VOLUME LXXIX – N. 1 GENNAIO – MARZO 2025

RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA

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

SIEDS
SOCIETÀ ITALIANA
DI ECONOMIA DEMOGRAFIA E STATISTICA

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

info@sieds.it

rivista@sieds.it

VOLUME FUORI COMMERCIO

INDICE

MEASURING SOCIO-ECONOMIC DEPRIVATION AT SUB-MUNICIPAL LEVEL THROUGH THE INTEGRATION OF CENSUS AND ADMINISTRATIVE DATA

Giancarlo Carbonetti, Flavio Biasciucci, Andrea Cutillo, Matteo Mazziotta, Valeria Quondamstefano, Maria Tiziana Tamburrano, Debora Tronu

Abstract. The integration of permanent population and housing census data with administrative sources offers an exceptional opportunity to produce detailed indicators of demographic, social, and economic phenomena down to the sub-municipal level. This approach, which combines accurate, timely, renewable, and geocoded data, enables in-depth analysis of key phenomena and allows for tracking their evolution across time and space.
The Italian National Institute of Statistics (Istat) has recently launched a project to measure socio-economic deprivation at the sub-municipal level. Leveraging existing databases, a series of individual indicators were developed to assess deprivation across four dimensions: economic status, employment, education, and housing. These indicators are synthesized into the Socio-Economic Deprivation Index of population (SED-Index), calculated at the enumeration area level. The availability of data at this level provides significant potential for analysis and supports policymakers in identifying areas with higher socio-economic deprivation within municipalities.

1. Introduction

The increasing demand for highly detailed geographic data, along with the growing availability of administrative data, has driven the Italian National Institute of Statistics (Istat) to modernize its processes for producing official statistics. Istat's approach is rooted in two significant advancements in official statistics in recent years: the new Permanent Census of Population and Housing, which enables annual collection of census data and the dissemination of accurate and timely information down to the municipal level; and the Integrated Register System, which allows for the systematic use of administrative sources to produce official statistics.

The combined use of these data sources will enable the production of territorial data with enhanced detail, updated annually, down to the enumeration area (EA). This provides new opportunities for spatial analysis of demographic, social, and economic phenomena and offers a powerful tool for local policymakers.

This article proposes an enumeration area EA-level index of household deprivation to study socio-economic inequalities within several large Italian municipalities. Several examples of such indicators already exist in the literature, e.g. the Severe

Material and Social Deprivation rate (SMSD), that is a Eurostat indicator that shows an enforced lack of necessary and desirable items to lead an adequate life. As it is based on a sample survey (Eu-Silc), it cannot achieve a high level of territorial disaggregation. An indicator that is based on census data, along with the Eu-Silc survey, is the European Deprivation Index (EDI) (Guillaume *et al.*, 2016). In the Italian context, deprivation indices have been developed at the municipal level based on the 1991 Population Census (Cadum *et al.*, 1999) and at the EA level based on the 2001 (Caranci *et al.*, 2010) and 2011 censuses (Rosano *et al.*, 2020).

However, these indices are no longer replicable due to changes in census methodology. To address this gap, Istat has launched a project to analyze household socio-economic deprivation at the sub-municipal level. This project introduces a new deprivation index based on a broader and more representative set of individual indicators, capturing various dimensions of deprivation: economic status, employment, education, and housing.

The paper is organized as follows: section 2 ad 3 present the new strategy of the Permanent Population and Housing Census (PPHC) and the sub-municipal data process through the integration of census and administrative data; section 4 illustrates the definition of household deprivation adopted, the individual indicators selected to measure socio-economic components of deprivation and the methodology used for the construction of the composite index; section 5 presents the results of an initial spatial analysis, relating to the municipality of Palermo, to identify the potential critical areas of concentration of household deprivation.

2. The sub-municipal data production process

In 2018, Istat introduced the Permanent Population and Housing Census (PPHC) to replace the traditional Census conducted every ten years. This new approach employs a combined methodology that integrates administrative data with sample surveys, ensuring annual data collection and dissemination (Falorsi, 2017; Gallo and Zindato, 2018). The development of this innovative census strategy was made possible by the extensive availability of administrative data and the opportunity to use Istat's statistical registers built from these sources.

At the core of the PPHC is the Basic Register of Individuals (BRI), which annually identifies usual residents in Italy as determined by the Census. Together with the Basic Register of Places (BRP) and thematic registers on education, employment and income, the BRI is the basis for census data production within a combined framework: indeed, this design includes two dedicated sample surveys (the Area Survey and the List Survey) conducted each year. The PPHC generates both a fully register-based

population count and census hypercubes, which are estimated by integrating data from the registers with on field-collected data.

In the context of the PPHC, producing population and housing data at the sub-municipal level relies on the linkage between the Basic Register of Individuals (BRI) and the Basic Register of Places (BRP) (Carbonetti *et al.*, 2023; Carbonetti *et al.*, 2024). This linkage enables matching individuals and households to their respective dwellings and buildings, establishing a unique and consistent geo-coding for all statistical units relevant to the census. This approach allows for the provision of individual variables and essential cross-tabulations for census data dissemination, even at highly detailed spatial levels. Administrative data sources linked to the BRI can thus be geocoded and represented by statistics down to the EA level.

The main features of the new sub-municipal data production model are:

- ✓ Consistency in small-area data across population, households, dwellings, and buildings;
- ✓ Integrated use of spatial data sources for geocoding;
- ✓ Use of administrative data on dwellings and buildings;
- ✓ Strong focus on data quality through geographical accuracy and data validation processes;
- ✓ The ability to replicate the process annually.

This process continues to evolve with the availability of new and updated data sources, as well as methodological advancements in data linkage and geo-referencing techniques.

3. Opportunities for sub-municipal data through the integration of census and administrative data

The dissemination of EA-level data through the PPHC represents a notable shift from previous censuses. While there is now less availability of highly detailed information with respect of the previous methodology, there is greater potential for integrating census data with the BRP and statistical registers, as well as for georeferencing data. This approach reflects an evolving framework aimed at providing more accurate and timely data on an annual basis, which enables spatial studies at very granular scales.

In particular, the linkage of administrative data with the census database allows for the construction of specific population indicators, offering unprecedented opportunities to study demographic, social, and economic phenomena. Furthermore, the geocoding of information at the EA level, combined with annual database updates, supports comparative analyses across both spatial and temporal dimensions. This advancement opens new opportunities for exploring socio-economic population

dynamics at a territorial level, particularly for following changes and make comparisons on spatial and temporal dimensions, as well as for identifying sub-municipal domains in which the different phenomena are most concentrated.

4. The sub-municipal Socio-Economic Deprivation Index of population

Istat, in collaboration with several municipalities[1], is conducting an experimental project to examine household socio-economic deprivation at a sub-municipal level (Biasciucci *et al.*, 2023). This study leverages the data available at Istat from census and administrative sources, along with the capability to integrate these sources and to geocode the information.

A group of thematic experts has defined household deprivation as a tangible condition of deprivation, distinct from mere exposure to risk[2]. The adopted definition of socio-economic deprivation is as follows: "A condition in which households and individuals experience difficulties in adequately meeting their basic needs due to insufficient economic, employment, educational, social, and housing resources and opportunities." Based on this definition, deprivation was categorized into four key dimensions: economic, occupational, educational, and housing.

The available data sources at Istat were then evaluated in order to identify the most suitable ones, in terms of data availability and quality, for calculating specific indicators at the highest territorial granularity (EA level). For the project's purposes, these sources must be accurate, timely, annually updatable, and geo-codifiable.

Nine individual indicators relating to different components of deprivation were defined and calculated at EA-level using data from census and administrative sources (Table 1). These indicators were then synthesised through the Adjusted Mazziotta-Pareto Index (AMPI) methodology, a non-compensatory composite index (De Muro *et al.*, 2011; Mazziotta and Pareto, 2016; Mazziotta and Pareto, 2017) which made it possible to obtain the sub-municipal Socio-Economic Deprivation Index of population (SED-Index) at the EA-level. The AMPI methodology proved to be robust in terms of the analysis to be carried out[3].

---

[1] The municipalities currently involved in the project are: Bologna, Florence, Gorizia, Messina, Perugia, Milan, Modena, Palermo, Prato, Olbia, Rome, Taranto, Verona, and Parma.

[2] Istat already produces an indicator which measures the exposure to risk. 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).

[3] The validation phase aims to assess the robustness of the composite index in terms of its ability to produce correct and stable measures and its discriminatory power. Indeed, the results and the ranking of statistical units according to a composite index may depend to a large extent on the choices made in the previous stages (choice of individual indicators, normalization and aggregation, etc.). For this

Table 1 – *Individual indicators of socio-economic deprivation: an initial proposal.*

Individual indicator		Prevalent component	Sources
Dep1	Individuals aged 67 and over living alone, without home ownership (%)	Economic	PPHC, Cadastre
Dep2	Households in which no member is employed or receiving a pension (%)	Economic	PPHC, INPS
Dep3	Households not living in owner-occupied dwellings (%)	Economic	PPHC
Dep4	Employees aged 35-64 with a temporary employment relationship (%)	Employment	PPHC, INAIL, INPS
Dep5	Employment rate 35-64 years	Employment	PPHC
Dep6	Individuals aged 25-64 without upper secondary education (%)	Educational	PPHC
Dep7	Individuals aged 15-29 not working and not attending any regular course of study (%)	Educational	PPHC, MIM, MUR
Dep8	Individuals aged 18-24 without upper secondary education and not in education (%)	Educational	PPHC, MIM, MUR
Dep9	Ratio of occupants to rooms in occupied dwellings	Housing	PPHC, Cadastre

Sources: Permanent Population and Housing Census (PPHC - Istat); National Social Security Institute (INPS); National Institute for Insurance against Accidents at Work (INAIL); Ministry of Education (MIM); Ministry of University and Research (MUR); Register of dwellings and buildings (Cadastre).

A specific procedure is designed to identify 'critical areas' of concentration of household deprivation within the municipality. These areas are drawn as clusters of contiguous enumeration areas around the most critical EAs (i.e. those with the highest value of the SED-Index) according to statistical (internal homogeneity) and geographical (shape and extension) rules that are being defined.

The results, presented in the form of maps and indicators, will be a powerful tool for municipal administrators in planning and evaluating local socio-economic policies. In addition, the possibility of repeating the study every year will make it possible to monitor changes in the phenomenon and assess the effectiveness of the policies implemented in the area.

5. A case study: socio-economic deprivation in the municipality of Palermo

This section presents the results of a preliminary study on socio-economic deprivation of population in the municipality of Palermo. For this study, only

reason, it is necessary to carry out statistical analyses to examine the robustness of the rankings when an elementary indicator is included/excluded or when the methods used to construct the composite index vary. In this respect, an influence analysis showed that the AMPI method is the most robust for this type of synthesis.

enumeration areas (EAs) in the centre and with residential buildings were included. Consequently, out of the total 3,600 EAs in 2021, the analysis focuses on 2,642 EAs. All EAs in suburban areas and those in urban areas with few or no inhabitants (e.g., churches, hospitals, barracks, monuments, parks) are thus excluded. Once individual indicators were calculated at the EA level (Table 1), they were aggregated using the AMPI methodology to generate the SED-Index. Table 2 presents the minimum, median, mean, and maximum values for the nine individual indicators and the SED-Index measure across the 2,642 EAs in Palermo.

Table 2 – *Minimum, median, mean and maximum of indicator values (individual and SED-Index) calculated for the 2,642 EAs of the municipality of Palermo involved in the study.*

Indicator	Minimum	Median	Mean	Maximum
Dep1	0,0	8,3	11,8	100,0
Dep2	0,0	14,2	16,6	100,0
Dep3	0,0	34,5	38,4	100,0
Dep4	0,0	0,00	0,6	20,0
Dep5	0,0	55,9	55,7	100,0
Dep6	0,0	40,0	40,4	100,0
Dep7	0,0	29,2	31,7	100,0
Dep8	0,0	12,5	19,6	100,0
Dep9	0,0	0,4	0,4	2,0
SED-Index	86,9	100,0	101,1	130,2

Source: Istat (2021).

Table 3 shows the distribution of the SED-Index values calculated for the EAs.

Table 3 – *Distribution of SED-Index calculated for the 2,642 EAs of the municipality of Palermo involved in the study.*

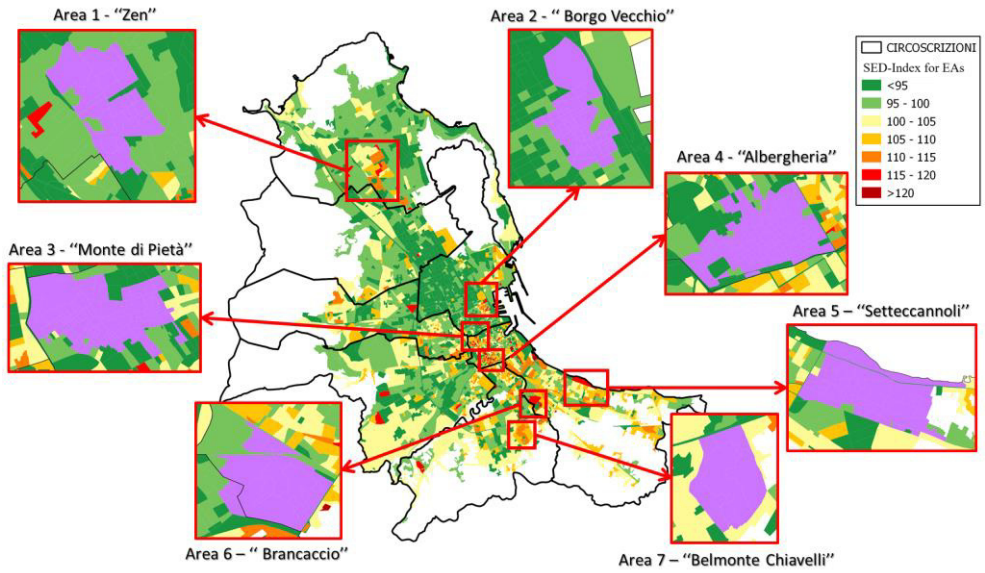
SED_Index	No. of EAs	%
< 95	597	22.6
95 - 100	727	27.5
100 - 105	588	22.3
105 - 110	409	15.5
110 - 115	221	8.4
115 - 120	68	2.6
> 120	32	1.2
Total	2,642	100.0

Source: Istat (2021).

It can be seen that in 50.1% of the cases the value is lower than the average municipal value, used as a reference (=100); these EAs therefore do not have situations of generalised deprivation. On the other hand, in 12.2% of the cases situations of high deprivation (SED-Index > 110) are observed; in particular, in 32 EAs (about 1.2%) the SED-Index indicates cases of very high deprivation (SED-Index > 120).

Figure 1 presents the territorial distribution of the SED-Index on the 2,642 EAs of Palermo involved in the study.

Figure 1 – *Spatial classification of SED-Index values for the 2,642 EAs of the municipality of Palermo involved in the study and identification of some potential critical areas.*



Through visual analysis of the territorial distribution, areas with the highest levels of household deprivation (SED-Index > 110) were identified as potentially critical. In collaboration with the Statistics Office of the Municipality of Palermo, seven "critical areas" were identified (highlighted in purple within red boxes in Figure 1), located both in the city center (areas 2, 3 and 4) and in more peripheral zones (areas 1, 5, 6, and 7). Figure 1 illustrates these areas and the administrative districts of Palermo where the critical areas are located.

For these seven critical areas, the individual indicators used to assess deprivation were recalculated (Table 4).

Table 4 – *Values of the individual indicators used to study deprivation calculated for the municipality of Palermo and the seven critical areas identified in the municipality.*

Indicator	Palermo	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7
Dep1	8.9	10.8	17.8	18.7	20.5	9.1	8.7	9.6
Dep2	13.8	23.3	22.1	26.8	29.6	15.3	18.5	20.9
Dep3	34.6	76.1	59.8	63.1	64.9	47.1	44.7	46.9
Dep4	0.6	0.5	0.7	0.6	0.8	0.8	0.5	1.7
Dep5	55.1	32.7	49.5	48.3	45.6	41.1	40.0	43.0
Dep6	41.9	76.2	59.7	57.7	63.4	64.9	64.7	58.4
Dep7	32.3	54.0	52.7	54.7	54.2	45.7	50.7	41.1
Dep8	19.8	41.6	47.9	51.4	44.0	33.0	38.3	32.7
Dep9	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5

Source: see Table 1.

In order to enrich the information on deprivation, some socio-demographic indicators have also been calculated. Although these are not included in the definition of socio-economic deprivation, they can be very useful in defining the socio-demographic profile of individuals and households living in critical areas, in order to make the most appropriate territorial policy choices to support the disadvantaged population living in these areas. These context indicators are (Table 5):

- *average number of members per household* (Dem1)
- *percentage of foreigners* (Dem2)
- *employment rate (15+)* (Dem3)
- *percentage of youth (0-24)* (Dem4)
- *percentage of elderly (65+)* (Dem5)
- *ratio of young (0-24)/old (65+)* (Dem6)
- *foreigners/Italians ratio* (Dem7)
- *percentage of single-member households* (Dem8)
- *percentage of households with 5+ members* (Dem9)
- *percentage of graduates (25-64)* (Dem10).

A first reading of the data relating to the critical areas shows that, compared with the average for the municipality as a whole, they have in common a higher percentage of 'households in which no member is working or receiving a pension from work' (Dep2), a lower 'employment rate of adults (35-64 years old)' (Dep5), a very low level of education (Dep8; Dem10), a higher presence of 'young people (0-24 years old)' (Dem4) and 'large households (with 5 or more members)' (Dem9). The critical areas also show high and similar values for the 'percentage of persons aged 25-64 with less than upper secondary education' (Dep6) and the 'ratio of persons to rooms in occupied dwellings' (Dep9).

Table 5 – *Values of some socio-demographic indicators calculated for the municipality of Palermo and the seven critical areas identified in the municipality.*

Indicator	Palermo	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7
Population	626,619	11,926	6,308	5,134	7,122	17,324	6,846	2,045
Households	259,647	3,904	2,627	2,379	3,215	6,429	2,580	703
Dem1	2.4	3.1	2.4	2.2	2.2	2.7	2.7	2.9
Dem2	3.8	3.1	22.1	19.3	21.1	0.6	1.0	2.3
Dem3	36.0	23.9	36.1	35.4	33.0	27.5	27.5	31.9
Dem4	24.5	33.4	28.3	27.5	28.7	26.1	28.0	34.3
Dem5	22.3	14.6	14.7	13.4	13.1	21.7	19.7	13.3
Dem6	1.1	2.3	1.9	2.1	2.1	1.2	1.4	2.6
Dem7	0.0	0.0	0.3	0.2	0.3	0.0	0.0	0.0
Dem8	31.7	18.1	37.9	47.1	47.2	23.4	25.5	20.8
Dem9	6.3	18.9	8.8	7.8	8.4	10.8	10.5	10.7
Dem10	22.0	3.2	14.4	17.2	13.7	7.7	7.5	8.6

Source: Permanent Population and Housing Census, Istat (2021).

The seven critical areas shown in Figure 1 can be divided into three groups:
a) the central critical areas (Areas 2, 3 and 4)
b) the south-eastern peripheral critical areas (Areas 5, 6 and 7);
c) the north-western peripheral critical area (Area 1).

The critical areas in group (a) are characterized by high levels of economic and educational deprivation. Among the resident population in these areas, there is a high rate of foreigners (Dem2: 19.3–22.1%) and employment rate higher than the other critical areas (Dep5: 45.6–49.5%; Dem3: 33–36.1%). These areas also show a significant share of single-person households (Dem8: 37.9–47.2%).

In contrast, the areas in group (b) have a very low percentage of foreign residents (Dem2: no more than 2.3%), a relatively high proportion of large households (Dem9: slightly over 10%), and employment rates below the municipal average (Dep5: 40-43%; Dem3: 27.5–31.9%).

Area 1, located within the "Zen" district, warrants separate discussion. This area displays high levels of economic deprivation, with 76.1% of households in rented accommodation (Dep3), as well as high levels of employment and educational deprivation: the employment rate for those aged 35–64 is 32.7% (Dep5), and 76.2% of individuals aged 25–64 have not completed the upper secondary education (Dep6). This area also has a notable proportion of young residents (Dem4: 33.4% are aged 24 or younger) and a considerably high average household size compared to other areas (Dem1: 3.1%), partly due to the presence of large households (Dem9: 18.9% with five or more members). The employment rate for individuals aged 15 and older (Dem3: 23.9%) is also lower than in other critical areas. Finally, in terms of education, there

are the lower university graduates between the individuals aged 25-64 years (Dem10: 3.2%).

6. Conclusions

In this paper, we present an experimental study conducted by Istat on the socio-economic deprivation of households at a highly detailed territorial level, enabled through the integration of data from the Italian permanent population census with administrative archives available to Istat. Recent advancements in methodology and IT solutions have significantly facilitated the merging of census data with administrative records, allowing for the precise geo-coding of households to their residential addresses and the creation of the Integrated System of Statistical Registers. This new data infrastructure supports spatial and temporal analyses at an unprecedented level of granularity across various topics. The availability of annual data series with a high level of territorial detail, both from census and administrative sources, homogeneous for all municipalities, will allow spatial and temporal analyses.

The study of household deprivation presented here exemplifies the new possibilities for analyzing territorial patterns through the integration of data from the Italian permanent census and administrative archives. An important aspect of this project is the involvement of municipalities, which possess in-depth knowledge of their local areas and can provide essential insights for interpreting the results accurately. Istat's collaboration with municipal statistical offices enhances this interpretive framework.

The definition of socio-economic deprivation adopted in this paper is: "A condition in which households and individuals experience difficulties in adequately meeting their basic needs due to insufficient economic, employment, educational, social, and housing resources and opportunities." Based on this definition, deprivation was categorized into four key dimensions (economic, occupational, educational, and housing) and measured through nine individual indicators. These indicators were then synthesised through the Adjusted Mazziotta-Pareto Index (AMPI) methodology, a non-compensatory composite index which made it possible to obtain the sub-municipal Socio-Economic Deprivation Index of population (SED-Index) at the EA-level. In this paper, the results for the municipality of Palermo are presented as an experimental exercise in order to show the potentials of this indicator. Indeed, this indicator can be calculated down to the enumeration areas of the municipalities, allowing the identification of some critical areas on which it was possible to focus attention for further investigation. This can, for example, be a very powerful tool for local policy makers in allocating economic resources to tackle social exclusion and to reduce inequalities.

It is important to emphasize that this study remains at an experimental stage, and that the individual indicators used may be supplemented or replaced by others that could prove more effective as the analysis of the extensive administrative sources at Istat progresses.

Acknowledgements

We would like to thank Dr. Girolamo D'Anneo, Head of the Statistics Office of the Municipality of Palermo and President of the Union of Statistics of Italian Municipalities (USCI), for his support in the identification and validation of critical areas in Palermo and in the statistical analysis phase.

References

BIASCIUCCI F., CARBONETTI G., FERRARA R., MAZZIOTTA M., QUONDAMSTEFANO V. 2023. New scenarios for measuring household deprivation in sub-municipal areas. *Italian Journal of Applied Statistics*, n. 3. Book of Short Papers. Vol. 35, No. 3 Supplement.

CADUM E., COSTA G., BIGGERI A., MARTUZZI M. 1999. Deprivazione e mortalità: un indice di deprivazione per l'analisi delle disuguaglianze su base geografica. *Epidemiol Prev*, Vol. 23, No.3, pp. 175-187.

CARANCI N., BIGGERI A., GRISOTTO L., PACELLI B., SPADEA T., COSTA G. 2010. L'indice di deprivazione italiano a livello di sezione di censimento: definizione, descrizione e associazione con la mortalità. *Epidemiol Prev*, Vol. 34, No. 4, pp. 167-176.

CARBONETTI G., DE MATTEIS G., DI ZIO M., FARDELLI D., FERRARA R., LIPIZZI F. 2023. Enumeration area imputation methods for producing sub-municipal data in the Italian permanent population and housing census. *Statistical Journal of the IAOS*, Vol. 39, No. 1, pp. 123-136. IOS Press.

CARBONETTI G., FERRARA R., GALLO G., VERRASCINA M. 2024. The evolution of the data production model in Istat. New perspectives for the analysis of population socio-economic phenomena. In *Proceedings of the 11th European Conference on Quality in Official Statistics (Q2024)*, Estoril, Portugal 4-7 June 2024.

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.

FALORSI S. 2017. Census and Social Surveys Integrated System. *Note by the National Institute of Statistics of Italy, presented at the UNECE/Eurostat Group of*

Experts on Population and Housing Censuses, Nineteenth Meeting, Geneva, Switzerland, 4–6 October 2017.

GALLO G., ZINDATO D. 2018. Italy case study. In *Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses*, UNECE New York and Geneva, 2018.

