources involved complex temporal harmonisation of the input databases, which refer to different time periods. Therefore, it was necessary to consider all the changes that occurred at the territorial level during the time period in question, and to trace all the information to the territorial configuration as of 12/31/2022. In particular, the analysis focuses on the manufacturing[2] and personal service sectors[3], which were the most affected by the pandemic. The final database consists in 7,347 municipalities with manufacturing activities on their territory and 7,799 municipalities with personal services activities.

3. Municipal profiling before, during and after the pandemic

3.1. The methodology used

Three time phases have been considered: 2015–2019, highlighting long-term trends and pre-pandemic stability at territorial level; 2019–2020, defining the immediate impact of the pandemic; and 2020–2022, defining the recovery phase. This considers the impact of the crisis and the ability to withstand endogenous shocks. The profiling of the municipalities was carried out on the basis of a combination of measures capturing employment trends and value-added dynamics at different intensity levels, for each period. This combination of measures results in the municipal profile being divided into six categories. With regard to employment trends, two categories were created: growth (positive variation above the national average) and stability or decrease (variation equal to or below the national average). Regarding value added, three categories were created based on the analysed period. For example, the categories emphasise growth in the periods 2015–2019 and 2020–2022, as well as the economic crisis in 2019–2020. In the first case, we considered high growth (positive variation above the national average), stability or low growth (variation in line with the national average or positive variation below the national average), and decrease (negative variation). In the second case, we considered growth (positive variation), stability or low decrease (variation in line with the national average or negative variation within the national average), and high decrease (negative variation below the national average) (Table 1).

[2] Classification of Economic Activity 2007 (Ateco), section C
[3] Classification of Economic Activity 2007 (Ateco), divisions: 55 - accommodation, 56 - restaurants, 49 - land transport, 50 - sea transport, 51 - air transport, 79 - travel agency and tour operator services, 90 - creative, artistic and entertainment activities, 91 - libraries, archives, museums, 92 - lottery, betting and gambling activities, 93 - sports, recreation and amusement activities.

Table 1 – *Municipal profiling for three periods considered.*

Legend	2015-2019 and 2020-2022	2019-2020
1	High growth VA* and growth PE**	Growth VA and PE
2	Stable or low growth VA and growth PE	Stable or low decrease VA and growth PE
3	Decrease VA and growth PE	High decrease VA and growth PE
4	High growth VA and PE stable or decreasing	Growth VA and PE stable or decreasing
5	Stability or low growth VA and PE stable or decreasing	Stability or slight decrease VA and PE stable or decreasing
6	Decrease VA and PE stable or decreasing	High decrease VA and PE stable or decreasing

*Value added **Persons employed

3.2. Municipal profiles in the manufacturing

In the manufacturing, the pandemic had a widespread effect that went beyond the growth patterns observed during the previous period (2015–2019). During these years, Italian municipalities exhibited a high degree of heterogeneity in their profiles with regard to value added and employment dynamics (Figure 1). In 2015-2019, over 2,600 municipalities (35%) experienced growth in industrial value added exceeding the national average, accompanied by positive employment dynamics. These municipalities were primarily located in Northern and Central Italy, as well as in certain areas of Southern Italy (mode 1 – intensive green colour). Conversely, over 2,000 municipalities (28% of all manufacturing municipalities), primarily in Southern Italy but not exclusively, experienced a decline in value added alongside negative employment dynamics (mode 6 – intense shade of red). A further 800 municipalities (11%), located in a widespread but specific manner across the national territory, experienced strong growth in value added alongside negative employment trends (mode 4 – intensive blue), suggesting ongoing restructuring processes in local production systems. Finally, almost 500 municipalities (7%), located in specific areas across the country, experienced employment growth alongside a reduction in industrial value added (mode 3 – light blue), suggesting potential limitations to future growth prospects.

The impact of the pandemic on local economies has greatly increased the number of municipalities with negative performance (Figure 2). In 2019-2020 period, almost 3,000 municipalities (40%) experienced a sharp decline in value added and negative employment dynamics (mode 6 – intense shade of red), while only 1,200 municipalities (17%) continued to record the best positive performance (mode 1 – intensive green colour). Additionally, over 900 municipalities (12%) non-positive performance in terms of both value added and employment (mode 5 – light shade of red), primarily located in Northern and Central Italy.