GUILLAUME E., PORNET C., DEJARDIN O. *et al.* 2016. Development of a cross-cultural deprivation index in five European countries. *J Epidemiol Community Health*, Vol. 70, No. 5, pp. 493-499.

ISTAT – 2020. *Le misure della vulnerabilità: un'applicazione a diversi ambiti territoriali*, Istat, Roma, 2020.

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.

MAZZIOTTA M., PARETO A. 2017. Synthesis of indicators: the composite indicators approach. In FILOMENA MAGGINO (Ed.) *Complexity in Society: From Indicators Construction to their Synthesis*. *Social Indicators Research Series*. Springer, pp.159-191.

ROSANO A., PACELLI B., ZENGARINI N., COSTA G., CISLAGHI C., CARANCI N. 2020. Aggiornamento e revisione dell'indice di deprivazione italiano 2011 a livello di sezione di censimento. *Epidemiol Prev*, Vol. 44, No. 2-3, pp. 162-170.

Giancarlo CARBONETTI, Istat, carbonet@istat.it
Flavio BIASCIUCCI, Istat, biasciucci@istat.it
Andrea CUTILLO, Istat, cutillo@istat.it
Matteo MAZZIOTTA, Istat, mazziott@istat.it
Valeria QUONDAMSTEFANO, Istat, quondamstefano@istat.it
Maria Tiziana TAMBURRANO, Istat, tamburra@istat.it
Debora TRONU, Istat, tronu@istat.it

INVESTIGATING THE RELATIONSHIP BETWEEN MATERIAL DEPRIVATION AND HOUSEHOLD CONTEXT AMONG OLDER PEOPLE IN ITALY

Maria Carella, Roberta Misuraca, Anna Paterno

Abstract. A large body of literature has delved into socioeconomic inequalities, emphasizing the multidimensional nature of deprivation. However, previous research has only considered deprivation in additive terms and has not adequately combined the variety and intensity of different deprivation indicators. Using data from the 2018 Multipurpose Survey on Households, Aspects of Daily Life (Italian National Institute of Statistics), we conducted a Latent Class Analysis to identify distinct deprivation profiles based on gradual patterns of hardship. This method allows for a better understanding of the interplay between varying degrees of severity across different deprivation indicators. In a second step, probit regression analyses were used to examine the impact of diverse living arrangements on latent deprivation classes among older individuals. Our findings reveal that older people living with others experience higher levels of housing deprivation but lower overall material deprivation than those living alone. Furthermore, living with family members, whether a spouse or descendants, is associated with higher levels of housing deprivation than cohabiting with strangers, while the reverse is true for overall material deprivation. These results underscore the need for further research to explore the varying degrees of deprivation across subpopulations, highlighting the importance of nuanced investigations in this field.

1. Introduction

Over the past two decades, an extensive body of literature has explored socioeconomic inequalities, focusing on the multidimensional nature of deprivation (Kim, 2016). In shaping various patterns of material deprivation among older individuals, scholars have highlighted the role played by individual socio-demographic characteristics, such as education level (Pham and Mukhopadhaya, 2018) and health conditions (Terraneo, 2017), as well as contextual factors like the role of the welfare system (Zaidi, 2012).

Regarding Italy, many studies have investigated material deprivation among older individuals from different points of view. Some of these have explored the determinants of material deprivation and its different measurement approaches (Istat, 2024; Mussida and Parisi, 2019; Terraneo, 2017), while a strand of research on this topic has focused on the additive concept of deprivation as the sum of different

degrees of hardship in different deprivation items (Andress *et al*. 2001). In this study, we aim to delve into the multifaceted nature of deprivation, considering the simultaneous combination of hierarchical degrees of hardship (frequency) in different deprivation indicators (variety). To achieve this purpose, we used data from the 2018 Multipurpose Survey on Households Aspects of Daily Life (Italian National Institute of Statistics) to identify distinct deprivation profiles based on their patterns of hardship. In the first phase, we conducted a Latent Class Analysis, while in a subsequent step Probit regression analyses were employed to explore the impact of diverse living arrangements (living alone, living with family members, i.e. spouse or descendants, and living with people other than family members, such as roommates or caregivers) among older individuals on latent deprivation classes (specifically, food deprivation, housing deprivation, and overall material deprivation) while controlling for variables such as income sources, education levels, and the existence of physical limitations. This perspective is better suited to offer a comprehensive understanding of poverty in its broader meaning of material deprivation.

2. Literature

In recent years, various approaches have been employed to assess disparities in poverty and the related concept of deprivation between different target groups (Navarro and Ayala, 2008). Starting from a conceptualization of poverty, earlier research predominantly employed an income-based metric to investigate poverty trends among the elderly, owing to its simplicity, international comparability, and widespread accessibility (Ringen, 1988). However, poverty metrics primarily reliant on income suffer significant drawbacks (Cheung and Chou, 2018). First, while household income-based measures offer insight into an individual's immediate financial status, they fall short of fully capturing the multidimensional nature of deprivation. This limitation appears especially pronounced among older individuals (Breheny *et al*., 2016), reflecting the common scenario of being "asset-rich but cash-poor" (Sullivan *et al*., 2008). To overcome these constraints, some studies have considered a living standard deprivation measure (Andress *et al*. 2001; Boarini and d'Ercole 2006), able to isolate the mere concept of disposable income and shed light on the effective living conditions of individuals (Pfoertner *et al*., 2011). Compared to income-based indicators, the deprivation approach based on living standards offers several advantages. First of all, it allows us to emphasize the role of deprivation as a multidimensional concept. Secondly, it directly pertains to an individual's actual living situation. Although numerous works have investigated the role of economic and social dimensions in a comprehensive notion of deprivation, a consensus on a definition of the term has yet to be reached (Pirani, 2013). The

acknowledged definition of material deprivation is grounded in the premise that deprivation is a multifaceted phenomenon, encompassing various aspects of daily life and not solely tied to an individual's economic status. It includes limitations across numerous material and sociodemographic indicators that are generally "*considered by most people to be desirable or even necessary to experience an adequate quality of life*" (Eurostat, 2023). These indicators include the ability to afford a meal, maintain a sufficiently warm household, own durable goods such as a car, telephone, or personal computer, and have housing amenities like an indoor flushing toilet, as well as the size and condition of the home (Istat, 2024; Kim, 2016). However, the existing studies tend to consider the concept of material deprivation only in additive terms, neglecting to take into account the simultaneous combination of hierarchical degrees of hardship (frequency) in different deprivation indicators (variety). A large strand of studies report that age has the potential to moderate the perception of the deprivation effects (Kwan and Walsh, 2018; Doebler and Glasgow, 2017). Traditionally, three different theories are called into question to explain the moderating effect that age can have on deprivation. According to the Life-Cycle Effect Theory, material deprivation may have a more pronounced impact on older age groups due to the greater difficulties they encounter in accessing alternative sources of income (Kwan and Walsh, 2018), strictly connected to declining work capacity, reduced physical mobility, and diminished social connections, particularly after retirement. Similarly, the Cumulative Disadvantage Theory has postulated that individual disadvantages (originating from characteristics such as ethnicity or socioeconomic class) accumulate over the life course from childhood to adulthood and tend to intensify in older age (Arber *et al*., 2014), heightening the adverse effects of deprivation in later life. Finally, other studies have identified age as a "leveler" (Herd, 2006). Specifically, the progressive frailty experienced by individuals in later life, regardless of their socioeconomic status, tends to equalize the potential adverse effects of deprivation between older age groups.

The discussion about the influence of different living arrangements on the material deprivation experienced by older individuals continues to generate diverse viewpoints. Some studies suggest that certain living arrangements, such as living with a partner, can mitigate the effects of material deprivation and poverty (Ku and Kim, 2020; Karagiannaki and Burchardt, 2020). The main idea is that family ties serve as a financial support mechanism for older individuals (Smeeding *et al*., 2008). In multigenerational extended families, the market incomes of younger family members can be shared with older adults (Smeeding *et al*., 2008). Karagiannaki and Burchardt (2020), exploring how living arrangements affect material deprivation across Europe, find significant disparities in various deprivation indicators, with individuals living alone facing higher levels of deprivation. Their study also shows that cohabiting with a partner offers economic support, mitigating the risk of material

deprivation. In this direction, Meemon and Paek (2020) investigating the association between different living arrangements and material well-being among individuals over 55 years in Thailand, reveal that living with a partner generally results in lower material deprivation. Older individuals living with a partner show typically better material well-being than those living alone, shedding light on the protective effect of familiar cohabitation. An interesting study in this vein is Ku *et al*. (2021). This study analyzes the shifts in income distribution among older adults in South Korea from 1996 to 2016. The authors identify the transition in living arrangements—from extended families to single-member households—as a key factor exacerbating income inequality, particularly in the presence of a decline in market income, increasing inequality, and poverty. Regarding Italy, a large strand of literature investigated material deprivation among individuals over 55 years old in Italy from different perspectives (Istat, 2024, 2020; Mussida and Parisi, 2019; Terraneo, 2017). One of these lines of research aims to identify the determinants of material deprivation and its different measurement approaches (Istat, 2024; Mussida and Parisi, 2019), often offering cross-country comparisons (Whelan and Maître, 2012). Whelan and Maître (2012) emphasize the significance of non-monetary measures of deprivation due to the limitations of income and related poverty metrics, including basic needs, consumption, household conditions, health, neighborhood environment, and access to public facilities. Focusing on older individuals, Bertoni *et al*. (2015) explore the determinants of the multidimensional concept of poverty and material deprivation across European countries. More specifically, they focus on specific indicators such as housing conditions, basic needs, and financial security, emphasizing the necessity of addressing multiple dimensions of deprivation to comprehensively grasp the economic challenges experienced by the older. The study shows a clear geographical gradient in the material deprivation of older adults in Europe. Scandinavian countries exhibit the lowest levels of deprivation, whereas higher levels are observed in Southern European countries (such as Italy), representing a significant factor in social exclusion among older individuals in these regions. The importance of considering a comprehensive approach to address the multifaceted nature of deprivation, especially for the older population, is also highlighted by other studies (Vignoli and De Santis, 2010). Regarding the relationship between living arrangements and material deprivation among older individuals in Italy, not many studies exist that directly investigate the differentiated effects of material deprivation among various household contexts among individuals over 55. One exception is Ivaldi (2016). This study examines two distinct forms of deprivation (material and social) on a regional basis in Italy. Generally, the findings indicate that older individuals living alone (or in single-parent households) tend to experience higher levels of both material and social deprivation, shedding light on the supportive role of family. In the Southern regions, a strong sense of family

attachment mitigates the negative effects of living alone, due to robust family networks.
Drawing from prior literature, two research hypotheses are formulated:
RQ1: Which sociodemographic characteristics of older people exhibit a stronger association with the different dimensions of material deprivation?
RQ2: Which dimension of material deprivation is more likely to be experienced by older individuals differentiating by living arrangements?

3. Data and methodology

We use data from the 2018 Italian Multipurpose Survey on Households "Aspects of Daily Life", conducted by the Italian National Institute of Statistics. After selecting individuals aged 55 and above, we obtained a sample of 16,494 individuals. Following the material deprivation concept and measures adopted by Eurostat (2023), we collected data on material and social indicators of deprivation such as *i)* the capacity to afford a meal with meat, chicken, and fish; *ii)* maintaining a sufficiently warm household; *iii)* having some durable goods such as a car, personal computer, telephone, and washing machine; *iv)* the presence of housing amenities like an indoor flushing toilet; *v)* the size and living condition of the home. Additionally, we included data on the degradation of the neighborhood (pollution, crime, violence, and noise). Finally, we also consider information concerning the ability to cover essential expenses such as housing costs. The measurement scale for items concerning food and warm households ranges from 1 to 3 (more than 2 times per week/1 time per week/ rarely or never), while dummy variables are considered for other deprivation items. Once collected information about the main deprivation indicators and their frequency, we employed a Latent Class Analysis (LCA) to derive distinct deprivation profiles[1]. Our goal is to explore how the joint influence of the diversity and intensity of various deprivation items may lead to distinct patterns of deprivation. To do that, different degrees of hardship relating to different deprivation indicators are taken into account, at the individual level. More in-depth, through LCA we are able to assign each member of the sample to the specific latent class that best fits their deprivation characteristics. These classes are mutually exclusive: each individual belongs to only one class. Through this methodology, we can cluster individuals exhibiting comparable deprivation patterns in terms of intensity of hardship and variety of deprivation indicators, allowing for further empirical analyses of these delineated subgroups. Therefore, we first evaluate different models with different numbers of classes, spanning from a single class to up to five classes.

[1] The LCA specification of our model is available upon request.

The optimal number of classes will be determined by considering the AIC and BIC criteria[2]. Consequently, we identify three distinctive deprivation profiles, that is *food deprivation*, *housing deprivation*, and *material deprivation*, labelled according to the features of their specific characteristics of deprivation. Regarding our main explanatory variable, we gathered information on the living arrangements of older individuals in Italy. We collapse the observations to obtain the following three categories: living alone, living with family members (i.e. spouse or descendants), and living with people other than family members (i.e. roommates or caregivers). In all our analysis, we also control for some demographic and territorial characteristics, such as gender, age, labor status, educational levels, presence of physical limitations, economic resources, and the region of residence.

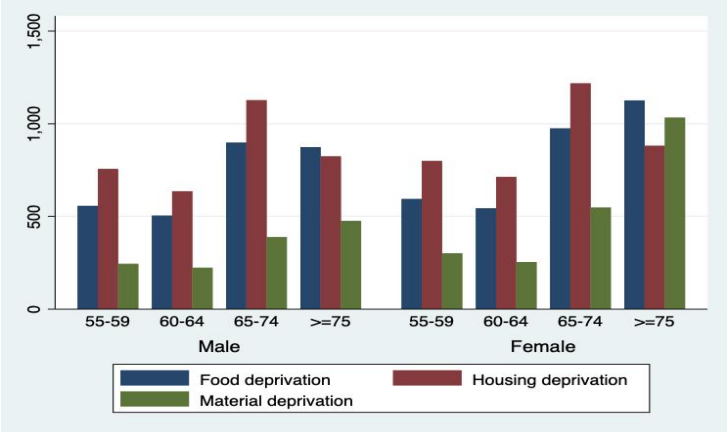
4. Results

The profiles derived from the LCA reflect the results of various combinations of the diversity and frequency of individual-level deprivation patterns[3]. These combinations encompass different degrees of deprivation across multiple items. The first profile is *food deprivation*, including individuals who have a very high likelihood of insufficient access (rarely or never) to an adequate quantity and quality of nutritious food, despite having proper housing conditions and an acceptable standard of living. This analysis specifically focuses on access to meals containing meat, fish, and chicken (Istat, 2024). The second profile is *housing deprivation*. This group includes individuals encountering significant housing difficulties and living in areas with neighborhood deterioration (elevated crime levels, violence, vandalism, and pollution). Additionally, individuals in this profile are very likely to reside in small and unaffordable homes and are less likely to own common consumer durables, such as televisions, mobile phones, and washing machines. Finally, the third profile is high *material deprivation*. This group includes individuals with a very high likelihood of experiencing all the previous deprivation indicators. It could be considered the most serious and comprehensive form of material deprivation. The most representative profile in our data is housing deprivation (41%), followed by food deprivation (37%). The most serious and comprehensive form of deprivation (material deprivation) is represented within the sample by 22%.[4]

[2] Given space limitations, Akaike's information criterion, Bayesian information criterion, and Goodness of fit measures for LCA are available upon request.
[3] Given space limitations, the parameter estimates obtained through LCA implementation and the conditional probabilities are available upon request.
[4] Given space limitations, the marginal probabilities of each class in our sample are available upon request.

Figure 1 shows the distribution of the deprivation profiles by age and gender (RQ1). These findings indicate that food deprivation is homogeneously distributed across genders even considering a light peak for women over 75 years although it shows higher values compared to more severe forms of deprivation (e.g., material). Housing deprivation, on the other hand, presents more elevated values. More in-depth, the risk of housing deprivation is higher in the 65-74 age group, while the likelihood of experiencing food deprivation is decreasing. Additionally, figure 1 reveals a clear peak in material deprivation within the over-75 age group, with a more pronounced effect observed among women over 75 years. This suggests that the oldest individuals experience worse conditions, mainly linked to increased social isolation and growing vulnerability in older age. This finding aligns with the Life-Cycle Effect Theory, which suggests that material deprivation could have a more significant impact on older age groups due to the greater difficulties related to declining work capacity, reduced physical mobility, and diminished social connections, particularly after retirement. Examining gender differences, our analysis also reveals that the most marginalized socio-demographic categories (women over 75) seem experiencing the most severe form of material deprivation.

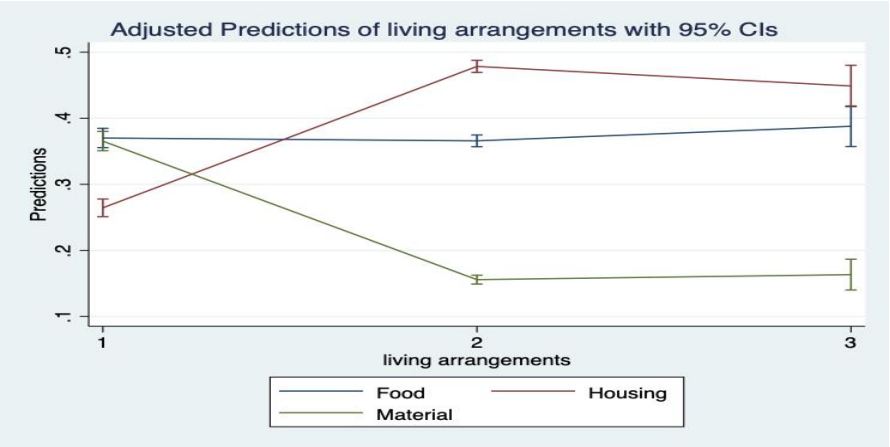
Figure 1 – *Sample distribution (sum) of the deprivation profiles by age and gender.*



Subsequently, we implement a Probit estimation to explore the impact of diverse living arrangements (as explanatory variables) among older individuals on latent deprivation classes (as dependent variables) obtained through LCA while controlling for some demographic and territorial characteristics, such as gender, age, labour status, educational levels, presence of physical limitations, economic resources, and

the region of residence. To assess the fit of the Probit model, we employed several measures of goodness of fit, including McFadden's pseudo R-squared, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC)[5].Figure 2 shows the average predicted value for food, housing, and material deprivation given various household contexts, such as living alone, living with family members (spouse or descendants), and living with people other than family members (e.g., roommates or caregivers).

Figure 2 – *Predictive Margins with 95% CI of the impact of diverse living arrangements among older individuals on latent deprivation classes.*



1: living alone; 2: living with family members; 3: living with people other than family members

We find that food deprivation is evenly distributed across older people regardless of the household context to which they belong. Conversely, older individuals living with others, instead of living alone, experience heightened levels of housing deprivation but a decrease in more serious overall material deprivation. In this context, living with others, whether family members or strangers, appears to exacerbate the challenges associated with cohabitation while improving the hardship associated with a more serious form of material deprivation. This result suggests that living with other adults could represent a crucial strategy for older individuals to avoid poverty and deprivation. In this case, the market incomes of younger household members are often shared with older adults. This is true above all in Italy

[5] For the food deprivation model, AIC: 19941.79; BIC: 20094.65 and Pseudo R-squared: 0.018. For the housing deprivation model, AIC: 19297.68; BIC: 19450.54 and Pseudo R-squared: 0.083. For the material deprivation model; AIC: 13451.7; BIC: 13604.56 and Pseudo R-squared: 0.150.

where the lack of opportunities caused by housing market policies and characteristics such as restrictions in the access to credit or rigid mortgage regulations and the unavailability of housing for young people discourage young adults from leaving the parental home. Additionally, figure 2 shows that living with family members, whether spouse or descendants, is associated with higher levels of housing deprivation compared to cohabiting with people other than family members (e.g., roommates or caregivers), while the reverse is observed for high material deprivation.

5. Conclusions

This study provides insights into the potentially pivotal role of graduality in understanding the multidimensional nature of material deprivation. Through the implementation of a novel approach based on the latent class method, we can go into the concept of material deprivation and segment the sample into distinct deprivation profiles (food, housing, and material deprivation), clustering individuals exhibiting comparable deprivation patterns in terms of intensity of hardship and variety of deprivation indicators. In line with previous literature (Myck *et al*., 2020; Kwan and Walsh, 2018), we find that the most severe forms of material deprivation are experienced by the population over 75 and the female population (RQ1). This could be mainly linked to increased social isolation and growing vulnerability in older age (Myck *et al*., 2020). We also find that older individuals living with others, instead of living alone, experience a decrease in more serious overall material deprivation. This evidence aligns with studies showing that living with other adults is vital for older individuals to avoid poverty and deprivation (Smeeding *et al*., 2008). Additionally, living with family members, compared to cohabiting with people other than family members (e.g., roommates or caregivers), is associated with lower levels of material deprivation (RQ2). This result supports recent European studies that highlight the positive impact of cohabitation with a partner or other family members (Karagiannaki and Burchardt, 2020).

Summarizing, this contribution, which is an evolution and in-depth analysis of a preliminary study on this topic (Carella *et al*. 2025), complements the body of research on the multidimensional nature of deprivation among aging adults. It highlights the importance of considering a multifaceted approach to material deprivation and the crucial role played by the combined interplay of variety and intensity of marginal deficiencies and varying degrees of severity across different deprivation items. Moreover, our approach aligns with recent studies that emphasize the importance of non-monetary measures of deprivation. Addressing deprivation in the aging population is crucial for formulating effective policies that promote well-

being and social inclusion. Additionally, it sheds light on the precarious conditions faced by elderly individuals living alone or with non-family members or caregivers.

Acknowledgements

We acknowledge funding from Next Generation EU, in the context of the National Recovery and Resilience Plan, Investment PE8 – Project Age-It: "Ageing Well in an Ageing Society" [DM 1557 11.10.2022]. The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

ANDREß H.J., LIPSMEIER G., LOHMANN H. 2001. Income, Expenditure and Standard of Living as Poverty Indicators-Different Measures, Similar Results?, *Journal of Contextual Economics*, pp.165-198.

ARBER S., FENN K., MEADOWS R. 2014. Subjective financial well-being, income and health inequalities in mid and later life in Britain, *Social Science & Medicine*, 100, pp. 12-20.

BERTONI M., CAVAPOZZI D., CELIDONI M., TREVISAN E. 2015. Development and validation of a material deprivation index, *Ageing in Europe-supporting policies for an inclusive society*, pp. 57-65.

BOARINI R., D' ERCOLE M. M. 2006. Measures of Material Deprivation in OECD Countries, *OECD Social, Employment and Migration Working Papers*, No. 37, OECD Publishing, Paris.

BREHENY M., STEPHENS C., HENRICKSEN A., STEVENSON B., CARTER K., ALPASS F. 2016. Measuring living standards of older people using Sen's Capability Approach: development and validation of the LSCAPE-24 (Living Standards Capabilities for Elders) and LSCAPE-6, *Ageing and Society*, pp. 307-332.

CARELLA M., MISURACA R., PATERNO A. 2025. Inequalities Among Older People: exploring differences in deprivation patterns. In: Alessio Pollice and Paolo Mariani (eds) *Methodological and Applied Statistics and Demography*, SIS 2024, Short Papers, Springer.

CHEUNG K.C.K., CHOU K. L. 2018. Child poverty among Hong Kong ethnic minorities, *Social Indicators Research*, Vol. 137, pp. 93-112.

DOEBLER S., GLASGOW N. 2017. Relationships between deprivation and the self-reported health of older people in Northern Ireland, *Journal of aging and health*, Vol. 29, No. 4, pp. 594-619.

EUROSTAT 2023. Living conditions in Europe - material deprivation and economic strain. Eurostat Report.

HERD P. 2006. Crediting care or marriage? Reforming Social Security family benefits, *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, Vol. 61, No. 1, pp. S24-S34.

ISTAT 2024. Condizioni di vita e reddito delle famiglie, anno 2023.

ISTAT 2020. Condizioni di vita dei pensionati, anno 2018.

IVALDI E. 2016. Material and social deprivation in Italy: An analysis on a regional basis, *Eastern European Business and Economics Journal*, Vol. 2, No. 3, pp. 248-268.

KARAGIANNAKI E., BURCHARDT T. 2020. Intra-household inequality and adult material deprivation in Europe. *CASE Papers,* Vol.218, Centre for Analysis of Social Exclusion, LSE.

KIM S. G. 2016. What have we called as "poverty"? A multidimensional and longitudinal perspective, *Social Indicators Research*, *129*, pp. 229-276.

KU I., KIM C. O. 2020. Decomposition analyses of the trend in poverty among older adults: the case of South Korea, *The Journals of Gerontology: Series B*, Vol.5, No. 3, pp. 684-693.

KU I., LEE W., LEE S. 2021. Declining family support, changing income sources, and older people poverty: Lessons from South Korea, *Population and Development Review*, Vol. 47, No. 4, pp. 965-996.