The post-pandemic period (2020-2022) is characterised by economic recovery (Figure 3). Compared to 2015-2019, the number of municipalities in the top two performance categories (modes 1 – intensive green colour and 2 – light green colour) increased from 3,300 before the pandemic to 3,800. Moreover, 1,100 fewer municipalities (900 versus 2,000) recorded a decline in value added associated with negative employment trends (mode 6 – intense shade of red). However, strong growth in both indicators (mode 1 – intensive green colour) declined in 300 municipalities (2,600 versus 2,300), mostly in the North-East, suggesting that certain local economies have not yet returned to pre-Covid levels.

Figure 1 - *Municipal profiles in manufacturing. Years 2015-2019.*

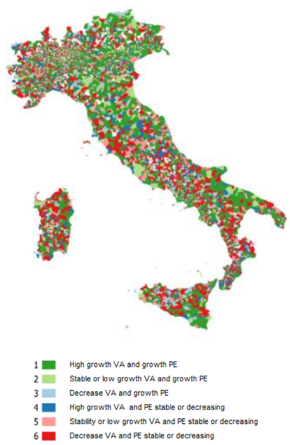
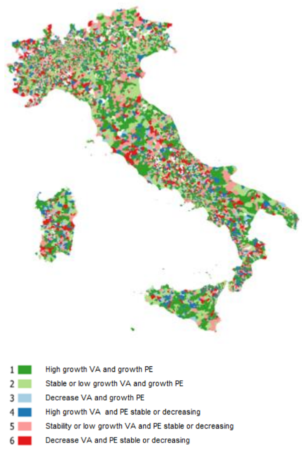


Figure 2 - *Municipal profiles in manufacturing. Years 2019-2020.*

Figure 3 - *Municipal profiles in manufacturing. Years 2020-2022.*



Source: our elaborations on Istat T-SBS Register

3.3. Municipal profiles in the personal services activities

Around 60% of municipalities in the personal service sector, which are mainly located in the North-East and Centre of Italy, showed positive performance dynamics between 2015 and 2019 (Figure 4). Specifically, 3,000 municipalities (40%) recorded growth in value added above the national average and positive employment dynamics (mode 1 – intensive green colour), while 1,500 municipalities (20%) experienced stable or low growth in value added and employment (mode 2 – light green colour). Conversely, 1,500 municipalities (20%) recorded a decline in value added and a negative employment trend (mode 6 – intense shade of red), which were mainly located in the North-West and South.

During the 2019–2020 period, 71% of municipalities experienced a decline or stability in both value added and employment (modes 5 – light shade of red and 6 – intense shade of red in Figure 5). Of these, in 45% of cases, the pandemic acted as an exogenous shock, marking a transition from stability or economic growth to decline. These areas are strongly specialised in tourism-related sectors and are spread throughout the country, with the highest concentrations in the North-East and Centre. In 15% of cases, the pandemic exacerbated existing crises, especially in North-West Italy, primarily in the traditionally industrial areas of Piedmont and Lombardy, as well as the inland areas of Liguria, but also in the Centre and South. However, 9% of municipalities that experienced growth in the pre-crisis period continued to grow, remained stable, or saw a slight decline in value added during the first year of the pandemic. Southern Italy has the highest concentration of municipalities of this type, which are not usually included in traditional tourist routes and have therefore been less affected by the severe mobility restrictions aimed at combatting the spread of the pandemic.

The post-pandemic period is characterised by economic recovery (Figure 6). Compared to 2015–2019, more municipalities now are into the top two performance categories (mode 1 – intensive green colour and 2 – light green colour) than before the pandemic (4,800 versus 4,500), and 900 fewer municipalities (600 versus 1,500) record a drop in value added accompanied by negative employment dynamics (mode 6 – intense shade of red). However, when considering municipalities recording strong growth in value added accompanied by positive employment trends (mode 1 – intensive green colour), the personal service sector is not returning to pre-Covid levels, with 300 fewer municipalities (2,700 versus 3,000), primarily located in the North-East.