KWAN C., WALSH C. A. 2018. Old age poverty: A scoping review of the literature, *Cogent Social Sciences*, Vol. 4, No. 1: 1478479

MEEMON N., PAEK S. C. 2020. Older adults living alone in Thailand: Socioeconomic inequality and its relation to unmet health needs. *Asia-Pacific Social Science Review*, Vol. *20*, No.4: 3.

MUSSIDA C., PARISI M. L. 2019. Features of personal income inequality before and during the crisis: An analysis of Italian regions, *Regional Studies*.

MYCK M., NAJSZTUB M., OCZKOWSKA M. 2020. Implications of social and material deprivation for changes in health of older people, J*ournal of Aging and Health*, Vol. 32, No. 5-6, pp. 371-383

NAVARRO C., AYALA L. 2008. Multidimensional housing deprivation indices with application to Spain, *Applied Economics*, Vol. 40, No. 5, pp. 597-611.

PFOERTNER T. K., ANDRESS H. J., JANSSEN C. 2011. Income or living standard and health in Germany: different ways of measurement of relative poverty with regard to self-rated health, *International journal of public health*, Vol. 56, pp. 373-384.

PHAM A. T. Q., MUKHOPADHAYA P. 2018. Measurement of poverty in multiple dimensions: The case of Vietnam, *Social Indicators Research*, Vol. 138, No.3, pp. 953-990.

PIRANI E. 2013. Poverty and social exclusion in Europe: Differences and similarities across regions in recent years, *Advances in Latent Variables-Methods, Models and Applications*.

RINGEN S. 1988. Direct and indirect measures of poverty, *Journal of Social Policy*, Vol. 17, No. 3, pp. 351-365.

SMEEDING T. M., GAO Q., SAUNDERS P., WING C. 2008. Elder poverty in an ageing world: Conditions of social vulnerability and low income for women in rich and middle-income nations. *LIS Working Paper Series* n. 497.

SULLIVAN D. H., ZIEGERT A. L. 2008. Hispanic immigrant poverty: Does ethnic origin matter? *Population Research and Policy Review*, Vol. 27, pp. 667-687.

TERRANEO M. 2017. Material Deprivation, Health and Well-Being in Older Age: A Comparative Analysis of 14 European Countries, Global Perspectives on Wealth and Distribution, in: WHITE R. (Ed.), *Measuring Multidimensional Poverty and Deprivation*, pp. 113-137. Palgrave Macmillan.

VIGNOLI D., DE SANTIS G. 2010. Individual and contextual correlates of economic difficulties in old age in Europe, *Population Research and Policy Review*, Vol. 29, pp. 481-501.

WHELAN C. T., MAÎTRE B. 2012.Understanding material deprivation: A comparative European analysis, *Research in Social Stratification and Mobility*, Vol. 30, No. 4, pp. 489-503.

ZAIDI A.2012.Exclusion from material resources: poverty and deprivation among older people in Europe. In *From exclusion to inclusion in old age*, The Policy Press, pp. 71-88.

Maria CARELLA, Department of Political Science, University of Bari Aldo Moro, maria.carella1@uniba.it

Roberta MISURACA, Department of Political Science, University of Bari Aldo Moro, roberta.misuraca@uniba.it

Anna PATERNO, Department of Political Science, University of Bari Aldo Moro, anna.paterno@uniba.it

LIVING CONDITIONS, LIFESTYLES AND SELF-RATED HEALTH IN ITALY[1]

Clodia Delle Fratte, Stefano Gerosa, Francesca Lariccia, Daniela Lo Castro

Abstract. We study the association between living conditions, lifestyles and self-rated health in Italy, using the module on lifestyles and health difficulties included in the 2022 edition of the European Survey on Income and Living Conditions (EU-SILC). In particular, we use logistic regression to estimate the association between three EU-SILC indicators of poverty or social exclusion (risk of poverty, material deprivation and severe housing deprivation), covering different aspects of socio-economic disadvantage, and self-rated health, controlling for a full set of individual covariates and for many lifestyle risk factors (body mass index, smoking, alcohol consumption, healthy diet, physical activity and social participation). We document the existence of a strong association between socio-economic deprivation and the risk of poor self-rated health. Moreover, we find a significant impact of lifestyles on this association, showing how observed inequalities in subjective health depend on the complex interactions between socio-economic deprivation and the adoption of healthy habits.

1. Introduction

Low socio-economic status (SES), as measured by income, education, occupational condition or a combination of such factors, is consistently found to be associated with higher morbidity, mortality and with a number of other adverse health outcomes, even in advanced countries with generous welfare programs and large public health infrastructure (Stringhini *et al.* (2017) and Vineis *et al.* (2020) for a general discussion; Chetty *et al.* (2016) for USA; Kinge *et al.* (2019) for Norway; Petrelli *et al.* (2024) for Italy). On the other hand, unhealthy lifestyles are found to be among the main risk factors for non-communicable diseases and for a large share of premature mortality and total years of life lost and disability-adjusted life years at the global level (GBD 2021 Causes of Death Collaborators (2024), GBD 2021 Risk Factors Collaborators (2024)). However, the interaction between socio-economic deprivation and lifestyles choice is still poorly understood (Foster *et al.* (2021)).

Lifestyles are often considered as mediators between SES and health outcomes, even if some recent studies found that unhealthy habits mediate only a small fraction

[1] The article exclusively expresses the authors' opinions. Although the paper is the result of joint work, sections are attributed as follows: paragraphs 1 and 5 to S. Gerosa, paragraph 2 to F. Lariccia, paragraph 3 to D. Lo Castro, paragraph 4 to C. Delle Fratte.

of the observed association between socio-economic deprivation and health (Zhang *et al.* (2021), Fang *et al.* (2023)). Moreover, not only unhealthy lifestyles are more prevalent among low SES populations, but also there is some evidence that the adverse health effect of such behaviours is amplified among socio-economic deprived individuals, coherently with the "differential vulnerability" hypothesis (Foster *et al.* (2018)). Finally, risk behaviours appear to cluster and interact among specific subpopulations (Meader *et al.* (2017)), suggesting the need for a careful evaluation of the impact not only of single risk factors, but also of extensive combination of lifestyles, in order to uncover what has been called the "behavioural constellation of deprivation" (Pepper and Nettle (2017)).

Our aim is to examine the joint relations between socio-economic deprivation, lifestyles and self-rated health in Italy, using cross-sectional sample data from the 2022 edition of European Survey on Income and Living Conditions (EU-SILC). Our chosen health outcome variable, self-rated health (SRH), despite being a subjective indicator, has been repeatedly shown to be a robust and independent predictor of objective health outcomes, and in particular of mortality (Idler and Benyamini (1997), Jylhä (2009)).

We assess socio-economic deprivation (SED) using three EU-SILC indicators (risk of poverty, material deprivation and severe housing deprivation), covering different and only partially overlapping aspects of the social disadvantage spectrum, and we build an extensive healthy lifestyle score index, combining information on multiple individual behaviours (body mass index, smoking, alcohol consumption, healthy diet, physical activity and social participation). We then evaluate the association between the risk of poor SRH, SED and lifestyles first in the whole sample, and then performing stratified and joint analyses among the subpopulations identified by our main SES and behavioural indicators, to evaluate the combined impact of living conditions and lifestyles.

2. Data and variables

We use data from the Istat EU-SILC (European Survey on Income and Living Conditions) conducted in 2022. The survey's priority is to produce annual comparable data, both cross-sectional and longitudinal, using harmonized definitions and methods, for the analysis of income distribution, social exclusion, living conditions and quality of life of households, as well as of economic and social policies adopted at the national and European levels (Eurostat (2022)).

The survey is conducted on a nationally representative sample of Italian population. The reference population includes households and their current members residing in the territory of the country at the time of data collection. All household members are surveyed, but only those aged 16 and more are interviewed.

Information on social exclusion and housing conditions is collected mainly at household level, while labour, education and health information is obtained for persons aged 16 and over (Istat (2021)). According to the Commission Regulation, the 2022 edition includes a rolling module on health and lifestyles[2], considered in this study.

We assess socio-economic deprivation (SED) using three EU-SILC indicators of poverty or social exclusion, considering both monetary and non-monetary aspects. In particular, people at risk of poverty are those living in households with a net equivalent income below the at-risk-of-poverty threshold[3]. An individual is in condition of material deprivation if he lives in a household experiencing at least three out of nine symptoms of deprivation[4]. Finally, severe housing deprivation is defined as the condition of those living in an overcrowded dwelling and exhibiting at least one of the housing deprivation measures[5].

In order to evaluate the associations of lifestyles with SRH we consider many lifestyle risk factors: body mass index, smoking, alcohol consumption, healthy diet, physical activity and social participation. We use body mass index (BMI) to measure healthy weight. According to the World Health Organization (WHO) definition, we calculate BMI as weight (kg) divided by height (m) squared. We then classify individuals as underweight (BMI<18.5), normal weight (18.5<=BMI< 25), overweight (25<= BMI< 30), obese BMI>=30). We define normal weight as the healthy category.

We evaluate smoking consumption distinguishing between never smoking in the last 12 months (which is considered to be a healthy lifestyle), and smoking at least once in the last 12 months. We are not able to distinguish former smokers using our survey questions. We estimate each individual's alcohol consumption distinguishing between drinking not daily in the last 12 months (i.e. healthy behaviour), and drinking daily in the last 12 months. We do not possess any information on the volume and frequency of alcoholic beverages drank.

---

[2] 2022 module on health: Commission Regulation (EU) N° 2020/1721 of 17 November 2020.

[3] At-risk-of-poverty rate: percentage of people living in households with a net equivalent income below an at-risk-of-poverty threshold, set at 60% of the median of the individual distribution of net equivalent income. The net income considered for this indicator complies with the European definition. The income reference year is the calendar year preceding the survey.

[4] Material deprivation rate: percentage of the population that cannot afford at least three of the following nine items: 1) to pay their rent, mortgage or utility bills; 2) to keep their home adequately warm; 3) to face unexpected expenses; 4) to eat meat or proteins regularly; 5) to go on holiday; 6) a television set; 7) a washing machine; 8) a car; 9) a telephone.

[5] Severe housing deprivation rate: percentage of population living in a dwelling which is considered as overcrowded, while also exhibiting at least one of the housing deprivation measures. Housing deprivation is calculated by reference to households with a leaking roof, neither a bath, nor a shower, nor an indoor flushing toilet, or a dwelling considered too dark.

We identify healthy diet using two separate questions on the frequency of fruit consumption and the frequency of eating vegetables or salad during a typical week in a given season. A dummy variable is then obtained, taking the value of one if both are consumed at least once a day and zero otherwise. We define regular physical activity as engaging at least 4 times a week in sports, fitness or recreational activities that cause at least a small increase in breathing or heart rate for a continuous period of at least 10 minutes.

Finally, we create a social participation score as a combination of four questions about the frequency of getting together or of contacts with family (relatives) or friends. From its distribution, we identify those respondents who exceed the first quartile as socially participants, obtaining a dummy variable that takes the value of one if there is social participation (i.e. active social participation), zero otherwise (i.e. low social participation)[6].

We include in this study also the following other covariates: sex (male, female); age in 10-year classes (16-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75+); citizenship (Italian, foreign); education level (lower secondary education or below, high school, university degree and above); employment status (employed, unemployed, inactive); region of residence (North, Centre, South/Islands). Moreover, we consider the covariate chronic diseases (absence, presence) by considering the individuals interviewed self-reported information about chronic diseases or long-standing (lasted or expected to last at least 6 months) health problems.

Our chosen health outcome variable is self-rated health (SRH). According to WHO's definition of health, in the EU-SILC the variable is collected by using the internationally standardized question "How is your health in general?" and asking respondents to assess their health status on a 5-point scale from very good to very bad. The dependent variable has been obtained dichotomizing respondents' assessment of their health status: the dummy variable is equal to one in the case of a respondent answering bad or very bad (i.e. poor SRH), zero otherwise.

3. Methods

Our statistical analysis is based on a sample of 38,808 individuals aged 16 years and above, obtained from an initial sample of 39,814 individuals, after excluding those refusing to respond to health related questions (i.e. SRH and presence of chronic diseases).

We start by evaluating the association between the risk of poor SRH, SED and lifestyles in the whole sample. Because of the outcome variable is binary, two

[6] Following the remark of a referee, we tested alternative cut-off values (10^{th} percentile and the median) for the dichotomization of the social participation variable: results are qualitatively robust to the choice of different thresholds.

logistic regression models[7] have been estimated in order to produce adjusted Odds Ratios (ORs[8]) relating living conditions and healthy lifestyle factors to SRH. Model 1 investigates the association of single indicators of SED with the response variable, adjusting for some control covariates. Model 2 additionally includes lifestyle items to correct for the effects of lifestyle.

Subsequently, composite score variables have been constructed from the single items of living conditions and lifestyles, so that stratified and joint analyses have been performed in order to examine how these scores modify their relations with SRH. The presence of SED is obtained as a sum of the single items of living conditions (risk of poverty, material deprivation and severe housing deprivation) and is grouped into two modalities: having at least one socio-economic distress signal, having none. To comprehensively assess the associations of multiple lifestyles with SRH, we build an extensive healthy lifestyle score (LS) index, combining information on multiple individual behaviours considered in the previous model. In particular we proceed by summing healthy lifestyles and we define three groups using the so calculated score: low LS (having up to 2 healthy lifestyles), medium (having 3 or 4), high (having 5 or 6).

We finally conduct stratified and joint analyses, to evaluate how lifestyles affect the risk of poor SRH across different socio-economic groups and how the same risk varies across subgroups defined by both lifestyles and SED signals. Stratified models are separately estimated on the at-risk-of-SED and not at-risk-of-SED subsamples (10,031 and 28,049 individuals respectively), with low LS as the reference category. The reference category for the joint model is the most disadvantaged group, i.e. participants showing both SED signals and low LS.

4. Results

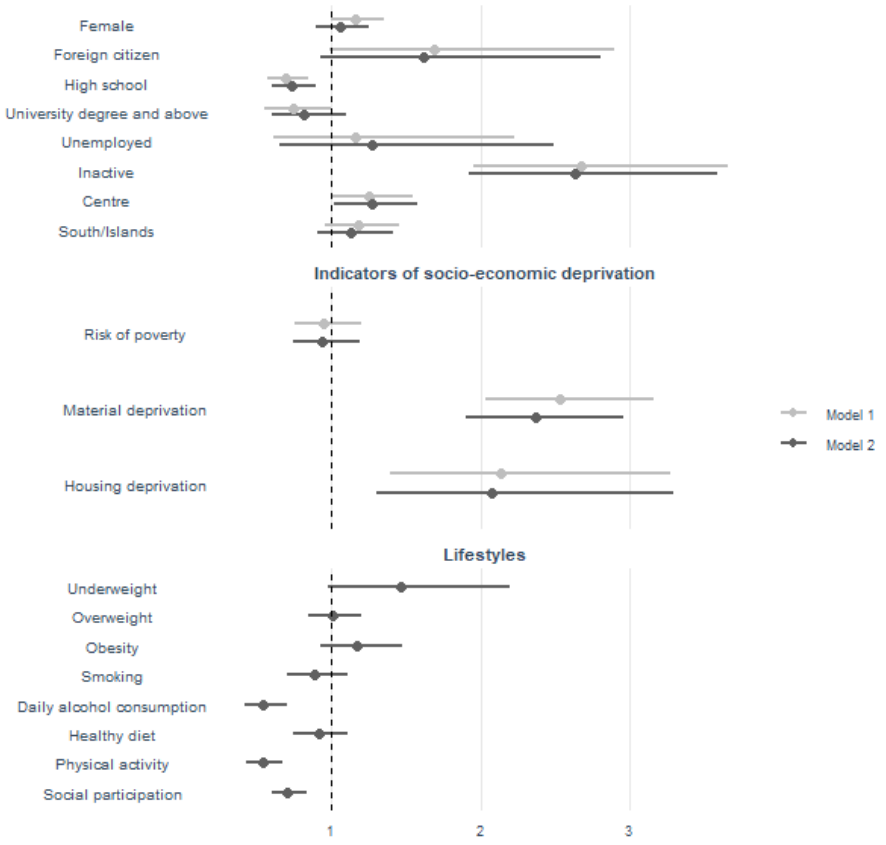
Figure 1 presents the results from the two logistic regression models. Controlling for age and chronic diseases, but not for lifestyles (Model 1), the characteristics associated with a higher risk of poor SRH at the 5% significance level are being inactive (OR=2.67, 95% CI=1.96-3.66; ref.: being employed) and living in the Centre (OR=1.25, 95% CI=1.01-1.54; ref.: North), while higher educational attainment is associated with a lower risk both for high school (OR=0.69, 95% CI=0.57-0.84; ref.: no high school degree) and for university degrees (OR=0.74,

[7] The package *survey* of the R software has been used to fit survey-weighted generalised linear models to data from a complex survey design.

[8] If the OR is equal to one, this means that there is no association between the exposure and the odds of a negative health perception (the odds of SRH given exposure and the odds of SRH given no exposure are the same). Values greater than one indicate positive association (the exposure is a risk factor), lower a negative one (the exposure is a protective factor).

95% CI=0.55-0.99; ref.: no high school degree). Being female (OR=1.16, 95% CI=0.99-1.35) and foreign citizenship (OR=1.69, 95% CI=0.98-2.89) are significant at the 10% significance level.

Figure 1 – *Logistic regression estimates of ORs and 95% confidence intervals without (Model 1) and with (Model 2) lifestyles.*



The dots provide a visual estimate of ORs, the lines represent the 95% confidence intervals. ORs with a confidence interval including 1 are not statistically significant at the 95% level. Age classes and chronic diseases have not been reported in this figure because of their high OR values.

Looking at indicators of SED, a remarkable difference appears between monetary and non-monetary ones: while individuals at relative risk of poverty do not show higher odds of poor SRH (OR=0.95, 95% CI=0.75-1.19), participants in condition

of material deprivation (OR=2.53, 95% CI=2.03-3.16) or experiencing severe housing deprivation (OR=2.14, 95% CI=1.39-3.27) display odds of poor SRH more than twice as high as those not presenting these SED signals.

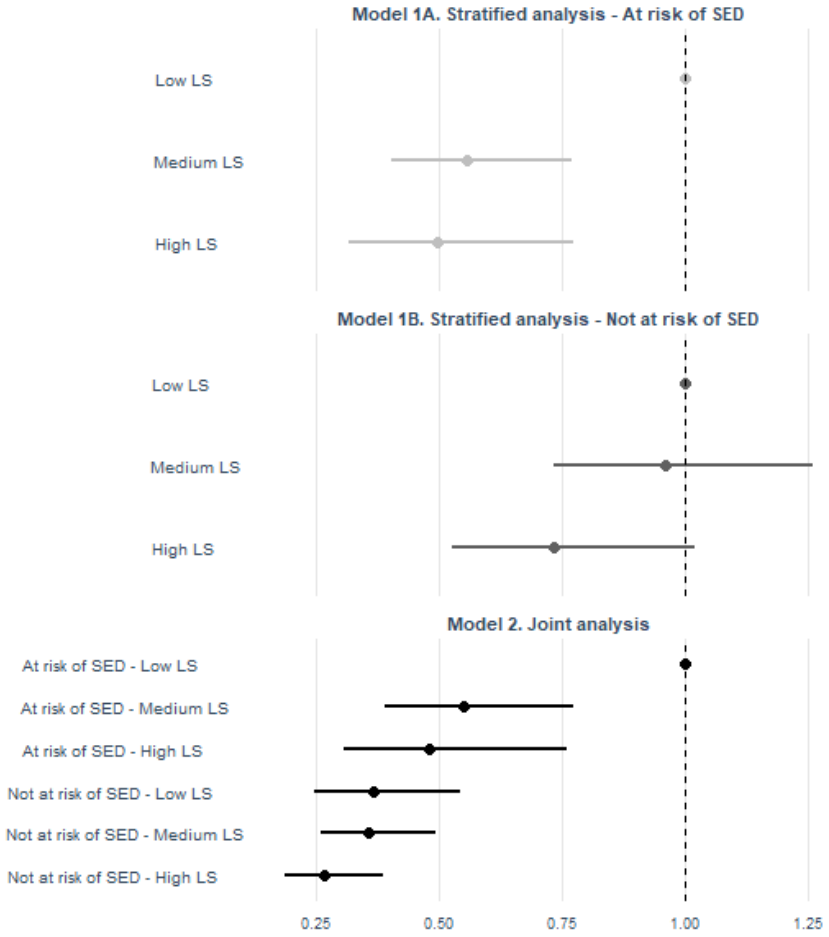
Adding lifestyle variables (Model 2) generally reduces the magnitude of the above-described associations and in some cases makes them lose statistical significance (e.g. female sex, foreign citizenship, university degree) while preserving the direction of the effects. Considering the impact of lifestyle variables, daily alcohol consumption (OR=0.54, 95% CI=0.41-0.70), regular physical activity (OR=0.54, 95% CI=0.43-0.67) and active social participation (OR=0.70, 95% CI=0.60-0.83) are all associated with lower odds of poor SRH. Being underweight is associated with a higher probability of being in poor SRH at the 10% significance level (OR=1.47, 95% CI=0.98-2.19; ref.: normal weight). Neither current smoking nor healthy diet habits seem to have significant effects on SRH[9].

Figure 2 and Tables 1 and 2 present the results from stratified and joint analyses. Stratified analysis (Table 1 and Figure 2 - Models 1A and 1B) show that healthier lifestyles are associated with lower odds of poor SRH both for the subgroup at risk of SED and for that not at risk, but the strength of the effect is much larger for the former. Moving from the low LS to the high LS category is associated with a 50% reduction in the odds of poor SRH for the socio-economically deprived population (OR=0.49, 95% CI=0.32-0.77), while the same improvement in the number of adopted healthy behaviours is linked with a 27% reduction for the not at-risk-of-SED subgroup (OR=0.73, 95% CI=0.53-1.02). These results support the hypothesis of differential vulnerability, for which SED amplifies the impact of unhealthy lifestyle behaviours.

Finally, the joint analysis (Table 2 and Figure 2 - Model 2) shows the combined effect of SED and lifestyles on the risk of poor SRH, using the most disadvantaged group as the reference point. The overall extent of observed inequalities in SRH is large, with the group with no SED signals and a high LS reducing by almost three-quarters the odds of poor SRH (OR=0.27, 95% CI=0.19-0.39) compared to the at-risk-of-SED and low LS group. Moreover, even the group with the most healthy behaviours among the at-risk-of-SED population has a smaller relative reduction in the risk of poor SRH (OR=0.48, 95% CI=0.31-0.76) than the group with the lowest LS but with no SED signals (OR=0.37, 95% CI=0.25-0.55), showing that socio-economic conditions play a bigger role than lifestyles in determining the risk of poor SRH.

[9] The unexpected results related to drinking and smoking behaviours are probably linked to our survey questionnaire, that does not allow us to distinguish former smokers and excessive drinking, as we discuss in Sections 2 and 5.

Figure 2 – *Logistic regression estimates of ORs and 95% confidence intervals of stratified (Models 1A and 1B) and joint (Model 2) analyses.*



Logistic models adjusted for sex, age, citizenship, education level, employment status, region of residence, chronic diseases.

Table 1 – *Association of lifestyle score with self-rated health stratified by presence of socio-economic deprivation (ORs, 95% confidence intervals, sample observations).*

Socio-economic deprivation	**Lifestyle score**		
	Low (0-2)	*Medium (3-4)*	*High (5-6)*
At risk of SED	1 (ref.) n=1,553	0.56 (0.40-0.77) n=5,954	0.49 (0.32-0.77) n=2,524
Not at risk of SED	1 (ref.) n=2,733	0.96 (0.73-1.26) n=15,884	0.73 (0.53-1.02) n=9,432

Table 2 – *Joint association of socio-economic deprivation and lifestyle score with self-rated health (ORs, 95% confidence intervals, sample observations).*

Socio-economic deprivation	**Lifestyle score**		
	Low (0-2)	*Medium (3-4)*	*High (5-6)*
At risk of SED	1 (ref.) n=1,553	0.55 (0.39-0.77) n=5,954	0.48 (0.30-0.76) n=2,524
Not at risk of SED	0.37 (0.25-0.54) n=2,733	0.36 (0.26-0.49) n=15,884	0.27 (0.19-0.39) n=9,432

5. Conclusions

We study the association between living conditions, lifestyles and the risk of self-rated poor health using EU-SILC, a cross-sectional nationally representative sample of the Italian population in 2022. Controlling for a large number of socio-demographic characteristics, we find that being at risk of poverty is not significantly associated with an increased risk of poor SRH, while both material deprivation and severe housing deprivation are associated with a more than doubled risk. Therefore, it seems that it is not low income *per se*, but weak access to material possibilities (presence of arrears on due household payments, incapacity to face unexpected expenses or to pay for holidays) and the distress due to overcrowded and low-quality housing that are associated with poor subjective health. Including lifestyles in our model only slightly reduces the magnitude of this association.

Looking at lifestyles, daily alcohol consumption, physical activity and social participation are found to be associated with significant reduction of the risk of being in poor SRH, while being underweight is weakly associated with an increased risk.

Our stratified and joint analyses uncover a complex picture of the risk of poor SRH across subgroups identified by the socio-economic deprivation index and the lifestyle score built aggregating healthy behaviours. Healthier lifestyles are associated with a reduced risk of poor SRH both within the socio-economic deprived and not socio-economic deprived subgroup, but the risk reduction associated with moving from the least healthy to the most healthy lifestyle group is much larger for the at-risk-of-SED population, pointing to the possibility of "differential vulnerability" of unhealthy behaviours across different populations. The joint analysis shows how lifestyles and socio-economic deprivation mutually reinforce each other in relation to the observed risk of poor subjective health. While the non-deprived population has a lower risk of poor SRH relative to the deprived population within each lifestyle score group, the combined effect of behaviours and socio-economic deprivation is very large: the subgroup of the non-deprived with at least 5 healthy lifestyles has a 73% lower risk of being in poor health relative to the subgroup with at least one signal of SED and less than 3 healthy behaviours.

These findings have direct policy implications, suggesting that a public health policy aimed at reducing the social and economic costs associated with poor health should concentrate on specific material aspects such as improving housing access and quality, and in identifying and supporting households facing multiple material risks related to debt, unpredictable economic shocks, low access to services.

Moreover, public health campaigns aiming at promoting healthy lifestyles should be specifically addressed to the most deprived segment of the population, because both risks and gains from behavioural changes are larger for this group. Our results show the urgent need for an explicit coordination of economic and health policies, with the objective of designing integrated interventions for the socially most vulnerable part of the population.

We acknowledge some limitations in our analysis. Being a cross-sectional study, we cannot claim to have identified any causal effect of socio-economic conditions and of personal behaviours on subjective health: there can be unobserved characteristics linked to the co-occurrence of low socio-economic and poor health conditions. Our chosen health outcome is self-reported, and not objective: as with any self-assessment, this can lead to measurement errors influenced by unobserved factors. The same caveat applies for our measurement of lifestyles: many of them are roughly measured, and in particular we do not have quantitative assessment of smoking and drinking behaviour and we do not possess information on objective health conditions.

Our study at the same time has multiple strengths and we consider it as a valuable reference point for further investigations. It is based on a nationally representative sample, and is part of a coordinated European survey using harmonized techniques, definitions and socio-economic indicators, making it possible to envisage

comparisons of the relationship between living conditions, lifestyles and health across European countries. It is a longitudinal survey, allowing for a future extension of our analysis considering changes over time in socio-economic conditions, lifestyles and health outcomes.

References

CHETTY R., STEPNER M., ABRAHAM S., LIN, S., SCUDERI, B., TURNER, N., BERGERON, A., CUTLER, D. 2016. The Association between Income and Life Expectancy in the United States, 2001–2014, *Journal of the American Medical Association*, Vol. 315, No.16, pp. 1750–1766.

EUROSTAT. 2022. *Methodological guidelines and description of EU-SILC target variables, 2022 Operation.* Luxembourg: Eurostat.

FANG W., CAO Y., CHEN Y., ZHANG H., NI R., HU W., PAN G. 2023. Associations of family income and healthy lifestyle with all-cause mortality, *Journal of Global Health*, Vol. 13:04150.

FOSTER H.M.E., CELIS-MORALES C., NICHOLL B.I., PETERMANN-ROCHA F., PELL J.P., GILL J., O'DONNELL C.A., MAIR F.S. 2018. The effect of socioeconomic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a prospective analysis of the UK Biobank cohort, *Lancet Public Health*, Vol. 3, pp. 576–585.

FOSTER H.M.E., POLZ P., MAIR F., GILL J., O'DONNELL C.A. 2021. Understanding the influence of socioeconomic status on the association between combinations of lifestyle factors and adverse health outcomes: a systematic review protocol, *BMJ Open*, Vol.11, No.5: e042212.

GBD 2021 CAUSES OF DEATH COLLABORATORS 2024. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021, *Lancet*, Vol. 403, No.10440, pp. 2100–2132.

GBD 2021 RISK FACTORS COLLABORATORS 2024. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021, *Lancet*, Vol. 403, No.10440, pp. 2162–2203.

IDLER E. L. AND BENYAMINI Y. 1997. Self-rated health and mortality: a review of twenty-seven community studies, *Journal of Health and Social Behavior*, Vol. 38, No.1, pp. 21–37.

ISTAT 2021. *L'indagine EU-SILC: innovazioni nella metodologia di rilevazione e stima.* Roma: Istat. https://www.istat.it/it/files//2021/04/Indagine-EU-SILC_Metodi_Ebook.pdf.

JYLHÄ M. 2009. What is self-rated health and why does it predict mortality? Towards a unified conceptual model, *Social Science & Medicine*, Vol. 69, No.3, pp. 307-316.

KINGE J.M., MODALSLI J.H., ØVERLAND S., GJESSING H., TOLLÅNES M., KNUDSEN A., SKIRBEKK V., STRAND B., HÅBERG S., VOLLSET S. 2019. Association of Household Income with Life Expectancy and Cause-Specific Mortality in Norway, 2005-2015, *JAMA*, Vol. 321, No.19, pp. 1916-1925.

MEADER N., KING K., MOE-BYRNE T., WRIGHT K., GRAHAM H., PETTICREW M., POWER C., WHITE M., SOWDEN A. 2016. A systematic review on the clustering and co-occurrence of multiple risk behaviours, *BMC Public Health*, Vol. 16:657.

PETRELLI A., VENTURA M., DI NAPOLI A., PAPPAGALLO M., SIMEONI S., FROVA L. 2024. Socioeconomic Inequalities in Avoidable Mortality in Italy: Results from a Nationwide Longitudinal Cohort, *BMC Public Health*, Vol. 24:757.

STRINGHINI S., CARMELI C., JOKELA M., AVENDAÑO M., MUENNIG P., GUIDA F. 2017. Socioeconomic Status and the 25 x 25 Risk Factors as Determinants of Premature Mortality: a Multicohort Study of 1.7 Million Men and Women, *Lancet*, Vol. 389, No.10075, pp. 1229–1237.

ZHANG Y.B., CHEN C., PAN X., GUO J., LI Y., FRANCO O., LIU G., PAN A. 2021. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies, *BMJ*, Vol. 373:n604.

Clodia DELLE FRATTE, Istat, dellefra@istat.it
Stefano GEROSA, Istat, gerosa@istat.it
Francesca LARICCIA, Istat, lariccia@istat.it
Daniela LO CASTRO, Istat, locastro@istat.it

ANALYZING HEALTH INEQUALITIES IN ITALIAN METROPOLITAN CITIES: A MULTIDIMENSIONAL APPROACH[1]

Daniela Vacca, Lucia Martina, Debora Tronu

Abstract. This study aims to verify whether, as with other previously investigated demographic and socio-economic phenomena (Cangialosi *et al.*, 2023; Istat, 2023), in Italian metropolitan cities, the relationships between the capital city and other surrounding territorial aggregates, divided based on geographical proximity, can explain health inequalities among the population. Health, defined as a state of complete physical, mental, and social well-being (WHO, 1948), is influenced by various determinants, including individual, social, economic, and environmental factors. It is plausible to hypothesize the existence of a territorial health gradient among municipalities with high availability of health infrastructures, low out-of-region hospital mobility, and favorable socio-economic conditions, compared to those with less accessibility to such services and less advantageous socio-economic conditions. Capitals (and sometimes municipalities in the first urban belt) often exhibit these advantageous conditions and play a central role in the organization of the territory (Logan *et al.*, 2002). Furthermore, the geographical positioning of territories is a significant factor in the proposed analysis, traditionally highlighting an increasing North-South discrepancy (Istat, 2024). A cluster analysis carried out on representative indicators of health conditions, health supply and economic situation facilitated the identification of homogeneous territorial groups and the description of their characteristics.

1. The concept of health and its determinants

The concept of health is fundamental to human life. The World Health Organisation (WHO, 1948) defines it as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity". This approach, which is multidimensional in nature, places the individual at the centre of analysis. Furthermore, it emphasises the interrelated importance of three inseparable components, namely physical, mental and social health (Engel, 1977; Marmot *et al*., 2006). Each of these components is influenced by health determinants, which are factors that can influence the onset or progression of disease without being its direct cause (Dahlgren *et al*., 1991). An analysis of these determinants allows the

[1] The paper has been jointly written by all the authors but § 1 and 2 can be attributed to D. Vacca, § 3 to L. Martina, § 4 to D. Tronu, § 5 all the authors. Indicators have been calculated by A. Cabras, E. Olla, G.Pala.

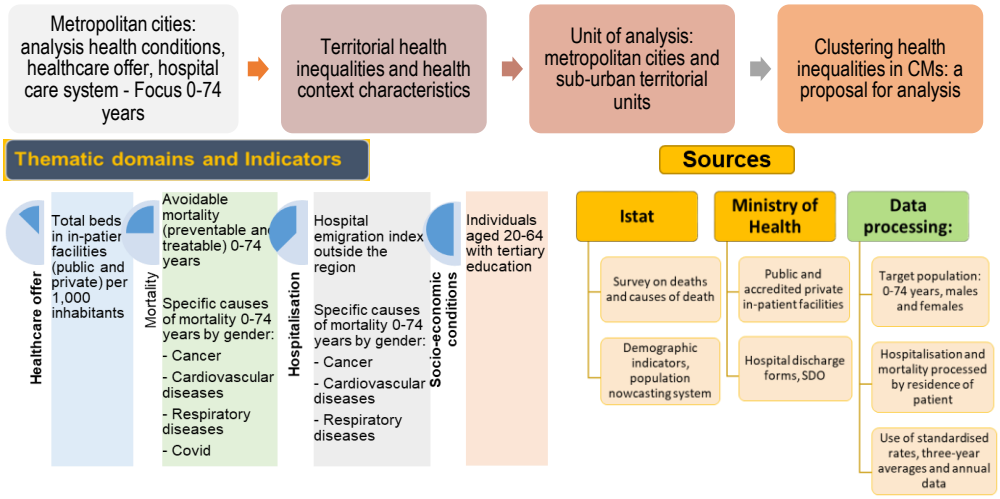
identification of the 'rules' by which disease occurs and develops, as well as the assessment of their presence, strength and direction. This enables the guidance of potential public health interventions. These determinants include a wide range of factors such as individual behaviours, social advantages or disadvantages, living and working conditions, access to health services, and socio-economic, cultural, environmental, and genetic influences (Commission on Social Determinants of Health, 2008). In recent decades, public health scholars have focused on social determinants of health (SDH) factors directly influenced by social policies and the socio-economic, cultural, and environmental conditions in which people live and work. Urban planning and regeneration programs, due to their direct impact on living environments, hold significant potential for improving public health outcomes (Kawachi *et al*., 2003). Strategies promoting inclusivity and community well-being are crucial for enhancing quality of life and reducing health disparities (WHO, 2016). Examples of such strategies include improving walkability, providing recreational areas, and ensuring access to nutritious foods, which collectively shape the quality of living environments (Barton *et al*., 2006). Cities, at both national and international levels, are pivotal in the implementation of health policies and programs that are crucial for achieving global public health objectives. The WHO identifies urbanization as a critical challenge for 21st-century health, with projections indicating that by 2050, two-thirds of the global population will reside in urban areas (WHO, 2016). Urban environments present a complex interplay of challenges and opportunities: while high population density can exacerbate issues such as pollution and disparities in access to services (Diez Roux *et al*., 2010), it also provides opportunities for targeted interventions and the optimization of health resources (Barton *et al*., 2006). Consequently, cities can serve as experimental settings for innovative strategies aimed at improving access to health services and fostering healthier living conditions by leveraging geographic proximity and urban infrastructure connectivity (Vlahov *et al*., 2007; Braveman *et al*., 2014). Public health, long a priority in policy-making, is now more than ever tasked with ensuring equitable and distributed levels of care and support, positioning it as a fundamental and integrative sector upon which all other decisions depend (Istat, 2022).

2. Data and Methods

This study analyses health conditions and healthcare provision in Italian urban contexts to identify the strengths and weaknesses of the areas examined. The 14 Italian metropolitan cities are compared according to a scheme that divides each area into four territorial typologies, defined in accordance with criteria of geographical contiguity. These types include capital cities, first belt municipalities, second belt municipalities and other municipalities. In total, 56 metropolitan areas were identified, comprising 1,268 municipalities (16% of the total municipalities) and

around 21 million inhabitants. The objective of the research is to ascertain whether the relationships between the Capital Municipalities and the surrounding areas can be used to explain the differences in the population's health status, in a manner analogous to the explanation of other demographic and socio-economic phenomena (Cangialosi et al., 2023; Istat, 2023). The study puts forward the hypothesis of a territorial health inequalities between municipalities with a high availability of health infrastructure and favourable socio-economic conditions and those with less access to these services and less favourable socio-economic conditions (Istat, 2024). In general, the capitals of metropolitan areas, which are central to spatial organisation (Rothwell et al., 2015), and municipalities in the first urban belt with greater connectivity to the central city, tend to have more favourable conditions.

Figure 1 – *Framework: Indicators and data sources.*



The geographic location of metropolitan cities is also regarded as a significant factor in the exacerbation of the health disparity between different regions of the country, particularly in its contribution to the North-South divide (Istat, 2019).

The testing of these assumptions is aligned with the modern multidimensional approach to studying health status, necessitating the application of a holistic and complex framework that considers all relevant dimensions of the phenomenon and its interactions. The adopted conceptual framework (Figure 1) is based on identifying a plurality of thematic domains within which to select the most pertinent and explanatory indicators (Istat, 2024), balancing data availability with the descriptive capacity of the selected measures. In particular, the health status of the population is

assessed using mortality indicators, disaggregated by cause and age. Avoidable mortality indicators are used, segmented into preventable and treatable components.

The number of hospital beds provides insight into the accessibility of health services. The rates of hospitalisation by cause and age, along with hospital mobility data, indicate the prevalence of specific diseases and the utilisation of health services. The level of education provides insight into the socio-economic context of the areas, indicating potential relations between education and health status or access to health service. Mortality and hospitalisation rates are age-standardised with the European standard population 2013. The key data for this analysis derive from multiple sources: infrastructural capacity and hospitalization data from the Ministry of Health, while death and cause-of-death data are provided by Istat. The analysis is conducted based on patient residence (hospital discharges) and deceased residence (deaths and causes of death). The indicators predominantly pertain to the population aged 0-74 and refer to the year 2021. For preventable mortality, the three-year average of 2017-2019 was considered in order to understand the effects of the Coronavirus Pandemic Disease 2019 (Covid-19) on the health and care landscape within the specified territory. Indeed, the experiences of the global population following the COVID-19 pandemic have expanded the audience of policymakers interested in quantitative information and brought attention from the general public (government bodies, private citizens, media organizations, etc.) to issues related to the understanding and measurement of public health. In order to capture the complexity of the analysis and to represent the thematic dimensions considered (health conditions, health services and socio-economic situation), a substantial number of indicators were used.

Cluster analysis was applied to these data, which enabled the identification of homogeneous groups and highlighted patterns and inequalities in health conditions and access to health services (Murray et al., 2020; Williams et al., 2018).

3. Cluster analysis

A cluster analysis was conducted using the k-means algorithm on twelve indicators representative of the different domains of interest (MacQueen, 1967).

The k-means algorithm, is a non-hierarchical clustering tool that is straightforward to operate. Its objective is to partition a dataset into k clusters based to their similarity. The goal of the algorithm is to minimize the total intra-cluster variance and maximize the variance between clusters[2]. The optimal number of

[2] The algorithm begins by randomly defining k centroids and assigning the closest samples to the j-th centroid, utilizing a Euclidean distance metric (commonly the square of the difference). Once the initial assignment of points to clusters has been made, the algorithm recalculates the centroid of each cluster as the average of the points assigned to that cluster. This process of assigning and updating the centroid

clusters into which to subdivide the dataset is not known a priori, but is an input necessary to start the algorithm. The clustering analysis is then performed for different values of k, starting with a minimum of 2, and the goodness of the clustering analysis is assessed a posteriori. In order to validate the results of clustering we considered the following measures: the Total Within Sum of Squared errors (SSW) and the Total Between Sum of Squared errors (SSB). Higher SSB values (and consequently lower SSW values, since SSW+SSB=TSS, constant) are more desirable, as they are indicative of greater internal group cohesion and data separation. These measures, however, although useful and quick to implement, are dependent on the number of clusters, and, therefore, are not suitable for comparing clustering results with a different number of clusters. In order to evaluate the clustering method and the choice of the k number the Silhouettes method was also used (Rousseeuw 1987). This method is based on the calculation of intra- and inter-cluster distances and provides a measure of how similar the data are to the assigned cluster in comparison to other clusters. It is done by calculating the silhouette value[3] for each data point and averaging the result over the entire dataset. The silhouette average can range between -1 and 1: a value close to 1 means the cluster is well formed, a value close to 0 means the position is unclear and a value close to – 1 means the dataset is badly partitioned.

Table 1 – *Values of Silhoutte, WSS, BSS e BSS/TSS (%) for different clustering with k varying from k=2 to k=6.*

k	Silhouette	WSS	BSS	TSS	BSS/TSS(%)
2	0.2821	455	205	660	31.1
3	0.2062	376	284	660	43.0
4	0.2204	324	336	660	50.9
5	0.2276	285	375	660	56.9
6	0.1920	257	403	660	61.1

To ascertain the optimal value of k, it is useful to examine the value of the Silhouette measure within the classes in the five- and six-group clusters.

Table 2 shows the values of the Silhouette measure for the five- and six-group clustering, for each k group. The highest Silhouette value was observed in the second group of the six-group cluster, with a value of 0.3753; the third group of this cluster

is repeated until it converges, i.e. until there is minimal change between iterations, or until the maximum number of iterations has been reached.

[3] The silhouette value for a single data point is calculated using the following equation:

Silhouette = (AverageOut – AverageIn)/ max (AverageOut, AverageIn)

AverageOut is the minimum average distance between the data point and data within other clusters, and AverageIn is the average distance between the data point and other data within the same cluster.

was also well represented, with a value of 0.3000. In contrast, the first class of the six-group cluster had the lowest silhouette value (0.0060), indicating a suboptimal positioning of some territories. Considering the different measures selected (Silhouette in total and in the groups, BBS/TSS%), the six-group cluster is assessed as optimal, having satisfactory values of the silhouette measures and BBS/TSS(%).

Table 2 – *Silhouette values for clustering with 5 and 6 groups for each group (k).*

k	5-groups		6-groups	
	n.	Silhouette	n.	Silhouette
1	23	0,2413	13	0,0060
2	10	0,2286	12	0,3753
3	10	0,1693	10	0,3000
4	9	0,2072	9	0,1420
5	4	0,3388	7	0,1360
6			5	0,1884
Total	**56**	**0,2276**	**56**	**0,1920**

4. Results

The cluster analysis, described in Section 3 and conducted on the twelve selected multi-source indicators, made it possible to identify homogeneous groups and to understand the key characteristics associated with them in terms of health status, health care provision and socio-economic conditions.

The results of the cluster analysis serve to substantiate the assumptions set forth in Section 2, which postulated a relation between enhanced health status and high accessibility to health services and superior socio-economic conditions. In the case of metropolitan cities, these conditions are predominantly concentrated in the Capital cities and the majority of territories situated in the Centre-North.

In order to interpret the results and describe the group profiles, the 'cluster centres' obtained from the cluster analysis performed were used (see Table 3). The analysis led to the classification of the 56 territories, comprising the capital cities, first- and second-belt municipalities, and other metropolitan city municipalities, into one of six clusters, as illustrated in the map (Figure 2).

Descriptive characteristics of each group are represented below.

The first group (CL1) includes the extra-urban territories mainly belonging to the Mezzogiorno, together with the municipalities bordering the capital cities of Genoa and Venice and the territories furthest away from them. They are characterised by acceptable health conditions of the population up to 74 years of age, but with higher levels of treatable mortality (6.9 deaths per 10 thousand inhabitants), which show

shortcomings attributable above all to the quality and effectiveness of health care, which is not always timely and effective (Table 3).

Table 3 – *Indicators - Cluster centers and values for capital city, MC and Italy.*

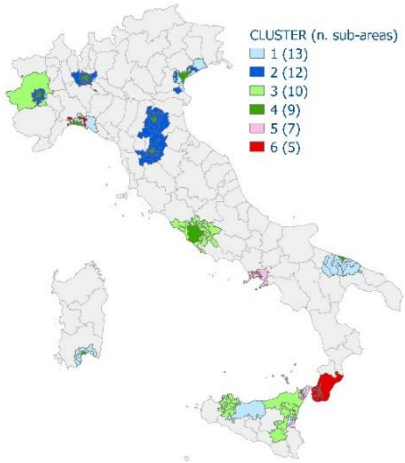
	Cluster centers						Capi-tal		
Indicators	CL1	CL2	CL3	CL4	CL5	CL6	city	MC	ITALY
Tot hospital beds 2021 (per 1,000 in.)	1.1	2.5	2.1	8.4	4.7	2.3	7.2	4.3	3.9
Hospital emigration index outside the region 2021 (%)	7.3	3.7	5.2	5.4	6.3	21.9	5.7	5.8	7.8
Hospitalizations from cancer 0-74 years 2021 (per 10,000 in.)	96.6	91.8	98.3	100.7	114.3	101.8	103.0	99.3	97.5
Hospitalizations from cardiovascular system disease 0-74 years 2021 (per 10,000 in.)	75.3	78.6	82.9	77.9	97.2	73.8	83.7	84.0	84.5
Hospitalizations respiratory disease 0-74 years 2021 (per 10,000 in.)	55.9	74.5	67.9	70.5	58.7	51.9	73.2	68.7	67.9
Preventable mortality 2017-19 (per 10,000 in.)	9.8	9.5	11.5	9.9	12.8	10.5	10.6	10.7	10.5
Preventable mortality ‘21 (per 10,000 in.)	12.7	11.6	14.6	12.2	17.1	12.4	13.6	13.6	12.8
Treatable mortality ‘21 (per 10,000 in.)	6.9	5.4	7.6	6.0	9.0	7.6	6.8	6.9	6.4
Mortality from cancer 0-74 years 2021 (per 10,000 in.)	12.1	10.9	12.3	11.2	14.3	11.9	12.0	12.1	11.7
Mortality from cardiovascular system disease 0-74 years 21 (per 10,000 in.)	5.0	4.0	6.3	4.8	7.8	6.1	5.6	5.5	5.2
Mortality from respiratory disease 0-74 years 2021 (per 10,000 in.)	0.9	0.9	1.5	1.0	1.7	1.6	1.3	1.2	1.1
Individuals aged 25-64 with tertiary education 2021 (%)	17.9	21.4	18.0	33.1	19.1	20.9	31.5	24.5	21.7

In contrast, primary prevention and public health interventions generate a good level of preventable mortality (12.7 deaths per 10 thousand inhabitants), although the pandemic has had a sustained impact on the health conditions of these areas (+2.8 percentage points in 2021 compared to 2017-2019). It includes 2.247 million people, 44.1% of whom belong to the first belt municipalities, 17.5% to the second belt municipalities and 38.3% residing in the ring areas (Table 4). The socio-economic context of the population is fragile with respect to both the level of tertiary education of the adult population (17.9%) and per capita income (12,819 euros per inhabitant), used as a further descriptive variable outside the cluster analysis.

The second cluster (CL2) includes the extra-urban territories of the metropolitan cities in the Centre-North, excluding Genoa. These territories have the best health conditions, with lower rates of all-cause and avoidable mortality, supported by an efficient and accessible healthcare system that does not rely exclusively on hospital

care. This is confirmed by the low hospitalisation rates for cancer and cardiovascular diseases, as well as the out-of-region hospitalisation rate (3.7%). Together with cluster 4, it has the highest rate of hospitalisation for respiratory diseases, at 74.5 admissions per 10 thousand inhabitants, likely due by the effects of the Covid 19 pandemic on hospital care.

Figure 2 – *Distribution of metropolitan city sub-areas among clusters.*



The socio-economic context is characterised by more favourable levels of tertiary education population ages 20-64 (21.4%), in line with the national average, and high levels of income (16,284 euros per inhabitant). About 3.970 million inhabitants reside in this group, of which 36% in the first belts, 30.6% in the second belts, and 33.4% in the territories further out from the capital.