Figure 4 - *Municipal profiles in personal service activities. Years 2015-2019.*

Figure 5 - *Municipal profiles in personal service activities. Years 2019-2020.*

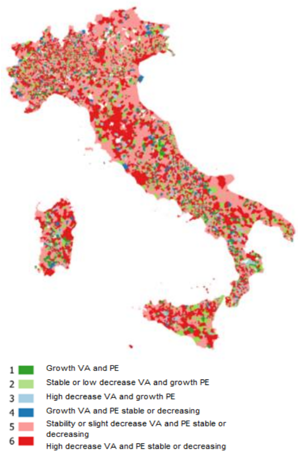


Figure 6 - *Municipal profiles in personal service activities. Years 2020-2022.*

Source: our elaborations on Istat T-SBS Register

4. Local spatial correlation analysis on apparent labour productivity levels

4.1. Analysis method

This analysis adopts apparent labour productivity—measured as the ratio of value added to the number of employees—as a central indicator of economic performance. The application of statistical methods of spatial analysis (LISA, Local Indicator Spatial Analysis) to the levels of apparent labour productivity at municipal level makes it possible to identify geographically contiguous clusters of municipalities that are characterised by specific development patterns.

Moran's I index is one of the most common measures of spatial association and is formalised as follows:

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

where n is the number of territorial units, y is the mean value of the variable, y_i is the value of the variable at location i, while w_{ij} is a measure of the relationships between locations i and j. As is well known, the values assumed by Moran's index are sensitive to different measures of contiguity or distance between units. Moran's I index can be interpreted as a measure of the correlation of variable y and its "spatial lag", defined by the average value of all values assumed by y in neighbouring areas. This index, considering all territorial units, measures spatial correlation at a "global" level, identifying spatial patterns at a regional or macro-regional level. It is therefore not particularly suitable for identifying agglomeration processes at the local level.

Local Indicator Spatial Analysis (LISA), developed by Anselin (1995) and included as a test algorithm in the GeoDA software, provide accurate information on agglomeration processes with respect to each territorial unit. From an analytical point of view, local spatial correlation indices (LISA) are a decomposition of Moran's "global" index.

Given a statistical significance level (p value = 0.05) it is possible to identify four types of territorial unit grouping characterised by the following levels of local spatial correlation between municipalities located at the centre and at the periphery of the grouping:

- "high-high" identifies the areas that present a statistically significant association (p value = 0.05 and I_i>0) with high values of productivity within a municipality and in its contiguous areas;
- "low-low" identifies the areas that present a statistically significant association (p value = 0.05 and I_i>0) with reduced values of productivity within a municipality and in its contiguous areas;
- "low-high" identifies areas that present a statistically significant association (p value = 0.05 and I_i<0) with reduced productivity values within a municipality and with high values in its contiguous areas;
- "high-low" identifies areas that present a statistically significant association (p value = 0.05 and I_i<0) with high productivity values within a municipality and with reduced productivity in its contiguous areas.[4]

The first two types identify so-called hot-spots, distinguishing established clusters (high-high) from emerging ones (low-low), while the other two types identify so-called "spatial outlayers".

4.2. Results

In 2022, the manufacturing sector included almost 700 municipalities characterised by high and widespread productivity levels, and thus higher growth

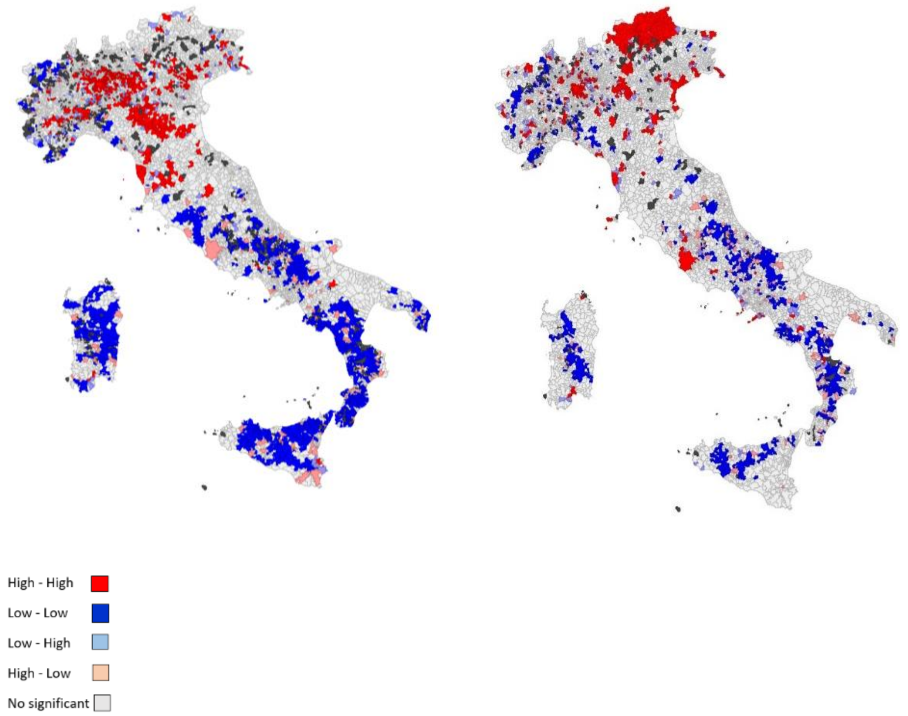
[4] The classification of a value as 'high' or 'low' is determined relative to the mean: if productivity exceeds the average, it is categorized as high; conversely, if it falls below the average, it is considered low.