The third cluster (CL3) comprises mainly extra-urban territories, with the exception of the municipality of Palermo, which cross Italy from North to South. It is characterised by poor health conditions, a higher recourse to hospital care for cancer and cardiovascular diseases, compared to the two previous clusters, a significant preventable mortality (14.6 per 10 thousand inhabitants), up by 3 percentage points compared to 2017-2019, accompanied by high values also of treatable and preventable mortality (7.6 deaths per 10 thousand inhabitants) and by cause. Socio-economic conditions are rather critical, with low levels of tertiary education of the adult population, similar to those in cluster 1 and associated with very low values of per capita income (10,356 euros per inhabitant). Population in this cluster is about 4 million inhabitants of which 15.7 % in the capital municipalities, 21.9 % in the first belts, 21.5 % in the second belts and 40.8 % in the Other Metropolitan City municipalities.

Table 4 – *Cluster territorial profile.*

Cluster description	Territories	N. of territories		Population (31/12/21)		Municipalities	
		a. v.	%	a. v.	%	a. v.	%
CL1: Suburban areas mainly belonging to the Mezzogiorno. To these are added I Belt and the Other Municipalities of the CMs of Genoa and Venice	Total (% on Total of groups)	**13**	**23.2**	**2,247,117**	**10.5**	**208**	**16.4**
	Capital cities of MC (% on Total of CL1)	0	0.0	0	0.0	0	0.0
	I Belt (% on Total of CL1)	**6**	46.2	991,457	44.1	60	28.8
	II Belt (% on Total of CL1)	**2**	15.4	393,933	17.5	22	10.6
	Other Mun. of MC (% on Total of CL1)	**5**	38.5	861,727	38.3	126	60.6
CL2: Suburban territories mainly in the Centre North	Total (% on Total of groups)	**12**	**21.4**	**3,970,645**	**18.6**	**280**	**22.1**
	Capital cities of CM (% on Total of CL2)	0	0.0	0	0.0	0	0.0
	I Belt (% on Total of CL2)	4	33.3	1,427,789	36.0	54	19.3
	II Belt (% on Total of CL2)	5	41.7	1,215,809	30.6	84	30.0
	Other Mun. of MC (% on Total of CL2)	3	25.0	1,327,047	33.4	142	50.7
CL3: Mainly extra-urban areas with the exception of Palermo that cross Italy from North to South	Total (% on Total of groups)	**10**	**17.9**	**4,037,722**	**18.9**	**557**	**43.9**
	Capital cities of CM (% on Total of CL3)	1	10.0	635,439	15.7	1	0.2
	I Belt (% on Total of CL3)	2	20.0	885,441	21.9	37	6.6
	II Belt (% on Total of CL3)	3	30.0	868,809	21.5	64	11.5
	Other Mun. of MC (% on Total of CL3)	4	40.0	1,648,033	40.8	455	81.7
CL4: Main municipalities with health advantage situations	Total (% on Total of groups)	**9**	**16.1**	**6,975,357**	**32.7**	**9**	**0.7**
	Capital cities of CM (% on Total of CL4)	9	100.0	6,975,357	100.0	9	100.0
	I Belt (% on Total of CL4)	0	0.0	0	0.0	0	0.0
	II Belt (% on Total of CL4)	0	0.0	0	0.0	0	0.0
	Other Mun.of MC (% on Total of CL4)	0	0.0	0	0.0	0	0.0
CL5: The South with critical health conditions	Total (% on Total of groups)	**7**	**12.5**	**3,541,784**	**16.6**	**104**	**8.2**
	Capital cities of CM (% on Total of CL5)	3	42.9	1,443,492	40.8	3	2.9
	I Belt (% on Total of CL5)	1	14.3	532,593	15.0	14	13.5
	II Belt (% on Total of CL5)	2	28.6	502,244	14.2	23	22.1
	Other Mun. Of MC (% on Total of CL5)	1	14.3	1,063,455	30.0	64	61.5
Cl6: High hospitalisation outside the region. Critical socio-economic context conditions. Reggio Calabria and Genoa II Belt	Total (% on Total of groups)	**5**	**8.9**	**568,349**	**2.7**	**110**	**8.7**
	Main Mun. (% on Total of CL6)	1	20.0	288,286	50.7	77	70.0
	I Belt (% on Total of CL6)	1	20.0	172,479	30.3	1	0.9
	II Belt (% on Total of CL6)	2	40.0	34,848	6.1	12	10.9
	Other Mun. Of MC (% on Total of CL6)	1	20.0	72,736	12.8	20	18.2
TOTAL		**56**	**100**	**21,340,974**	**100.0**	**1**	**100.0**

The fourth group (CL4) is that of large urban centres, as it includes only capital cities, 9 out of 14, and affects about 7 million inhabitants. This is the group with the best health offer and good health conditions. The rates of preventable (12.2 deaths per 10,000 inhabitants), treatable (6 deaths per 10,000 inhabitants) and all-cause mortality are low and only slightly lower than those of cluster 2. The hospital supply in the area of residence is the highest (8.4 beds per thousand inhabitants) and associated with a high use of hospital care for cancer and respiratory diseases. The socio-economic context is excellent. In fact, these territories have the highest level of tertiary education, with one third of people aged between 25 and 64, and the highest per capita income (17,500 euros per inhabitant).

The fifth cluster (CL5) consists of territories in the South with more critical health conditions. It involves the entire metropolitan city of Naples and extends to the pole municipalities of Catania and Messina, including the first belt of the latter. These territories are characterised by very high values of preventable (17.1 deaths per 10,000 inhabitants), treatable (9 deaths per 10,000 inhabitants) and all-cause

mortality. Recourse to hospital care also reaches maximum levels with reference to hospitalisation for cancer (114.3 admissions per 10 thousand) and for cardiovascular diseases (97.2 admissions per 10 thousand). The epidemiological profile of these territories highlights the need of these areas for urgent interventions of primary prevention, public health and improvement of the accessibility and efficiency of the healthcare systems, also confirmed by a high recourse to hospital care outside the region of residence (6.3%). The resident population in this group is about 3.5 million inhabitants of which 40.8 % in the capital municipalities, 15 % in the first belts, 14.2 % in the second belts and 30 % in the Other Metropolitan City municipalities.

In conclusion, the sixth cluster (CL6) includes all the territories of the Metropolitan City of Reggio Calabria and the second belt of Genoa and involves about 568,000 people, mainly residing in the city of Reggio di Calabria and its first belt (81.1%). The cluster is characterised by the population's strong propensity to use hospital care outside the region and a significant treatable mortality rate. Health conditions are poor with various health and health-related fragilities. There is a high hospitalisation for cancer (101.8 admissions per 10 thousand inhabitants), a high preventable mortality (12.4 deaths per 10 thousand) suggesting that public health and primary prevention interventions should be strengthened although the impact of Covid 19 was lower than in other groups (+1.9 percentage points compared to 2017-2019). The main criticality of the group is the high treatable mortality (7.6 deaths per 10 thousand inhabitants) and cardiovascular (6.1 deaths per 10 thousand) and respiratory (1.6 deaths per 10 thousand) diseases. The socio-economic conditions, as in group 5, are disadvantaged, characterised by low average levels of tertiary education of the adult population (20.9%) and low per capita income (10,900 euros per inhabitant).

5. Summary

This analysis has highlighted several significant trends regarding the distribution of healthcare services and health conditions in Italian metropolitan cities and their suburban areas. Hospital services are predominantly concentrated in major urban centers, while the propensity for hospital mobility increases as one moves away from the main cities. Some areas exhibit a very high rate of out-of-region hospitalizations. The analysis underscores the importance of developing targeted health policies to address territorial inequalities and improve access to healthcare services, especially in disadvantaged areas. Specific interventions need to be implemented to improve health conditions in the most critical regions, with a particular focus on Southern Italy. These interventions should aim to reduce avoidable mortality and enhance healthcare infrastructure to ensure equitable access to hospital care. By addressing the initial assumptions, the study confirms that the relationships between the capital city and other surrounding territorial aggregates, divided according to geographical

proximity, can indeed explain health inequalities among the population. The health differences identified suggest that municipalities with high availability of healthcare infrastructure and favourable socio-economic conditions exhibit better health outcomes compared to those with limited access to these services and less advantageous socio-economic conditions. The classic North-South divide also plays a significant role in influencing health conditions and outcomes, further emphasizing the need for region-specific health interventions.

References

BARTON H., GRANT M. 2006. A health map for the local human habitat. *Journal of the Royal Society for the Promotion of Health*, Vol. 126, No. 6, pp. 252-253.

BRAVEMAN, P., GOTTLIEB L., 2014. Defining equity in health. *Journal of Epidemiology and Community Health*, Vol. 58, No. 3, pp. 254-258. doi:10.1136/jech-2014-204288

CANGIALOSI D., TRONU D., VACCA D. 2023. The ageing of the population in urban contexts: a study on Italian metropolitan cities. *Rivista Italiana di Economia Demografia e Statistica,* Vol. 77, No. 3, pp. 22-32, July-Sept.

COMMISSION ON SOCIAL DETERMINANTS OF HEALTH. 2008. Closing the gap in a generation: health equity through action on the social determinants of health. *Final report of the Commission on Social Determinants of Health*. Geneva: World Health Organization.

DAHLGREN, G., WHITEHEAD, M. 1991. *Policies and strategies to promote social equity in health.* Stockholm: Institute for Futures Studies.

DIEZ ROUX, A. V., & MAIR, C. 2010. Neighborhoods and health. *Annals of the New York Academy of Sciences*, Vol. 1186, No. 1, pp. 125-145. doi:10.1111/j.1749-6632.2009.05333.x

ENGEL, G.L. 1977. The need for a new medical model: a challenge for biomedicine. *Science*, Vol. 196, No. 4286, pp. 129-136. doi:10.1126/science.847460

ISTAT. 2024. Condizioni di salute e offerta sanitaria nelle città metropolitane. *Statistica Focus*, *May 21 2024*

ISTAT. 2023. Gli anziani nelle città metropolitane. Profilo sociodemografico e analisi comparativa fra i contesti urbani. *Statistica Focus*, August 03 2023

ISTAT. 2022. *Attività antropiche e salute delle coste. Indicatori di rischio e sostenibilità per aree costiere e insulari in Italia e nell'Unione Europea*. Roma: Istituto nazionale di statistica. ISBN 978-88-458-2084-7

ISTAT. 2020. *Rapporto sul territorio 2020*. Collana: Letture statistiche. Roma: Istituto nazionale di statistica. ISBN 978-88-458-2014-4; doi.org/10.1481/Istat.Rapportoterritorio.2020

ISTAT. 2019. *La salute nelle regioni italiane. Bilancio di un decennio*. Roma: Istituto nazionale di statistica. ISBN 978-88-458-1970-4

KAWACHI, I., BERKMAN, L. F. 2003. Neighborhoods and health. *Oxford University Press*. ISBN 978-0-19-513838-2

LOGAN, J.R., ALBA, R.D., ZHANG, W. 2002. Immigrant enclaves and ethnic communities in New York and Los Angeles. *American Sociological Review*, Vol. 67, No. 2, pp. 299-322. http://dx.doi.org/10.2307/3088897

MACQUEEN, J. 1967. Some methods for classification and analysis of multivariate observations, in: Le Cam, L. and Neyman, J. *Proceedings 5th Berkeley Symposium on Mathematical Statistics and Probability. Vol. 5.1, pp.* 281-297. University of California Press, Berkeley

MARMOT, M., WILKINSON, R.G. 2006. *Social determinants of health*. Oxford: Oxford University Press. ISBN 978-0-19-856589-5

MURRAY, C.J., LOPEZ, A.D. 1996. *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020*. Cambridge, MA: Harvard School of Public Health on behalf of the World Health Organization and the World Bank. ISBN 978-0-674-3544

ROUSSEEUW, P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*. Vol. 20, pp.53-65.

ROTHWELL, J.T., OLFERT, M.R., PARTRIDGE, M.D. 2015. Settled to core: Metropolitan nonmetropolitan disparities in human capital, amenity, and knowledge spillovers. Regional Studies, Vol. 49, No. 2, pp. 270-287.

VLAHOV, D., FREUDENBERG, N., PROIETTI, F., OZONO, S., COOKSON, S., CONNOLLY, M., GALEA, S. 2007. Urban as a determinant of health. *Journal of Urban Health*, Vol. 84, No. 1, pp. 16-26. Doi 10.1007/s11524-007-9169-4

WILLIAMS, A., LARRIER, D., ALTMEYER, W. 2018. A survey of clustering algorithms for big data: Taxonomy and empirical analysis. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 30, No. 8, pp. 1470-1484.

WORLD HEALTH ORGANIZATION (WHO). 1948. *Constitution of the World Health Organization*. Geneva: World Health Organization.

WORLD HEALTH ORGANIZATION (WHO). 2016. *Urbanization and health*. Geneva: World Health Organization.

Daniela VACCA, Istat, daniela.vacca@istat.it
Lucia MARTINA, Istat, lucia.martina@istat.it
Debora TRONU, Istat, debora.tronu@istat.it

WELL-BEING AT SCHOOL

Nidia Batic

Abstract. The concept of personal well-being is multidimensional and encompasses both objective (economic, cultural, environmental, physical) and subjective (psychological, emotional, affective, socio-relational) conditions experienced by the individual. Numerous studies have confirmed the existence of a strong positive relationship between classroom climate and well-being, as well as between well-being and learning. This research is part of an action research framework and does not aim to draw inferential conclusions. Based on the assumption that any discomfort arising within a classroom must first be analyzed (research) and subsequently addressed through targeted interventions (action), the focus of this study was on the social and relational dimensions of well-being in the classroom.
Between February and April 2024, a non-probability sample of 553 children from third, fourth, and fifth grades in several primary schools in the northeast was surveyed using a structured questionnaire with 15 questions. The practical implications of this research are significant, as children provide clear insights into the conditions that foster their well-being. These findings offer valuable guidance for teachers and students alike on the actions that can be taken to improve classroom quality of life, and consequently, enhance learning outcomes.

1. Introduction

The school is a socialization agent that significantly influences pupils, their relationships, and serves as a learning environment for their future. It can be defined as "a main scenery for building social, emotional, and behavioural support among children because students spend a substantial amount of time there" (Tomé *et al.*, 2021, p. 3). It represents an educational context that provides students with the opportunity to engage in cognitive, emotional, relational, and behavioral experiences, which impact their quality of life and well-being (Tobia *et al.*, 2018). Within this context, the teacher plays a central role, not only in the transmission of knowledge but also in fostering interpersonal relationships with pupils (Kanizsa and Zaninelli, 2020) which contributes to the creation of a positive classroom climate, where positive relationships with classmates and teachers support pupils in their learning journey (Kanizsa, 2007). Furthermore, the school is also a physical environment that accommodates students, meaning that physical spaces, materials, and furnishings also affect their experiences and performance (Baker and Bernstein,

2012). Thus, the relationships between peers, teachers, and the environment are crucial elements for establishing a positive classroom climate. It is the outcome of the moods, desires, expectations, behaviors, and values of all individuals—students and teachers alike. Achieving a balance among these components fosters both personal and collective well-being (Polito, 2021).

1.1. Well-being and learning

Well-being is generally understood as synonymous with the quality of people's lives or happiness (Veenhoven, 2000) and can be assessed using both objective measures (such as economic conditions, cultural resources, or health status) and subjective measures (such as perceived quality of life, satisfaction, and happiness) (Statham and Chase, 2010). The World Health Organization's 1946 charter established well-being as a fundamental component of health, stating: "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity". This definition laid the foundation for numerous studies on well-being, all characterized by a multidimensional approach (Hossain *et al.*, 2023) that encompasses cognitive, psychological, social, physical, and material aspects.

Personal well-being is essentially the result of subjective perceptions, closely linked to emotions, which are in turn strongly connected to the ability to interact socially, learn, and internalize knowledge (Stefanini, 2013). The learning process, in fact, is not merely about "storing information"; it is also shaped by emotions, which influence how thoughts and decisions are formed (Vygotsky, 1966). There is a strong parallel between the development of emotions and intellectual functions, as both are involved in every human action (Piaget, 1991). Gardner underscores this idea by noting that emotions, such as enthusiasm for new experiences, play a key role in determining academic success, while less stimulating experiences tend to have little lasting impact (Gardner, 2010). In particular, events experienced with medium-to-high emotional involvement, and thus higher emotional activation, are more likely to be remembered (Palomba and Stegagno, 2007). As Pekrun *et al*. (2002) state, "Academic emotions are significantly related to students' motivation, learning strategies, cognitive resources, self-regulation, and academic achievement, as well as to personality and classroom antecedents" (Pekrun *et al.*, 2002, pag. 91). Positive academic emotions are also associated with higher levels of attention (Vuilleumier 2005), better memorization and learning (Phelps, 2004), improved reasoning (Jung *et al.*, 2014) and enhanced problem-solving abilities (Isen *et al.*, 1987). To summarize with Goleman, Emotional Intelligence governs people's lives, guides their choices and determines their well-being (Goleman, 2020). Students' emotional well-being is shaped by various factors, including a positive attitude towards school

and learning, a healthy self-perception in the academic context, the absence of anxiety and worries about school, and the absence of bullying or discriminatory behaviors (Polito, 2021).

The effect of school climate on learning is significantly mediated by students' well-being. A high-quality school climate is associated with increased student interest in learning and participation in educational activities. Moreover, the perceived quality of relationships between teachers and students is a strong predictor of student engagement (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 and later on as adults by gaining employment, leading a socially engaged life, and contributing to the nation". Other research has found that "that well-being experience influence engagement in school activities beyond the effect of school-climate perception" (Lombardi *et al.*, 2019). Conversely, a negative and hostile classroom climate hampers learning, causing students to retreat into individualism. In such environments, classrooms become spaces of discomfort, making both teaching and learning significantly more challenging (Polito, 2021).

Children's well-being is central to the Convention on the Rights of the Child, adopted by the UN General Assembly in 1989[1] and serves as a strategic prerequisite for realizing all children's rights. This emphasis on well-being is also echoed in the recommendation of 28 November 2022 from the Council of the European Union, which states that "special attention should be paid to well-being at school, as it is a key component of academic success" (p. 22).

2. Objectives, tool and method

"Quality, welcoming atmosphere, inclusion, perceived self-efficacy, positive climate, participation, sharing, dialogue—these are just some of the terms that resonate with the concept of well-being for students and teachers" (Cristalli, 2021), regardless of the order or grade of school. This research focuses on the social and relational dimensions of well-being in the classroom, specifically exploring the "ideal" conditions that help children feel good in their relationships with classmates and teachers. To achieve this, a questionnaire was administered in which children were asked to identify both what makes them feel good and what makes them feel bad. It is important to note that the absence of elements that promote well-being does not necessarily equate to malaise, and vice versa. Following a pre-test conducted with 40 children, the final questionnaire included 15 structured multiple-choice

[1] Ratified in our legal system by Law No. 176 of 27 May 1991, the Convention on the Rights of the Child was adopted in New York on 20 November 1989.

questions. The administration of the questionnaire [2] took place between February and April 2024, involving 33 classes from primary schools in the provinces of Belluno, Gorizia, Pordenone, Trieste, Treviso, Udine, and Venice. The sample comprised 553 pupils, of whom 36.3% were third-graders, 36.2% were fourth-graders, and 27.5% were fifth-graders, with a distribution of 50.3% boys and 49.7% girls.

The research project is situated within an action-research framework (Lewin, 1946), where the objective is not inferential in nature. In this context, the research aims to understand the nature of a problem and gather the necessary information to address it—specifically, to determine effective strategies for promoting well-being in primary school classrooms.

3. Data analysis

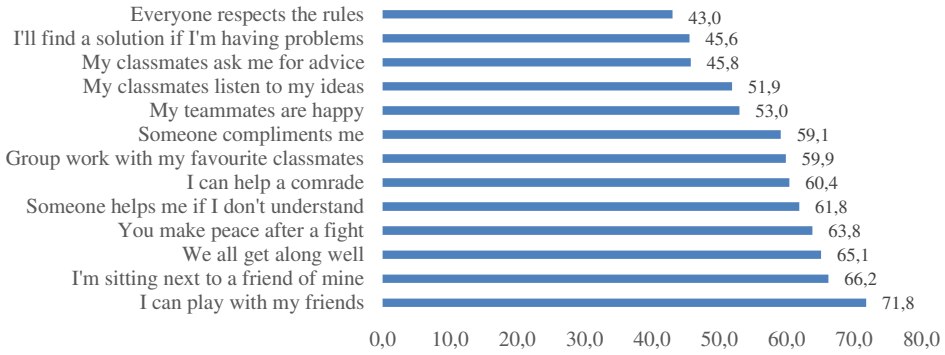
3.1. Well-being factors

The focus of this research was on the social-relational dimension of well-being in the classroom. Key conditions that enhance this sense of well-being include the opportunity for students to share time and space with their friends, as well as fostering an environment where everyone gets along and collaborates with one another (fig. 1). There is a strong consensus among both males and females across all ages, with only three exceptions: males (64.0%) are more inclined than females (55.6%) to feel good when working in a group with preferred peers ($\chi^2 = 4.052$)[3]; males (50.0%) also appreciate it more than females (41.5%) when classmates ask for advice during group work ($\chi^2 = 4.068$). Additionally, fourth-graders (68.0%) are more sensitive to compliments than fifth-graders (55.3%) and third-graders (53.2%) ($\chi^2 = 10.343$).

[2] The questionnaires were administered by trainee students from the Educational Sciences program at the University of Udine, in the presence of the class teachers. Prior to administering the questionnaire, the research project received approval from the school principals, and an information form was signed by the families, granting consent for their children to participate. The data were collected anonymously and processed in aggregate form, fully complying with privacy regulations according to the GDPR (2016/679) and Legislative Decree 101/2018.

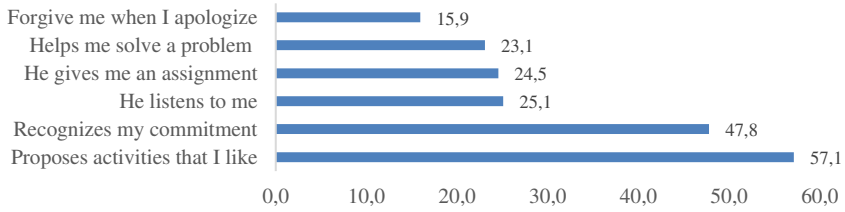
[3] To verify the existence of statistically significant differences in opinions among children, the chi-square test was applied to all responses based on gender and class. Only significant test values with $\alpha = 0.05$ are reported in the text. Additionally, to determine whether gender or age profiles could be identified, Cramer's V test was used, which highlighted a "fair association" in only two cases, as will be detailed in the paper.

Figure 1 – *Well-being factors in the classroom (percentages).*



Feeling good at school also involves having a positive relationship with the teacher. To explore this, the children were asked to identify the behaviors of teachers that contribute to their well-being in class (fig. 2). Overall, opinions were consistent, except in terms of the ability to propose enjoyable activities, which was indicated more frequently by males (63.4%) than by females (50.7%) ($\chi^2 = 8.939$).

Figure 2 – *Elements of well-being in the relationship with the teacher (percentages).*



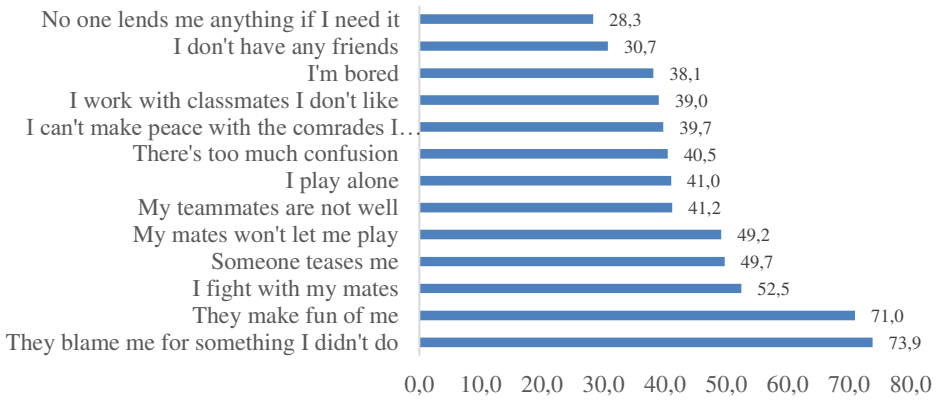
Regarding age, older children tend to feel good when the teacher assigns them work (57.0%), while this sentiment is less pronounced among fourth-graders (49.5%) and third-graders (39.1%) ($\chi^2 = 14.923$). Conversely, third-graders feel more rewarded by their teachers when their efforts are recognized, with 33.5% expressing this sentiment, followed by fifth-graders at 22.8% and fourth-graders at 17.0% ($\chi^2 = 11.332$).

3.2. Factors of discomfort

The children were asked to identify not only the situations that promote well-being but also those that make them feel uncomfortable at school. In fact, it can sometimes be easier to pinpoint unpleasant situations and, from those, infer the

opposite behaviors that contribute to well-being (fig. 3). There is only one notable difference between the genders: males (42.6%) report feeling more affected by boredom than females (33.6%) (χ^2 = 4.754). Considering age, certain negative situations diminish in impact as students progress from third to fifth grade. For instance, feelings of discomfort arise when being excluded from games (χ^2 = 6.302), being the target of teasing (χ^2 = 7.855), or when no one lends them something they need (χ^2 = 7.205).

Figure 3 – *Situations in which children do not like to stay at school (percentages).*

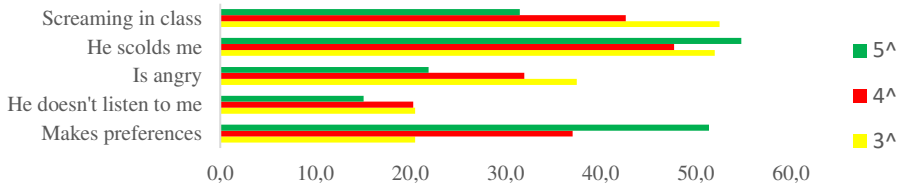


To feel good in the classroom, children also identify certain behaviors that should be avoided, as they cause annoyance and discomfort. Both boys and girls agree that they feel bad when a classmate is treated poorly (65.2%), when someone damages their classmates' materials (64.8%), or when there is disruption during the teacher's explanation (61.4%). Fewer inconveniences arise from failing to comply with certain rules, such as respecting traffic queues while moving (56.6%), raising hands before speaking (48.7%), or damaging school property (46.7%). However, differences emerge between classes: as students age, the annoyance caused by behaviors related to a lack of respect tends to decrease. This includes not raising a hand to speak (χ^2 = 7.245), not remaining silent in the classroom (χ^2 = 7.184), not respecting the line while moving (χ^2 = 8.273), and damaging the environment (χ^2 = 10.750). It can be inferred that these behaviors significantly diminish or disappear from third-graders to fifth-graders, to the extent that they no longer represent disruptive elements. Nevertheless, there is still a collective disapproval when someone hurts a classmate or damages their belongings.

In the relationship with the teacher, being reprimanded causes more distress, and few report a lack of listening as a factor contributing to their discomfort (fig. 4). This may be attributed to teachers being attentive and responsive to students' needs, which

diminishes the sensation of feeling unheard. There are no gender differences noted; however, perceptions vary by class level. Students in fifth grade report feeling less discomfort compared to earlier grades when the teacher gets angry ($\chi^2 = 9.593$), shouts in class ($\chi^2 = 15.202$), or shows favoritism ($\chi^2 = 36.131$, Cramer's V=0,258).

Figure 4 – *Situations that cause discomfort in the relationship with the teacher (percentages).*



Since well-being results from both psychological and physical conditions, children were asked about experiencing headaches at school, and the data provide important insights: only 21.6% of children report never having headaches, while 15.0% say they experience them almost daily or all the time. The majority (64.4%) report having headaches occasionally. The frequency distribution indicates that females tend to experience headaches more than males ($\chi^2 = 13.935$), with no significant differences observed between classes. There may be issues related to eyesight—such as struggling to read the blackboard (19.0%) or reading for too long (15.9%)—or learning-related problems, including anxiety about a test (27.2%) or difficulties with a subject (33.0%). However, the challenges posed by a noisy classroom environment are the most significant, with many children reporting distractions from loud voices (69.3%) and noise in general (50.1%). Mathematics is the subject that creates the most tension for these children, leading to headaches (35.0%), followed by Italian (18.2%) and science (13.1%). No gender differences are observed, but third-graders are significantly more sensitive to noise ($\chi^2 = 9.459$), loud voices ($\chi^2 = 7.046$), and excessive light in the classroom ($\chi^2 = 6.239$) compared to fourth and fifth-graders.

3.3. Behaviors to improve well-being in the classroom

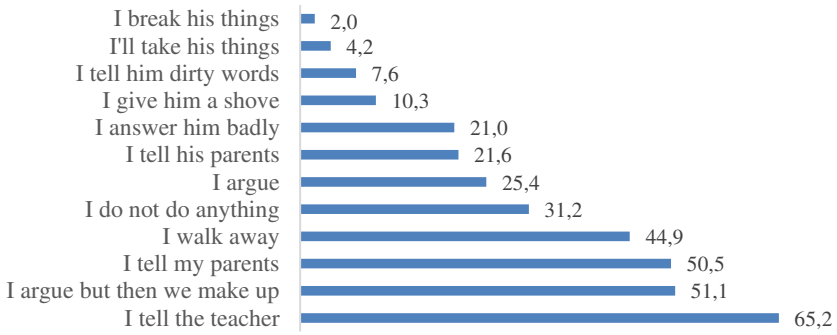
The classroom climate can also be influenced by relationships among classmates, with children generally preferring to resolve conflicts either by intervening directly or by asking the teacher for help (fig. 5). Gender differences are evident in how children choose to handle their issues: boys are more likely to stay away from conflicts (20.9%) compared to girls (8.4%) ($\chi^2 = 17.429$), while girls are more

inclined to ask the disputing parties to resolve their conflicts (79.3%) than boys (66.1%) (χ^2 = 12.114). Regarding age, there is a significant variation in opinions about the appropriateness of seeking teacher intervention (χ^2 = 16.626) or informing the parents of children who argue (χ^2 = 9.961). This tendency is more common among third-graders and decreases by fifth grade.

Figure 5 – *What should be done when classmates argue (percentages).*

Let us now consider children's reactions when someone makes them angry (fig. 6). It is interesting to note that a quarter of the children react by arguing (25.4%), while 51.1% argue but eventually make up. Although few children are aggressive or vindictive, the majority respond not only by reacting directly but also by seeking the teacher's help (65.2%).

Figure 6 – *How children react when a classmate makes them angry (percentages).*



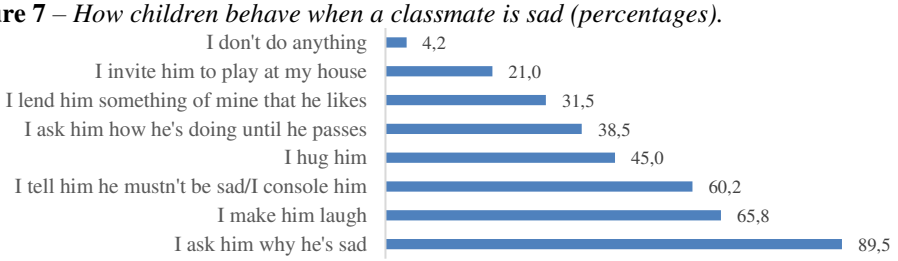
Males (15.9%) are more likely than females (4.7%) to react by pushing (χ^2 = 18.552) or swearing (10.8% vs. 4.4%, χ^2 = 8.209). In contrast, females (59.3%) more frequently tell their parents about conflicts than males (41.9%) (χ^2 = 16.705). Additionally, females tend to reconcile after a quarrel more often than males (55.6% vs. 46.6%, χ^2 = 4.539). Age also influences reactions to provocation, with older children displaying more moderate behaviors compared to younger ones. This is particularly evident in their responses, such as pushing (χ^2 = 6.907), informing the parents of the child who angered them (χ^2 = 28.286), or talking to their own parents

(χ^2 = 10.006) or the teacher (χ^2 = 17.700), and even choosing to do nothing (χ^2 = 10.985).

The children in the sample not only avoid quarrels but also demonstrate a high level of generosity by helping their classmates in times of need. For instance, if someone lacks school supplies, 79.0% offer to lend something, with an additional 13.8% doing so only if it is for a friend. There are no statistically significant distinctions based on gender or age.

Generosity and the desire to support others are also evident when a classmate is experiencing sadness or difficulty. Only 4.2% (23 children in total) would ignore their peer's discomfort. In contrast, the majority would seek to help in various ways, starting with simply asking why their peer is sad (89.5%). Other responses include making them laugh (65.8%), attempting to console them (60.2%), or offering a hug (45.0%) (fig. 7).

Figure 7 – *How children behave when a classmate is sad (percentages).*



Females show a greater interest than males in the emotional state of their classmates, as evidenced by their willingness to ask why someone is sad (94.9% vs. 84.2%, χ^2 = 16.975), try to comfort them (65.1% vs. 55.4%, χ^2 = 5.424), and offer hugs (58.2% vs. 32.0%, χ^2 = 38.243, Cramer's V = 0.263). Conversely, males are more likely than females to invite their classmates to play at home (24.5% vs. 17.5%, χ^2 = 4.093). There is a notable consistency in behavior across different ages; however, differences do exist in the level of concern for a classmate's sadness (χ^2 = 10.514) and the willingness to lend something to make them happy (χ^2 = 11.778).

Even when a classmate is angry, children's first response is to show interest in him/her by asking what is going on (83.8%), trying to calm him/her down (69.6%), and encouraging him/her not to act violently or damage others' belongings. Both males and females generally agree on these approaches; however, the recommendation to refrain from violence is more common among third and fourth graders (χ^2 = 8.518), as is the suggestion not to break things (χ^2 = 17.371).

To conclude, we asked what teachers should do when children do not respect rules of appropriate behavior in order to restore a positive classroom climate. When a classmate is mocked, the teacher should apologize (48.6%) and/or reprimand the

student (42.4%). If children argue, the teacher should encourage them to apologize (42.7%) and also talk to them to understand the reason behind the conflict (35.8%). If students are noisy during the lesson, the teacher should scold them (52.0%) and remind them of the importance of rules (29.8%). In cases where students use swear words, parents should be informed (53.8%) alongside the scolding (31.9%). If children harm their classmates, it is essential to notify their parents (33.1%) and reprimand the children (28.2%). When speaking turns are not respected, students should be reprimanded (47.2%), and the significance of rules should be reiterated (28.6%).

4. Conclusion: from research to action

This research has identified factors that contribute to children's well-being at school, as well as those that cause discomfort, and has illustrated how children react to enhance the well-being of their entire class. It is now essential for teachers to integrate this information into their classrooms and use it as a foundation for designing actions that promote well-being and improve learning outcomes.

The findings indicate that children's expressions of discomfort vary with age; however, the importance of friendships remains significant for all. Teachers are therefore tasked with fostering social relationships and guiding students in recognizing and managing their emotions. Implementing a self-awareness program can start as early as primary school.

Another critical area to address is the education around respecting rules. While respecting rules was initially perceived as the least important factor for creating well-being (43.0%), teachers play a pivotal role in managing inappropriate behaviors (such as fighting or using foul language). Thus, a comprehensive "Respect Project" should be integrated into the entire primary school curriculum, focusing on social interactions (both with peers and adults), community relationships, and connections with the natural environment.

Teachers can promote well-being by introducing activities that spark students' curiosity and interest without compromising essential teaching and planning requirements. They can easily create an environment that minimizes physical discomfort, keeping in mind that, from an action-research perspective, any reported discomfort by even a single child necessitates intervention, as the goal is to ensure that every child feels included and comfortable.

Lastly, educators must reflect on their teaching styles and their relationships with students. This includes avoiding shouting, recognizing student efforts, treating all students equitably without showing favoritism, and, most importantly, actively listening to their needs and concerns.

References

BAKER L., BERNSTEIN H. 2012. *The Impact of School Buildings on Student Health and Performance*. McGraw-Hill. https://eddesignaward.com/research/ wp-content/uploads/2021/12/the-impact-of-school-buildings-on-student-health-and-performance_lindsay-baker.pdf

CRISTALLI B. 2021. *Benessere a scuola per studenti e insegnanti*. https://www.focus-scuola.it/benessere-a-scuola-per-studenti-e-insegnanti/

COUNCIL OF THE EUROPEAN UNION 2022. *Council Recommendation on Pathways to School Success and replacing the Council Recommendation of 28 June 2011 on policies to reduce early school leaving*. https://www.consilium.europa.eu/media/60390/st14981-en22.pdf.

FATOU N., KUBISZEWSKI V. 2018. Are perceived school climate dimensions predictive of students' engagement?, *Social Psychology of Education*, Vol. 21, pp. 427–446.

GARDNER H. 2010 (first edition 1983). *Formae mentis. Saggio sulla pluralità della intelligenza*. Milano: Feltrinelli.

GOLEMAN D. 2020 (first edition 1995). *Intelligenza Emotiva. Che cos'è e perché può renderci felici*. Milano: Rizzoli.

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.

ISEN A. M., DAUBMAN K. A., NOWICKI G. P. 1987. Positive affect facilitates creative problem solving, *Journal of personality and social psychology*, Vol. 52, No. 6, pp. 1122-1131.

JUNG N., WRANKE C., HAMBURGER K., KNAUFF M. 2014. How emotions affect logical reasoning: evidence from experiments with mood-manipulated participants, spider phobics, and people with exam anxiety, *Frontiers in psychology*, Vol. 5, Article 570, https://doi.org/10.3389/fpsyg.2014.00570.

KANIZSA S. (2007). *Il lavoro educativo: l'importanza della relazione nel processo di insegnamento-apprendimento*. Milano: Bruno Mondadori.

KANIZSA S., ZANINELLI F. L. (Eds) 2020. *La vita a scuola*. Milano: Raffaello Cortina Editore.

LEWIN K. 1946. Action Research and Minority Problems, *Journal of Social Iusses*, Vol. 2, pp. 34-46.

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.

PALOMBA D., STEGAGNO L. 2007. Emozioni e memoria: riscontri neuro e psicofisiologici, in E. AGAZZI E. and FORTUNATI V. (Eds) *Memoria e saperi: percorsi transdisciplinari*. Roma: Meltemi, pp. 229-242.

PEKRUN R., GOETZ T., TITZ W., PERRY R. P. 2002. Academic emotions in students' selfregulated learning and achievement: a program of qualitative and quantitative research, *Educational Psychologist*, Vol. 37, pp. 91-105.

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

PIAGET J. 1991 (first edition 1936). *La nascita dell'intelligenza nel fanciullo*. Firenze: Giunti e Barbera.

POLITO M. 2021. *Benessere in classe e apprendimento*. Independently Published by Amazon.

STATHAM J., CHASE E. 2010. *Childhood wellbeing: A brief overview*. Loughborough University: Childhood Wellbeing Research Centre, www.researchgate.net/profile/June_Statham/publication/242676811_Childhood_Wellbeing_A_brief_overview/links/549bd87c0 cf2b80371372fc7.pdf

STEFANINI A. 2013. *Le emozioni: Patrimonio della persona e risorsa per la formazione*. Milano: Franco Angeli.

TOBIA V., GRECO A., STECA P., MARZOCCHI G.M. 2018. Children's Wellbeing at School: A Multi-Dimensional and Multi-Informant Approach, *Journal of Happiness Studies*, Vol. 20, No. 3, pp. 841-861.

TOMÉ G., ALMEIDA A., RAMIRO L., GASPAR T., GASPAR DE MATOS M. 2021. Intervention in Schools promoting mental health and well-being: a systematic review, *Global Journal of Community Psychology Practice*, Vol. 12, No. 1, pp.1-23.

VEENHOVEN R. 2000. The four qualities of life, *Journal of Happiness Studies*, Vol.1, pp.1–39.

VUILLEUMIER P. 2005. How brains beware: neural mechanisms of emotional attention, *Trends in cognitive sciences*, Vol. 9, pp. 585-594.

VYGOTSKY L. S. 1966 (originale 1934). Pensiero e linguaggio. Firenze: Giunti e Barbera.

WORLD HEALTH ORGANIZATION 2020. *Basic documents*. Forthy ninth edition.

Nidia BATIC, Università di Udine, nidia.batic@uniud.it

GENDER DISPARITIES IN SCHOOL-TO-UNIVERSITY TRANSITION IN ITALY: THE ROLE PLAYED BY THE SOCIO-ECONOMIC CONDITION AND THE TYPE OF HIGH SCHOOL[1]

Valentina Tocchioni, Samuele Milone, Gabriele Lombardi

Abstract. The issue of the gender gap in university enrolment is widely investigated, and it is acknowledged that the school-to-university transition tends to be higher for female students than for males. Gender inequalities in access to the higher education system could derive from the persistence and transmission of gender disparities relying on several factors; in particular, in this work, we intend to assess potential gender inequalities in the school-to-university transition, which could be moderated by different students' socioeconomic conditions or type of high school. Using the integration between two administrative data sources, the INVALSI database and the Anagrafe Nazionale Studenti (ANS) database, our final dataset comprises a total of 420,261 students of grade 13 who graduated in the school year 2018/2019, of which 215,565 are females (51.3%). Among those, 52.6% of all students enrolled in the following 2019/2020 academic year.
Our results suggest that the proportion of female students enrolling in university is consistently higher than that of male students, regardless of the type of diploma or the socioeconomic conditions of the family of origin. Nevertheless, the gender disparity in the school-to-university transition is more pronounced among students coming from technical and vocational high schools and among those with low/middle socioeconomic and cultural status, leaving male students with less prestigious backgrounds even further behind.

1. Introduction

The past few decades have constituted a golden era for Europe regarding the proliferation of individuals with university degrees. Two decades ago, in 2003, only 24.6% of Europeans between the ages of 25 and 34 possessed a tertiary education degree. However, by 2023, this value had risen to 43.1% (Eurostat, 2024). This marked increase offers a promising outlook for the future, showcasing a positive trend that underscores Europe's educational progress. Indeed, these data are often cited to highlight the progress the European Union is making in increasing the

[1] This article is the joint work of the authors. In particular, V. Tocchioni dealt with conceptualization and supervision; S. Milone dealt with data curation, formal analyses and visualisation; G. Lombardi dealt with supervision. All authors contributed to the writing: Tocchioni wrote section 1, 3, 4; Milone wrote section 1, 2, 3; Lombardi wrote section 1. All authors reviewed the final version of the writing.

number of educated individuals, even though it has not yet achieved the strategic goal of 45% in all member states (Council of the European Union, 2021).

The benefits of a highly educated population are extensively documented in academic literature (Bynner *et al.*, 2002), which points to numerous advantages, including enhanced economic growth (Barro, 2013), increased innovation (Bartel and Lichtenberg, 1987), and greater social cohesion (Green and Preston, 2001). A population with higher educational attainment levels is generally better equipped to tackle complex global challenges, from technological advancements to environmental sustainability. This progress in education is often viewed as a crucial element in maintaining Europe's competitiveness on the worldwide stage.

However, it is crucial to note that the overall statistic of 43.1% masks significant disparities between and within countries. For instance, while some nations have surpassed the 45% target, others, such as Italy, lag significantly behind, creating a patchwork of educational attainment across the continent. Moreover, within countries, there are pronounced differences in educational attainment by gender, with women often having higher rates of university education than men, which can lead to new societal and economic dynamics. In Italy, 37.1% of women aged 25 to 34 have a university degree, while the percentage among men drops significantly to 24.4%. This difference (12.7%) is slightly higher than the European average of 11.2% (Eurostat, 2024).

The transition through the higher education system is an essential investment at the individual level, too, which may also highly favour social displacement for students from underrepresented groups (Briggs *et al.*, 2012). Nevertheless, recent research has highlighted that a general increase in educational attainment does not necessarily translate to equitable progress across all social groups (Breen *et al.*, 2010). Disparities based on gender, socioeconomic status, and geographic location remain pronounced, suggesting that while more people obtain degrees, the benefits are not evenly distributed.

Looking at socioeconomic status (SES), it is acknowledged that students from families with higher status have a greater likelihood of enrolling in university (Contini *et al.*, 2018), detecting more severe difficulties in transitioning for those students who come from low socioeconomic backgrounds and geographical areas where the public schools are not well resourced (McGhie, 2017). Nonetheless, Bletsas and Michell (2014) attribute such evidence to the existence of a perverse mechanism: the fact that low SES students are considered as lacking in motivation and preparation compared to their higher SES peers - even if they perform just as well when attending (see, among others: Bradley *et al.*, 2008; Kleemola *et al.*, 2023) - justifies public disinvestment for their educational environments, so that the risks of reproducing such a mechanism further increase. Low SES students in transition are the most affected by financial strain, time pressures, unclear expectations, low

confidence, preparedness, family support and aspirations. Nevertheless, they are also more likely to independently challenge the limitations of their identity and career path (Devlin and McKay, 2014).

Finally, among the critical factors affecting the school-to-university transition, the high school's curriculum also plays a role. For Italy, empirical research highlighted that the two curricula with the lowest rates of university enrolment are the technological and vocational tracks (Aina *et al.*, 2011; Contini *et al.*, 2018), where female students are typically underrepresented, as well as high SES students.

Our objective is to explore how the family socioeconomic background and the high school curriculum may moderate gender disparities affecting the individual propensity to enrol in the university. Gender can interact with family background and high school curriculum, potentially amplifying or reducing their effects. Moreover, concerning Italy, very few studies have addressed the topic of university enrolment by gender, being more concentrated on other issues about students' academic careers, such as the persisting existence of gender segregation across fields of studies (Triventi, 2010; Barone and Assirelli, 2020).

Previous empirical studies showed that also in the Italian context, women are less likely to leave school early, more resilient, more likely to enrol on university, and more likely to adopt effective coping strategies for adapting themselves during the transition through tertiary education (Borgna and Struffolino, 2017; Cabras and Mondo, 2018); at the same time, students from families with a higher socioeconomic background and students coming from lyceums have a greater likelihood of enrolling in university (Contini *et al.*, 2018). Still, it is not straightforward how the type of high school and the socioeconomic condition of the family of origin moderate the likelihood of university enrolment by gender. For this reason, we ask ourselves the following two research questions: to what extent does the relationship between gender and university enrolment differ according to the student's type of high school? Does the association between gender and university enrolment vary with the student's socioeconomic background?

2. Data and methods

2.1. Data

Our study examines gender differences in Italy's transition from high school to university. To do so, we draw upon two administrative data sources: the first, namely the Istituto Nazionale per la VALutazione del Sistema educativo di Istruzione e di formazione (INVALSI; in English: National Institute for the Evaluation of the Education and Training System) database, contains information about all students of

grade 13th in the school year 2018-2019 who took the INVALSI test and graduated in the same year in an upper secondary school located in the Italian territory; while the second is the Anagrafe Nazionale Studenti (ANS; in English: National Student Registry) database, which includes all students enrolled in an Italian university in the academic year 2019-2020, thus in the year following graduation.

By doing so, our dataset comprises 420,261 students in grade 13th, of which 215,565 are females (51.3%). University enrollees are 220,984 (52.6% of all students considered), and female students are 124,746 (56.5%).

2.2. Methods

As for the methodology, first, we considered a school-level indicator: the entry rate, computed as the ratio between the number of students who took the INVALSI test in a specific year and the number of students enrolling in university the following year. This indicator (referred hereafter to as the female/male entry rate) has been calculated separately for the female and male populations for the three school years available[2].

Second, to investigate whether and how the association between gender and university enrolment varies by type of high school and socioeconomic condition of the family of origin in Italy, we estimated multilevel logistic regression models for the school-to-university transition. The chosen model is the following:

$$\begin{aligned} logit\left(P_{enrol_{ij}}\right) = \quad & \beta_0 + \beta_1 Gender_{ij} + \sum_{q=2}^{Q} \delta_q D^{(q)}_{ESCSquartile_{ij}} + \\ & \sum_{c=2}^{C} \gamma_c D^{(c)}_{curriculum_j} + \sum_{q=2}^{Q} \eta_q Gender_{ij} D^{(q)}_{ESCSquartile_{ij}} + \\ & \sum_{c=2}^{C} \zeta_c Gender_{ij} D^{(c)}_{curriculum_j} + +\beta_2 Ita\_score_{ij} + \\ & \beta_3 Math\_score_{ij} + \sum_{m_v=2}^{M_v} \theta_{m_v} D^{(m_v)}_{v_{ij}} + \sum_{n_v=2}^{N_v} \vartheta_{n_v} D^{(n_v)}_{v_j} + u_{0j} \end{aligned} \tag{1}$$

This model posits that the probability of enrolling on the university of a student (the i^{th} student in j^{th} school) can be described as the effects of school-level and student-level covariates, as well as their interactions, and the effect of a residual error term of being in school j (u_{0j}). $D^{(q)}_{SESquartile_{ij}}$ represents a set of dummies for the categories q that the socioeconomic and cultural status (ESCS index) into quartiles can take for student i in school j; similarly, $D^{(c)}_{curriculum_j}$ represents a set of dummies, one for each category c that the curriculum variable can take for school j. In the

[2] The three available years are: scholastic year 2018/19, 2020/21, and 2021/22. In the subsequent analysis, the chosen model will be applied only to the first year due to numerous missing values for the control variables in the subsequent years.