potential (high-high). These municipalities were located in the most industrialised areas of Piemonte, Lombardia, Veneto and Emilia-Romagna. Some clusters are also found in specific areas of central Italy (Figure 7). The most widespread clusters are in Tuscany, whereas in the other central regions they are much more limited, covering parts of the Perugia municipality and the Frosinone province. Clusters of municipalities with high levels of productivity are present to a very limited extent in Southern Italy. On the other hand, the clusters of municipalities that are characterised internally by low and diffuse productivity levels (low-low), and thus lower growth potential, are 1,073 and are predominantly located in the South and Islands. The clusters of municipalities that are characterised internally by a high level of productivity in the centre but also by lower levels of productivity on the periphery (high-low) are 120 and are mainly located in the Centre, South and Islands. These clusters have a high growth potential but limited spill-over effects on the surrounding area. Clusters characterised by low productivity in the centre and high productivity in the periphery (low-higt) are 198 and are also mainly located in the Centre, the South and Islands.

In 2022, the clusters of municipalities within the personal services activities with high levels of productivity (high-high) are 511 (Figure 7). A large area of agglomeration characterised by high levels of productivity is found in Trentino Alto Adige and other specific areas in northern and central Italy. In the South and Islands, areas with high levels of productivity are found in specific areas such as Capri in Campania and Vasto in Abruzzo. Emerging clusters, which are characterised by low and diffuse productivity levels (low-low) are 667 and thus lower growth potential, are mainly located in central and southern Italy. The clusters of municipalities that are characterised internally by a high level of productivity in the centre but also by lower levels of productivity on the periphery (high-low) are 144 and are mainly located in the Centre, South and Islands. These clusters have a high growth potential but limited spill-over effects on the surrounding area. Clusters characterised by low productivity in the centre and high productivity in the periphery (low-high) are 194 and are also mainly located in the Centre, South and Islands.

However, it is important to remember that spatial analysis of data using the LISA method is purely exploratory. This is because local agglomeration patterns detected at one or more dimensions can be caused by spurious spatial correlations with other variables. For example, the presence of urban areas or transport infrastructures can influence business location.

Figure 7 – *Local spatial correlation indicators (LISA) for productivity levels at municipal level in the manufacturing sector (left) and in personal service activities (right). Year 2022*[1].



Source: our elaborations on Istat T-SBS Register.

5. Conclusions

The pandemic was an unprecedented external shock that had a profound impact on local economies, with the personal services sector being one of the worst affected. Indeed, in 2020, over 70% of municipalities experienced a decline or stagnation in value added alongside negative employment trends in this sector. However, in the post-pandemic period, signs of economic recovery emerged, particularly in the personal services sector, which was driven by the tourism rebound. Nevertheless, the performance of the best-performing municipalities is still lower than in the pre-pandemic period, indicating that certain local economies has not yet fully returned to its 2015–2019 level.

The analysis also revealed significant differences in the impact of the pandemic on the local economies and their subsequent recovery paths. These differences were influenced by the severity of the economic downturn in 2020 and the structural characteristics at the local level. In around 1,000 municipalities, mainly in the North-East and across both manufacturing and services, economies that had previously shown stability or growth transitioned to stagnation or decline. Meanwhile, other areas experienced persistent crisis conditions, with around 900 manufacturing sector municipalities – primarily in Lombardy and Piedmont, in areas known for their 'Made in Italy' and 'Manufacturing' Local Systems – and approximately 500 personal services sector municipalities, particularly in the North-West and Mezzogiorno. In contrast, certain areas exhibited a clear post-pandemic resurgence. Around 1,300 manufacturing municipalities, primarily situated in the South – particularly in Sardinia's Non-Manufacturing Local Systems and Sicily's Non-Specialised Local Systems – displayed positive recovery trends. Similarly, around 1,130 service-focused municipalities, mainly in the North-West, also recorded encouraging post-pandemic developments.