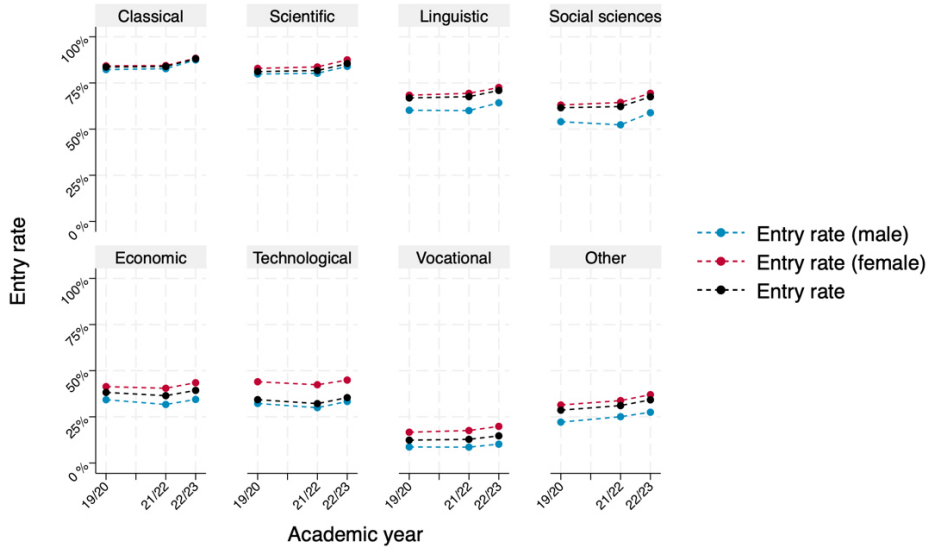
equation, we have also included the interaction between gender and ESCS quartile, as well as the interaction between gender and curriculum.

Finally, we considered four control variables (v) for student i in school j: scores on the INVALSI standardized tests in Italian and Maths (as continuous), migration background (native, first-generation immigrant, or second-generation immigrant), first-term grades in Mathematics and Italian (divided into four categories). Additionally, we included two control variables for school j: the macro area where the high school is located (North-West, North-East, Centre, South, and Islands) and whether the high school was private or public.

3. Results

In Figure 1, entry rates are presented by high school curricula, with the four different types of lyceums in the first row, and the technical and vocational curriculum in the second row.

Figure 1 - *Entry rate by high school curriculum.*



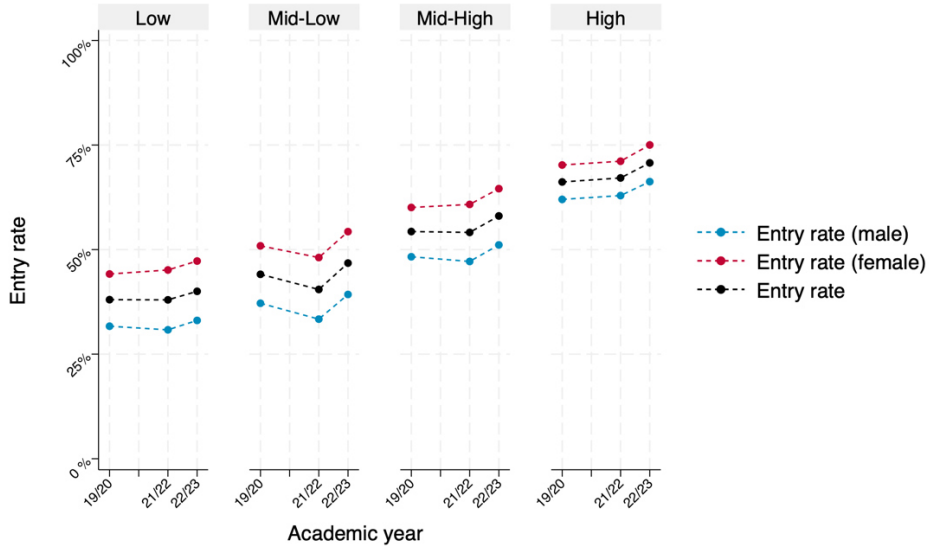
Note: own elaborations on INVALSI data and ANS data.

Regardless of the school curriculum, the percentage of female students enrolling in university is consistently higher than that of male students. This trend highlights a gender disparity in the school-to-university transition, with a prevalence of females

accessing higher education. Moreover, the entry rate varies considerably according to the school curriculum, ranging from a minimum in the Vocational track (13.3%) to a maximum in the Classical track (85.3%).

Figure 2 presents the entry rate values for the four quartiles based on the ESCS index, where "low" represents the lowest socioeconomic and cultural status, and "high" means the highest. Here, it is evident that the strong influence of the family's ESCS on students' likelihood of enrolling in university is evident: as the family ESCS rises, so does the entry rate. This gap is particularly striking when comparing the extreme quartiles, with students with the lowest ESCS having an entry rate roughly half of the entry rate observed for students with the highest ESCS (68.1%). This stark contrast highlights the crucial role that socioeconomic and cultural factors play in shaping educational outcomes and access to higher education opportunities.

Figure 2 - *Entry rate by ESCS.*

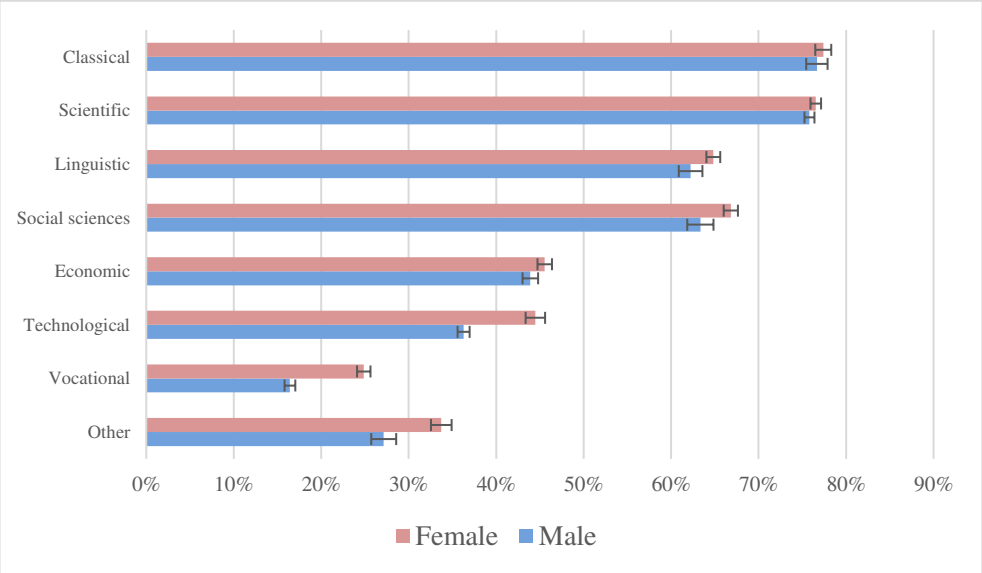


Note: own elaborations on INVALSI data and ANS data.

In the two following figures, predicted values for the probability of enrolling in university after graduation and deriving from the results of the logistic model (1) are presented by gender and high school curriculum (see Figure 3) and by gender and ESCS (see Figure 4), whereas in Table 1 odds ratios from the estimated model are reported. As expected, the highest probability of enrolling in university is for male and female students who come from a classical or a scientific track, showing at most

1.2 percentage points (p.p.) between the two genders in favour of females. For other types of lyceums (namely, linguistic and social sciences), differences in the probability of enrolment are still modest. On the contrary, female students are more likely to transition to university than their male counterparts coming from a technical or a vocational track, with the highest gap for the last type of high school curriculum (9.1 p.p. in favour of female students).

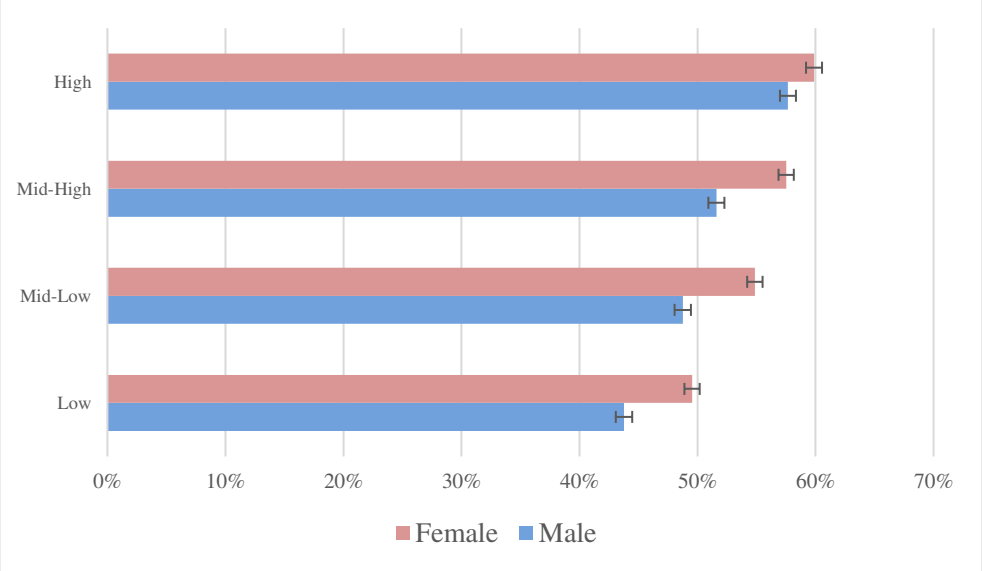
Figure 3 - *Predicted probability of enrolling to the university by gender and high school curriculum.*



Note: Control variables: high school macro area, student's migratory background, first-term grades in Maths and Italian, scores on the INVALSI tests in Maths and Italian, private/public high school.

Looking at the heterogeneity by socioeconomic and cultural background (see Figure 4), the gender gap in the predicted probability of transitioning to university is similar for the three groups of low, mid-low and mid-high ESCS students, who have a disparity in predicted likelihood in favour of female students around 4.2-4.6 p.p.; instead, the difference is reduced to 2.2 p.p. (nearly halved) among students with a high ESCS. Thus, once the student's socio-demographic characteristics and school career are controlled for, the gender gap is modest, concerning what is illustrated in Figure 2. Moreover, it is essential to note that the gaps between ESCS levels are much higher than within ESCS levels by gender.

Figure 4 - *Predicted probability of enrolling to the university by gender and ESCS.*



Note: Control variables: high school macro area, student's migratory background, first-term grades in Maths and Italian, scores on the INVALSI tests in Maths and Italian, private/public high school.

4. Conclusions and discussion

In this work, we investigated the issue of the gender gap in university enrolment, in particular, assessing how potential gender inequalities in the school-to-university transition could be moderated by different students' socioeconomic conditions or types of diplomas.

We found that, regardless of the type of diploma or the socioeconomic conditions of the family of origin, the propensity of female students to enrol in university is consistently higher than that of male students, thus confirming a higher investment in the education of girls than boys across the different high school curricula and social strata. In particular, this universal advantage of girls across social strata highlights that the selection effect, which once led only girls from higher social classes to achieve higher education due to female emancipation that began in those classes, is no longer visible. Nevertheless, we found both horizontal heterogeneity among high school curricula and vertical heterogeneity among socioeconomic conditions of the family of origin.

Table 1 – *Model results of enrolling to the university by gender and ESCS (odds ratio).*

	OR	SE
Gender (base: Male)		
Female	1.075*	0.044
ESCS quartile (base: Low)		
Mid-Low	1.222***	0.020
Mid-High	1.370***	0.022
High	1.750***	0.030
Curriculum (base: Classical)		
Scientific	0.954	0.035
Linguistic	0.501***	0.022
Social sciences	0.526***	0.024
Economic	0.238***	0.009
Technological	0.173***	0.007
Vocational	0.060***	0.002
Other	0.113***	0.006
Gender × ESCS quartile		
Female × Mid-Low	1.013	0.022
Female × Mid-High	1.007	0.022
Female × High	0.869***	0.020
Gender × Curriculum		
Female × Scientific	0.999	0.041
Female × Linguistic	1.074	0.051
Female × Social sciences	1.119**	0.056
Female × Economic	1.026	0.044
Female × Technological	1.351***	0.060
Female × Vocational	1.616***	0.074
Female × Other	1.311***	0.070

Note: *** p<.01, ** p<.05, * p<.1. *Control variables: high school macro area, student's migratory background, first-term grades in Maths and Italian, scores on the INVALSI tests in Maths and Italian, private/public high school.*

The horizontal heterogeneity between the two genders in the attended high school is reflected in two points. First, we identified a very narrow gender gap in pathways where the transition to tertiary education is very high, e.g., classical and scientific tracks: in this case, the advantage in continuing studies is very similar between genders, given a student's social class and academic performance. For those high school tracks, continuing in higher education is very common, and most students who attend these schools decide very early (even if their decision may be revised) to enrol to university, irrespective of their gender. Second, we found a pronounced gender gap in pathways where the transition to tertiary education is low, like in technical and even more in vocational tracks, where the propensity to continue studies in higher education is more remarkable for girls. This striking gender gap may be partly explained by the persisting existence of gender segregation across fields of studies in Italy, which is already visible in technical and vocational

tracks (Triventi, 2010; Barone and Assirelli, 2020). Indeed, boys tend to be more inclined towards technical pathways that require more physical strength and offer better employment prospects in demand, supply, and salaries, such as in the industrial or IT sector. Conversely, girls tend to gravitate towards care-oriented pathways, with worse employment prospects in demand, supply, and salaries. Consequently, having a tertiary education degree can enhance girls' skills and their chances of better job opportunities in the labour market.

Regarding the vertical heterogeneity among socioeconomic classes, the gender gap is smaller in the highest social class than the others, thus reversing the selection effect favouring only high SES girls existing in the past. Nowadays, boys from higher social classes have more to gain from continuing their studies than boys from lower classes once their socio-demographic characteristics and school abilities are considered. Thus, whereas students from high social-class families continue their studies regardless of their abilities, students from low social-class families tend to drop out. In terms of mechanisms, this underscores the importance of social origin in transmitting socioeconomic advantages of the affluent classes and consolidating their original social position through education, acting equally across both genders.

To sum up, the gender disparity in the school-to-university transition is more pronounced among students coming from technical and mainly vocational high schools and among those with a mid or low socioeconomic and cultural status, leaving male students with less prestigious backgrounds even further behind in terms of tertiary education (and all the consequent impacts on employment prospects and earnings) compared to female students with the same characteristics. With our work, we intend to shed light on these dynamics, looking at which factors contributed to gender-unbalanced enrolments at universities in Italy.

Acknowledgements

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU– Project Title “From high school to university: Assessing peers' influence in educational inequalities and performances” – CUP F53D23006150006 - Grant Assignment Decree No. 1060 adopted on 07/17/2023 by the Italian Ministry of University and Research (MUR).

The data used in this study have been processed by the RESEARCH PROTOCOL FOR THE STUDY "From high school to the job placement: analysis of university careers and university mobility from Southern to Northern Italy" among the MUR,

the Ministry of Education and Merit, the University of Palermo as the lead institution, and the INVALSI Institute. The reference researcher is Bruno Bertaccini.

References

AINA C., BAICI E., CASALONE G. 2011. Time to a degree: students' abilities, university characteristics or something else? Evidence from Italy, *Education Economics*, Vol. 19, No. 3, pp. 311–325.

BARONE C., ASSIRELLI G. 2020. Gender segregation in higher education: an empirical test of seven explanations, *Higher Education*, Vol. 79, No. 1, pp. 55–78.

BARRO R.J. 2013. Education and Economic Growth, *Annals of Economics and Finance*, Vol. 14, No. 2, pp. 301–328.

BARTEL A.P., LICHTENBERG F.R. 1987. The Comparative Advantage of Educated Workers in Implementing New Technology, *The Review of Economics and Statistics*, Vol. 69, No. 1, p. 1.

BLETSAS A., MICHELL D. 2014. Classism on campus? Exploring and extending understandings of social class in the contemporary higher education debate. In BROOK H., FERGIE D., MAEORG M., MICHELL D. (Eds.) *Universities in Transition: Foregrounding Social Contexts of Knowledge in the First Year Experience*, University of Adelaide Press, pp. 77–96.

BORGNAC., STRUFFOLINO E. 2017. Pushed or pulled? Girls and boys facing early school leaving risk in Italy, *Social Science Research*, Vol. 61, pp. 298–313.

BRADLEY D., NOONAN P., NUGENT H., SCALES B. 2008. *Review of Australian higher education: final report.* Canberra, A.C.T.

BREEN R., LUIJKX R., MÜLLER W., POLLAK R. 2010. Long-term Trends in Educational Inequality in Europe: Class Inequalities and Gender Differences, *European Sociological Review*, Vol. 26, No. 1, pp. 31–48.

BRIGGS A.R.J., CLARK J., HALL I. 2012. Building bridges: understanding student transition to university, *Quality in Higher Education*, Vol. 18, No. 1, pp. 3–21.

BYNNER J., DOLTON P., FEINSTEIN L., MAKEPEACE G., MALMBERG L., WOODS L. 2002. Revisiting the Benefits of Higher Education.

CABRAS C., MONDO M. 2018. Coping strategies, optimism, and life satisfaction among first-year university students in Italy: gender and age differences, *Higher Education*, Vol. 75, No. 4, pp. 643–654.

CONTINI D., CUGNATA F., SCAGNI A. 2018. Social selection in higher education. Enrolment, dropout and timely degree attainment in Italy, *High Educ*, Vol. 75, No. 5, pp. 785–808.

COUNCIL OF THE EUROPEAN UNION. 2021. Council Resolution on a strategic framework for European cooperation in education and training towards the European Education Area and beyond (2021-2030), 2021/C 66/01.

DEVLIN M., MCKAY J. 2014. Reframing "the problem": students from low socio-economic status backgrounds transitioning to university. In BROOK H., FERGIE D., MAEORG M., MICHELL D. (Eds.) *Universities in Transition: Foregrounding Social Contexts of Knowledge in the First Year Experience*, University of Adelaide Press, pp. 97–126.

EUROSTAT. 2024. Population by educational attainment level, sex, age and citizenship (%), Online data code: edat_lfs_9911, Last data update: 13/06/2024.

GREEN A., PRESTON J. 2001. Education and Social Cohesion: Recentering the Debate, *Peabody Journal of Education*, Vol. 76, No. 3/4, pp. 247-284.

KLEEMOLA K., HYYTINEN H., TOOM A. 2023. Critical thinking and writing in transition to higher education in Finland: do prior academic performance and socioeconomic background matter?, *European Journal of Higher Education*, Vol. 13, No. 4, pp. 488–508.

LIPS H.M. 2004. The Gender Gap in Possible Selves: Divergence of Academic Self-Views Among High School and University Students, *Sex Roles*, Vol. 50, No. 5/6, pp. 357–371.

MCGHIE V. 2017. Entering university studies: identifying enabling factors for a successful transition from school to university, *High Educ*, Vol. 73, No. 3, pp. 407–422.

TRIVENTI M. 2010. Something changes, something not. Long-term trends in gender segregation of fields of study in Italy, *Italian Journal of Sociology of Education*, Vol. 2, No. 2, pp. 47-80.

Valentina TOCCHIONI, University of Florence, valentina.tocchioni@unifi.it
Samuele MILONE, University of Florence, samuele.milone@unifi.it
Gabriele LOMBARDI, University of Florence, gabriele.lombardi@unifi.it

EXPLORING NATURAL EXPERIMENTS: GENDER OF OFFSPRING AND THE CHALLENGES OF THE STOPPING RULES[1]

Giambattista Salinari, Gianni Carboni, Virginia Zarulli

Abstract. The sex of children has frequently been used in the field of social sciences to conduct natural experiments. The key hypothesis behind this methodology is that the sex of the firstborn (or the first k births) is an exogenous variable, meaning it is not influenced by family characteristics observed before the birth of the children. Recent analyses have questioned the supposed exogeneity of the sex of children, proposing that stress experienced during pregnancy results in higher male embryo mortality, thereby leading to a higher probability of female births. This hypothesis casts doubt on the results reached by studies that have used the sex of children to conduct natural experiments. However, the analyses supporting this hypothesis have not properly considered the problems arising from stopping rules, specifically the tendency of some families to continue having children until a child of the desired sex is born. In this work, we show, using an indirect approach, that if stopping rules are properly taken into account, the sex of offspring is not associated to parental stress.

1. Introduction

Randomized experiments are generally considered the most reliable method for estimating a causal effect of interest (Imbens and Rubin 2015). The logic behind these experiments is very simple: a treatment is randomly administered to individuals in a study group. Individuals randomly selected to receive the treatment will form the treatment group. The remaining individuals will constitute the control group. The random assignment of the treatment ensures that the control group and the treatment group share, on average, the same pre-treatment characteristics, except for sampling errors. Thus, any observed differences in the outcome of interest between the control and the treatment group will reflect only the effect of the treatment itself, and not any pre-existing differences between the two groups.

In the field of social sciences, the possibility of conducting randomized experiments often clashes with ethical limits. For example, it is not possible to randomize

[1] Giambattista Salinari produced the research questions and the study design, conducted the analysis, provided methodological expertise for the study and drafted the first version of the manuscript. Gianni Carboni and Virginia Zarulli contributed significantly through providing overarching critical feedback, interpreting the results with insightful conceptual enhancements, editing and revising the final manuscript.

individuals based on different levels of education to then verify the effect of this variable, say, on fertility, because, obviously, one cannot impose on an individual the level of education they must have.

To overcome these difficulties, the social sciences have developed a set of techniques known as *natural experiments* (Dunning 2012). The basic idea is to use a "natural randomizer" to replace the random assignment of treatment used in randomized experiments. These "natural randomizers" allow us to study situations that would otherwise be hard to study because of the interconnection of two or more variables and the impossibility to detangle the causal effect. Today, numerous instances of this approach exist, with random variability often sourced from actual lotteries, as in the classic example of Angrist's (1990) study on the income of Vietnam War veterans and studies on migrations from Tonga (McKenzie et al. 2010) are such examples. The sex of children has also been considered as a random source of variability, for example in the study of the effect of fertility on female labor supply (Angrist and Evans 1998), the study of divorce and separation (Dahl and Moretti 2008) and the study of the effect of fertility on cognitive decline at old age (Bonsang and Skirbekk 2022).

Hamoudi and Nobles (2014), however, argued that the sex of children cannot be considered a random source of variability to be employed in natural experiments. To bolster their argument, they used biological considerations, claiming that stress during pregnancy increases male embryo mortality compared to females, thus suggesting that a lower probability of male births may occur in conflict-ridden families. To validate this hypothesis, these authors conducted analyses showing that the sex of the children was associated with the stress levels measured before birth.

With this work, we show that Hamoudi and Nobles (2014) might have underestimated the role played by the so-called stopping rules: those strategies by which parents continue to have children until the desired sex, or the desired combination of sexes is reached. Herein we demonstrate that in the presence of such strategies, the analysis proposed by Hamoudi and Nobles (2014) suffers from an insidious and little-known form of bias called *collider bias*.

2. Problems stemming from stopping rules

It is known that parents may have a preference regarding the sex of their children, like desiring boys or a mix of genders. When these preferences exist, parents may use stopping rules that consist of continuing to have children until they get the desired outcome. Such rules have been identified in many societies. Research by Angrist and Evans (1998) and Dahl and Moretti (2008) confirm this behaviour in the US population. For instance, Angrist and Evans, discovered that only 37.2% of

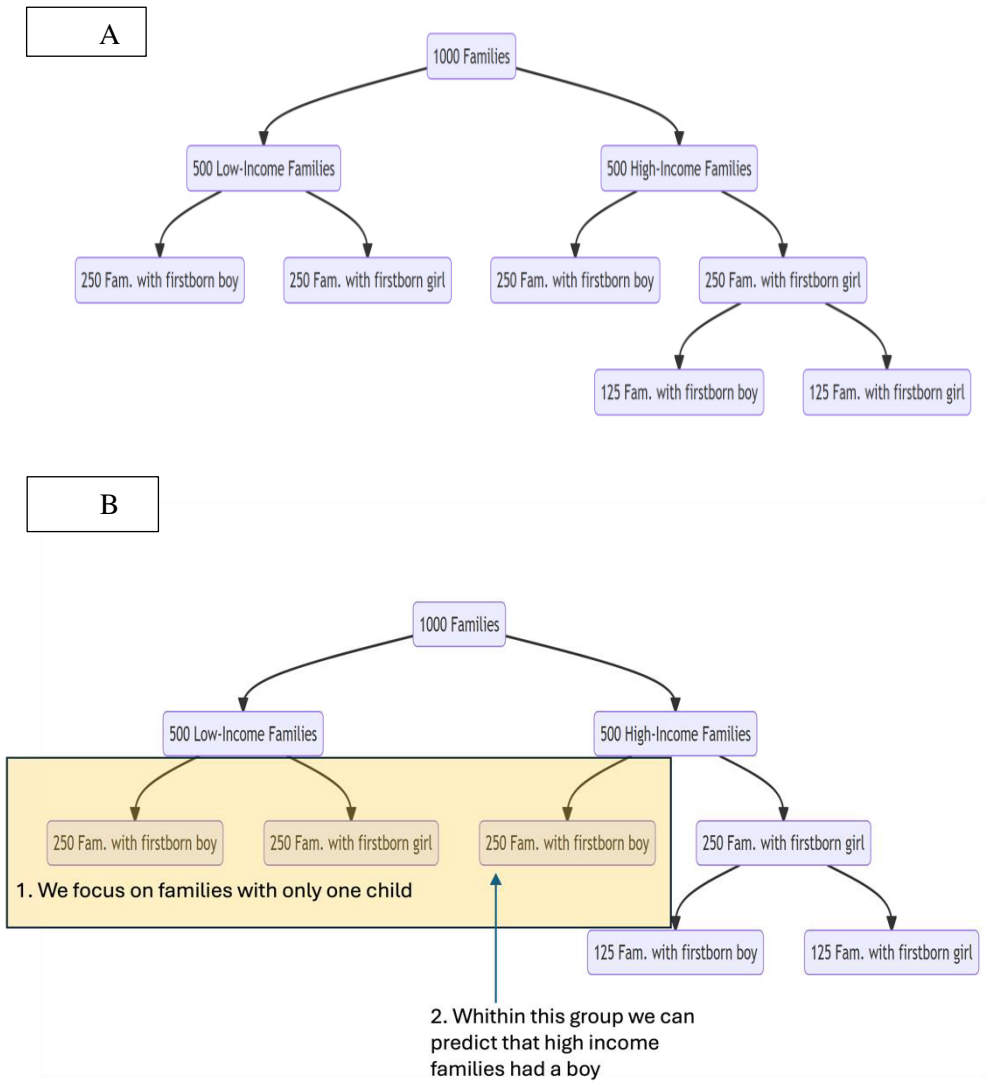
women with two children of different sexes go on to have a third child, compared to 43.2% of women who have two children of the same sex. In Denmark, it was found that families with two same-sex children had the highest fertility rates, those with two children of the same sex followed by a child of the opposite sex had the lowest rates, and that families with two boys showed a slight preference for girls, as they had higher fertility rates compared to families with two girls (Jacobsen et al. 1999).