Finally, spatial correlation analysis revealed clusters of municipalities characterised by high productivity levels. In the manufacturing sector, these clusters are found in the most industrialised areas of Piedmont, Lombardy, Veneto and Emilia-Romagna. Clusters of municipalities with high productivity levels in the personal services sector are located in a large agglomeration area in Trentino-Alto Adige and in other specific areas in northern and central Italy.

In conclusion, the Italian entrepreneurial system, already heavily influenced by pre-existing structural factors, has been severely tested by the pandemic, exacerbating existing structural and economic differences between Italian regions.

References

ANSELIN L. 1995. Local indicators of spatial association — LISA, *Geographical Analysis*, No. 27, pp. 93–115.

ANSELIN L., SYABRI I., SMIRNOV O., REN Y. 2002. Visualizing spatial autocorrelation with dynamically linked windows, *Computing Science and Statistics*, Vol. 33.

BANCA D'ITALIA. 2022. L'impatto del Covid-19 sui fallimenti e le uscite dal mercato delle imprese italiane, https://www.bancaditalia.it/media/notizia/l-impatto-del-covid-19-sui-fallimenti-e-le-uscite-dal-mercato-delle-imprese-italiane/?dotcache=refresh.

BECATTINI G. 1987. *Mercato e forze locali: il distretto industriale*. Bologna: Il Mulino.

BECATTINI G. 1979. Dal settore industriale al distretto industriale. Alcune considerazioni sull'unità di indagine dell'economia industriale, *Rivista di Economia e Politica Industriale*, Vol.1, No. 1, pp. 7–21.

BECATTINI G. 1975. *Lo sviluppo economico della Toscana con particolare riguardo all'industrializzazione leggera*. Firenze: IRPET–Guaraldi.

CADDEMI D. 2023. Imprese e pandemia: Analisi della sopravvivenza delle imprese europee alla pandemia Covid-19, *Journal of Applied Economics*, Vol. XLII, No.3, pp. 1-20.

CONFINDUSTRIA, CERVED. 2022. *Rapporto regionale PMI 2022: a rischio la ripresa economica nel biennio 2022–23,* Roma, Settembre 2022.

LUCAS R. E. Jr. 1988. On the Mechanics of Economic Development, *Journal of Monetary Economics*, Vol. 22, No. 1, pp. 3–42.

ISTAT. 2024. Risultati economici delle imprese e delle multinazionali a livello territoriale - Anno 2022. *Statistica Report*, ISTAT. https://www.istat.it/tavole-di-dati/risultati-economici-delle-imprese-e-delle-multinazionali-a-livello-territoriale-anno-2022/

ISTAT. 2023. *Rapporto annuale 2023 - La situazione del Paese.* Roma: Istat.

ISTAT. 2022a. *Rapporto annuale 2022 - La situazione del Paese.* Roma: Istat.

ISTAT. 2022b. *Rapporto sulla competitività dei settori produttivi. Edizione 2021*. Roma: Istat.

ISTAT. 2022c. Situazione e prospettive delle imprese nell'emergenza sanitaria Covid-19. *Statistica Report*, ISTAT. https://www.istat.it/comunicato-stampa/situazione-e-prospettive-delle-imprese-nellemergenza-sanitaria-covid-19-giugno-ottobre-2020/

PORTER M. E. 1990. *The Competitive Advantage of Nations*. New York: Free Press.

RICARDO D. 1817. *On the Principles of Political Economy and Taxation*. London: John Murray.

ROMER P. M. 1986. Increasing Returns and Long-Run Growth, *Journal of Political Economy*, Vol. 94, No. 5, pp. 1002–1037.

SMITH A. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations.* London: W. Strahan and T. Cadell.

SOLOW R. M. 1956. A Contribution to the Theory of Economic Growth, *The Quarterly Journal of Economics*, Vol. 70, No. 1, pp. 65–94.

Elisabetta BILOTTA, Istat, bilotta@istat.it
Ilaria STRACCAMORE, Istat, straccamore@istat.it

SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS
ISSN ASSEGNATO: 0035-6832
E-ISSN ASSEGNATO: 3103-1757

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).

LXI Povertà e disuguaglianze: popolazioni e territori nell'era delle transizioni (Istituto Nazionale di Statistica e Università degli Studi di Roma Tre, Roma, 28-30 maggio 2025).