One way to clarify the nature of the problems introduced by stopping rules is to resort to an example where the role of this type of rules is "extremized." Let's assume that the families of a given population conform to the following rule: a) all families have at least one child; b) all families with high income (or belonging to a certain ethnic group, or with low marital conflict, etc.) have an additional child if the firstborn is female; c) the probability of having a male child is 50%. Notice that nowhere in this rule it is stated that family income has a causal effect on the sex of children. We are thus assuming that the sex of a newborn and the family income are independent. Figure 1.A shows an example of this type of stopping rule, where it is assumed that there are 1000 families, divided into 500 high-income families and 500 low-income families. This example demonstrates that the stopping rule results in high-income families having a higher average number of children compared to low-income families. However, the sex ratio remains unaffected by the stopping rule. Among high-income families, there are 750 children, with an equal distribution of 350 females and 350 males, mirroring the distribution in low-income families, which have 250 males and 250 females. The fact that stopping rules do not alter the sex ratio has been mathematically proved by various articles (see Grech et al. 2018 for a summary). So, in general, we cannot predict the sex of a child based on the income of the family to which the child belongs. And this is consistent with the example assumption, according to which family income does not cause the sex of newborns.

Things change, however, if we condition (focus) the analysis on family size. This means that if we consider only those families with, say, only one child, then we can predict the sex of a newborn based on the information about family income.

To illustrate this point, let's consider families with only one child (see Figure 1.B). From this group, we select a family known to have a high income. We can infer that this family has a boy because, if they had a girl, they would have likely continued to have more children in pursuit of a male, given their high income. Therefore, conditioning on the number of children creates the impression that family income can predict whether a male child is present, even though there is no causal link between income and the sex of the children by design.

Figure 1 - *A hypothetical example of a stopping rule.*



Note: In this example, we suppose we observe 1000 families divided into 500 low-income families and 500 high-income families. These families conform to the following rules: a) all families have at least one child; b) all high-income families have an additional child if the firstborn is female; c) the probability of having a male child is 50%.

The fact that it seems possible to predict the sex of a child based on the family's income in the presence of stopping rules is a statistical artifact known as collider bias. In general, given three variables A, B, and C, B is said to be a collider (or collision node) when both A and C have a causal effect on B, i.e.:

$$A \rightarrow B \leftarrow C$$

In this simple causal graph,[2] it is possible to prove through a criterion called *d-separation* (Pearl 2009) that A and C are marginally independent. At the population level, the information about the value of A does not allow predicting the value of C. And this is correct, because it is evident that there is no causal effect of A on C or vice versa. With the d-separation criterion, it is also possible to show that, if the analysis is conditioned on B (for example, by selecting a subpopulation characterized by a specific value of B), within this subpopulation, A and C will be associated. Within the subgroups defined by the values of B, it is thus possible to predict the value of C from the knowledge of the value of A, even though A is not causing C or vice versa. A and C are thus linked by a spurious association (correlation), not produced by the existence of a causal connection.

In our example of Figure 1, the reason why it seems possible to predict the sex of a child based on family income, in the presence of stopping rules, is closely related to what we have just discussed. Let's denote B as the family size. In our example, B is determined (caused) by the sex of the first child (A) and the family's income level (C). When we condition the analysis on family size, a spurious association between the sex of the first child and family income is triggered.

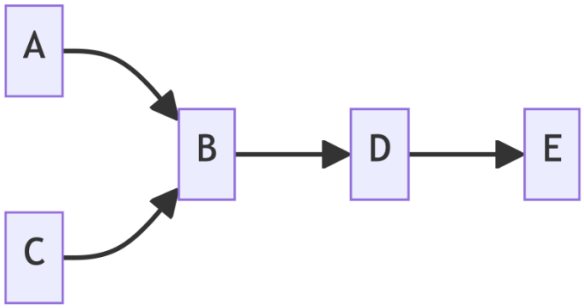
In reality, the conditions leading to the formation of collider bias are more complex than what discussed so far. Here it will suffice to notice that, in general, collider bias will occur whenever the analysis is conditioned on a collider or on a descendant of a collider, a variable that is directly or indirectly caused (through the mediation of other variables) by a collider. Referring to the causal diagram in Figure 2, we can say that a spurious association (bias) between A and C will emerge if the analysis is conditioned on B, or on D, or on E.

In our example, if the analysis is conditioned on any variable that is directly or indirectly influenced by family size, this will produce a spurious correlation between family income level and the sex of the newborn.

The existence of stopping rules can thus complicate the identification of the determinants of children's sex.

[2] In the context of causal graph theory, an arrow (→) indicates a direct causal effect, nodes represent variables, direction shows causality flow, and acyclicity ensures there are no loops.

Figure 2 – *Collider bias.*



3. The association between marital stress and the sex of children

Hamoudi and Nobles (2014) argued that marital stress experienced during pregnancy can lead to higher mortality for male embryos compared to female embryos, thereby altering the sex ratio at birth. The authors use panel data from the National Longitudinal Study of Youth (1979) that reports the level of marital stress in 1992 and estimate a linear probability model of the following kind:

$$E(Y|x,w) = \alpha + \beta X + \gamma W, \qquad (1)$$

where Y is a binary variable representing the birth of a male child after 1992, X measures marital stress, and W is the total number of children had after 1992. The authors justify the use of a linear probability model stating that the choice of the functional form (e.g. logit, probit, etc.) does not affect the result of their analysis. The authors include additional "control" variables in their model, such as birth order and age at marriage, but for the moment, we ignore these additional variables, because what is relevant here is the effect produced by the inclusion in the model of W.

We can first suppose that the variable X (marital stress) is a determinant of family size (W). For example, in the analysis conducted by Mencarini and Tanturri (2006) on a sample of Italian women, family relationship instability is cited as a reason for not having (another) child by 34% of childless women and by 24% of women with only one child. Depending on whether we consider women with parity 0 or 1, this is the second and third most important reason given for not having (more) children.

We can also suppose that the sex of a child is a determinant of family size if there are groups within the population that are applying stopping rules.

Finally, if we assume that marital stress does not influence the sex of the children (the null hypothesis within the test being conducted), we conclude that the causal diagram representing the causal connections between the variables in model (1) is:

$$Y \rightarrow W \leftarrow X.$$

In this diagram, the variable W turns out to be a collider. The fact that this collider has been included as a control in the model is equivalent to conditioning the analysis on this variable. But as we discussed in the previous paragraph, conditioning the analysis on a collider triggers a spurious correlation between X and Y. This means that we do not know if the significant correlation found between marital stress and the sex of the child resulted from the biological process imagined by Hamoudi and Nobles (2014) or from the spurious correlation introduced by the collider bias. Additionally, note how the inclusion of other variables worsens the situation. The variable birth order is clearly a descendant of family size (for example, if W=1, then birth order will necessarily be 1). Thus, the model is also conditioning the analysis on a descendant of a collider.

The problem with model (1) lies in the inclusion of a bad control, namely W. Cinelli et al. (2022) recently published a list of the most frequently encountered bad controls in empirical analyses, showing the solutions to adopt to avoid distortions in the results. In the case considered here, regarding the relationship between stress and the sex of newborns, the solution is simply to remove the bad controls from the model.

4. Testing the association between stress and the sex of children

Verifying the link between stress during pregnancy and the sex of newborns is not straightforward, primarily due to the lack of adequate data. When a healthy individual faces a sudden challenge, such as a threat to their physical integrity, or during a natural disaster like an earthquake, there is an activation of their hypothalamus-pituitary-adrenal (HPA) axis resulting in the production of cortisol (the stress hormone). The concentrations of this hormone in blood, saliva, or even hair can indicate the level of stress in an individual. However, the samples of individuals for whom these measures are taken and for whom we know the sex of their offspring are generally small, making it challenging to identify the effect of cortisol on the sex of births. If such an effect exists, it is considered small and therefore requires large samples to be detected (Hamoudy and Nobles 2014).

To address this issue, we decided to test the effect of stress on the sex of offspring indirectly. It is known that exposure to adverse conditions during childhood can lead to chronic stress conditions during adulthood, when cortisol levels remain elevated even in the absence of external threats or challenges. Conditions most commonly

associated with chronic stress include financial hardship during infancy, poor quality of the relationship with the parents, absence of one or both parents, parental union dissolution, and experiences of parental violence (Lupien et al., 2009; Miller et al., 2007; Sassler et al., 2009). Miller et al. (2007) further analyse how the chronic stress induced by these adverse childhood and youth conditions is associated with a range of negative outcomes in adulthood, including heavy alcohol use, drug use, sexual risk-taking, anxiety, depression, interpersonal violence, and self-oriented violence.

We then used retrospective information collected during the seventh wave (2017) of the Survey on Health Ageing and Retirement in Europe (SHARE) to obtain information on the conditions experienced by the respondents during childhood and the sex of their children throughout their lives. Additionally, since this survey focuses on the population aged 50 and over, we have information on the entire reproductive history of these women.

We proceeded in two different phases. Firstly, we sought to verify if the conditions experienced during childhood were indeed significantly associated with depression at the time of the interview. This preliminary analysis serves to confirm that the hypothesized association between adverse childhood conditions, chronic stress, and depression (explicitly formulated in Hamoudi and Nobles 2014) was also supported in our data. To perform this test, we applied a chi-square test. First, we considered experiencing financial hardship during childhood. (the data are described in Table 1).

Table 1 – *Financial hardship suffered by a woman during childhood and the likelihood of depression.*

	Financial hardship during childhood	
Depression	Yes	No
1-2	129	4444
3-4	112	3302
5-6	61	1662
7-8	33	688
9-10	20	239

Note: data derived from wave 7 of SHARE.

The chi-squared statistic for this table is 23 with 4 degrees of freedom, which leads to a rejection of the null hypothesis of independence between these two variables (p-value < 0.001). The average score on the depression severity scale is 3.35 for individuals who did not experience financial difficulties during childhood, and 3.91 for those who did. Therefore, financial hardships experienced during childhood appear to have a long-term association with depression, consistent with the hypothesis that financial hardship during childhood is associated with a condition

of chronic stress. We can thus conclude that at least a fraction of people who experienced financial hardship during childhood have developed a chronic stress condition.

We should then expect that the sex distribution of births in this subpopulation is skewed in favour of female births.

Table 2 – *Financial hardship suffered by a woman during childhood and the sex of her children.*

	Financial hardship during childhood	
Sex	Yes	No
Male	184	5413
Female	177	5044

Note: data derived from wave 7 of SHARE.

We can, again, test this hypothesis by applying a second chi-square test of independence to the data shown in Table 2. This time, the value of the chi-squared statistic is 0.59 with only one degree of freedom. This test does not reject the null hypothesis of independence (p-value = 0.81). So, experiencing financial hardship during childhood does not seem to produce any detectable effect on the sex distribution of children born to a woman. Note that in testing the association between financial hardship experienced by the mothers and the sex distribution of the children, we have not controlled for the existence of potential confounders. This is for several reasons. First, the (null) hypothesis that we are considering is that the sex distribution of the children is an exogenous variable (because this is the role given to this variable in the natural experiments that we cited in the introduction). As such, the sex distribution of the children, being a pure source of random variability, is supposed to not be affected by any variable. Similarly, it cannot be confounded by any other variable, because a confounder, by definition, is a variable that affects simultaneously the treatment, in this case stress, and the outcome, in this case the sex of the child. Second, by avoiding conditioning the analysis according to the value of other variables, we protect our results from a collider bias. As we have seen in the two previous sections, it suffices to include in a model a variable which is indirectly affected by family size to trigger such a bias. Since we do not really know the complete set of variables indirectly affected by family size, the safest approach is to avoid conditioning the analysis on any other variables. Based on these considerations, we thus believe that the approach based on a simple chi-squared test is also the safest way to test the association between chronic stress and the sex distribution of children.

Table 3 – *Testing the association between adverse condition experienced during childhood with depression (Test 1) and the sex distribution of children (Test 2).*

	Test 1		Test 2	
Adverse conditions experienced during childhood	**Sample size**	**Depression**	**Sample size**	**Sex imbalance**
Father no job	10667	yes	10667	no
Lived without mother at 10	11228	yes	11228	no
Lived without father at 10	11228	yes	11228	no
Mother understand worries	11017	yes	11017	no
Father understand worries	10515	yes	10515	no
Relationship with mothers	11070	yes	11070	no
Relationship with father	10575	yes	10575	no
Mother physical harm	11070	yes	11070	no
Father physical harm	10643	yes	10643	no

Note: The first three adverse conditions experienced during childhood were coded as binary variables. For the next six conditions, the variables are ordinal and represent a scale ranging from 1 to 5. In the two columns labelled "Depression" and "Sex imbalance," "yes" and "no" indicate whether the null hypothesis of independence was rejected (yes) or not (no).

We repeated the two types of tests shown earlier for nine additional adverse conditions experienced during childhood. The results of these tests are shown in Table 3. For all the adverse conditions examined, an association with a state of depression in adulthood was identified (the women considered were interviewed at the age of 50 or older). The data presented here thus seem to confirm the causal pathway that links adverse conditions experienced during childhood and adolescence to the onset of a chronic stress condition and a depressive state during adulthood. Nevertheless, when comparing the sex distribution of the children born to women who suffered adverse conditions during childhood with those who did not suffer from these conditions, we find that these distributions are indistinguishable from each other. Therefore, we do not find evidence of the causal pathway proposed by Hamoudi and Nobles (2014), which suggests that adverse conditions experienced during childhood lead to a chronic stress condition that, in turn, would lead to an imbalance in the sex distribution of offspring.

5. Conclusions

In their analysis, Hamoudi and Nobles (2014) argued that the sex of newborns is influenced by maternal stress levels and, therefore, cannot be considered an exogenous source of variability. This claim challenges a long-standing tradition of

research (Angrist and Evans 1998; Dahl and Moretti 2008; Bonsang and Skirbekk 2022) that has used the sex of children to conduct natural experiments. In this paper, we aim to assess the robustness of Hamoudi and Nobles' claim from both a theoretical and empirical perspective.

From a theoretical point of view, we first aimed to demonstrate how stopping rules can produce subtle distortions. We showed how such behaviours can create the illusion that variables such as income or stress determines the sex of children. This phenomenon occurs when the analysis is conditioned on family size or any variable that is directly or indirectly influenced by family size. It is thus possible that the innovative analysis conducted by Hamoudi and Nobles might suffer from this type of problem. However, it should be noted that this is not necessarily the case. The bias might still be small and not significantly alter the results of the analysis. For this reason, we believe this analysis should be repeated by removing bad controls from the model.

From an empirical point of view, we employed an indirect method to verify the effect of stress on the sex ratio of births. Hamoudi and Nobles (2014) assert, based on extensive literature, that adverse conditions experienced during childhood and adolescence increase the likelihood of suffering from chronic stress. Therefore, we sought to verify the existence of an association between a set of these adverse conditions and the sex ratio of children for a sample of approximately 10,000 women. However, our analyses were unable to identify any sex imbalance among women who suffered from these conditions. It is possible that this is due to the (indirect) method used, or the fact that the sample size is still too small. Whatever the reason, the hypothesis of an association between stress and the sex ratio of births seems to require further research work.

References

ANGRIST J.D. 1990. Lifetime Earnings and Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records. *The American Economic Review*, Vol. 80, No. 3, pp.313-336.

ANGRIST J.D., EVANS W.N. 1998. Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review,* Vol. 88, No. 3, pp. 450-477.

BONSANG E., SKIRBEKK V. 2022. Does Childbearing Affect Cognitive Health in Later Life? Evidence From an Instrumental Variable Approach. *Demography*, Vol. 59, pp. 975-994.

CINELLI C., FORNEY A., PEARL J. 2022. A Crash Course in Good and Bad Controls. *Sociological Methods & Research* https://doi.org/10.1177/00491241221099552.

DAHL G.B., MORETTI E. 2008. The demand for sons. *The review of economic studies*, Vol. 75, pp. 1085-1120.

DUNNING T. 2012. *Natural experiments in the social sciences: a design-based approach*, Cambridge University Press.

GRECH V., JAMESB W.H., LAURI J. 2018. On stopping rules and the sex ratio at birth. *Early Human Development*, Vol. 127, pp.15-20.

HAMOUDI A., NOBLES J. 2014. Do Daughters Really Cause Divorce? Stress, Pregnancy, and Family Composition. *Demography,* Vol. 51, No. 4, pp. 1423-49.

IMBENS G.W., RUBIN D.B. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences. An Introduction*. New York, Cambridge University Press.

JACOBSEN R., MØLLER H., ENGHOLM G. 1999. Fertility rates in Denmark in relation to the sexes of preceding children in the family, Human Reproduction Vol. 14, No. 4, pp.1127-1130.

MENCARINI L., TANTURRI M.L. (2006). High Fertility or Childlessness: Micro-Level Determinants of Reproductive Behaviour in Italy. *Population,* Vol. 61, No. 4, pp. 389-415.

MCKENZIE D., GIBSON J., STILLMAN S. 2011. How Important is Selection. Experimental vs. Non-Experimental Measures of the Income Gains from Migration. *Journal of the European Economic Association*, Vol. 8, No. 4, pp. 913–945.

PEARL J. 2009. *Causality. Models, Reasoning and Inference*. Cambridge, Cambridge University Press.

Giambattista SALINARI, University of Sassari, gsalinari@uniss.it
Gianni CARBONI, University of Sassari, gcarboni3@uniss.it
Virginia ZARULLI, University of Padua, virginia.zarulli@unipd.it

COUPLES' FERTILITY DIFFERENTIALS BY EDUCATION: DO STEPCHILDREN MAKE A DIFFERENCE? [1]

Alessandra Trimarchi, Laurent Toulemon

Abstract. A growing number of studies has explored both partners' education as determinant of couples' fertility, acknowledging the fact that the decision to have a child is couple based. Still, those studies have solely focused on children born to the couple, without considering stepchildren. As a result, in studying couples' birth rates by educational pairing, previous studies did not account for the complexity of family composition, which also affects partners' decision to have a common child. In this paper, we aim at tackling family complexity and its association with education. Using Generations and Gender Surveys (GGS) data of 14 European countries, we analyse the association between educational pairing and couples' fertility based on different definitions of couples' children. Applying standard fertility analysis, overall results show a decline in childlessness among younger cohorts when stepchildren are considered, with strong educational difference. We found that among the younger cohorts, highly educated homogamous couples have less often stepchildren (born from one partner before the union) and remain less often without shared children. Stepchildren, instead, are more common among low educated couples, and among the heterogamous couples. We also found diversity among heterogamous couples: there are fewer stepchildren when one partner is highly educated, stepchildren more often come from the woman, especially when she is low educated.

1. Introduction

The diffusion of family behaviours such as divorce and cohabitation, and ensuing constitution of single-parent families or stepfamilies represents an important societal change in Europe, taking place over the second half of the 20th century (Thomson 2014). The constitution of stepfamilies is mainly due to the diffusion of divorce and separation of couples with children, but also re-partnering, since now it is more common to partner with someone who has already children (Goldscheider *et al.* 2009). The incidence of union dissolution, re-partnering, and fertility is not equal across social strata (Perelli-Harris *et al.* 2010; Trimarchi and Van Bavel 2018). According to the "*Diverging Destinies*" thesis, the diffusion of these family

[1] Authors have jointly contributed to the article. A.T. has drafted sections 1, 2, 3. L.T. has drafted sections 4, and 5. Both authors have edited and revised the whole paper.

behaviours may contribute to an increase in social inequalities because they tend to be more common among the lower strata of society (McLanahan 2004; McLanahan and Percheski 2008). The lower educated are more likely to divorce and separate, but also to have children more quickly within the union relative to the highly educated that instead tend to postpone and then eventually end their union before having children.

The way social inequality is linked to these family behaviours, however, substantially depends on fertility behaviours of those individuals who have experienced partnership disruption. This implies that the educational pairing of those who form stepfamilies also varies, and it affects the overall fertility behaviour of these couples. Recently, a growing number of studies has explored both partners' educational characteristics as potential determinants of couples' fertility behaviour, acknowledging the fact that the decision to have a(nother) child is couple based (Osiewalska 2017; Nitsche *et al.* 2018; Trimarchi and Van Bavel 2020).

Still, these studies have especially focused on children born to the couple, without considering stepchildren in their counts. As a result, previous studies do not account for the complexity and variety of family composition, which also affects partners' decision to have a common child (Toulemon 2014). This is an important gap, since considering or not stepchildren among couples' children may affect the educational gradient in couples' fertility and childlessness. We aim to fill this gap, by specifically asking to what extent a different definition of couples' fertility, which also considers the number of stepchildren present in the union, may lead to different educational gradients in fertility over two different cohort groups. We examine standard fertility indicators, i.e., mean number of children and proportion childless, using the first round of Generations and Gender Survey (GGS) data of 14 European countries.

2. Background and hypotheses

2.1. Education and stepfamilies

In many European countries, previous studies have found that in more recent cohorts, the lower educated are more likely to divorce and separate (Harkonen and Dronkers 2006; Matysiak *et al.* 2014). This change over time has been strongest among low educated women, who, in the past, were less likely to divorce. Educational and gender differentials also occur in repartnering rates (Raley and Sweeney 2020). Substantial research has shown that mothers are less likely to repartner than fathers (e.g., Di Nallo 2019), even if the fact that repartnering is more frequent among fathers does not necessarily affect gender differences in completed

cohort fertility, controlling or not for the level of education (Andersson 2023). Previous research on fertility in higher order unions focusing on a couple level analysis is still rare, while abundant research addresses women's (or men) fertility in higher order unions (e.g., Kreyenfeld *et al.* 2017; Jalovaara and Kreyenfeld 2020).

A low level of education has been associated with a faster rate of repartnering among women, which is interpreted as a way to overcome economic strains (McNamee and Raley 2011; Shafer and James 2013). Beyond the timing of repartnering, education may also affect preferences for new partners (Raley and Sweeney 2020). In Belgium, for instance, highly educated men were found more likely to be in the role of biological father rather than stepfather, which is, instead, more common among the less educated (Schnor *et al.* 2017). In sum, the constitution of stepfamilies, and the transition to "stepfamily-fertility" is the outcome of several processes (union formation, childbearing, union dissolution, and repartnering), which all depend on age and on education, at the individual level, but also at the contextual level, according to variations on the mating market (Van Bavel 2021).

2.2. Hypotheses

Differently from previous research, here we focus on a couple-approach. Hence, we formulate hypotheses distinguishing between homogamous (partners having the same level of education) and heterogamous couples (partners having a different level), focusing on two large birth cohorts. First, based on previous findings, we expect that among the younger cohorts, highly educated homogamous partners remain less often without any common child and have less often stepchildren relatively to their less-educated counterpart (*H1*). In other words, we expect that highly educated homogamous couples less often have stepchildren (born from one partner before the union), and remain less often childless relatively to other pairings, when the indicator of childlessness only considers shared children.

Regarding heterogamous couples, we expect that, among the younger cohorts, the presence of a highly educated partner (man or woman) decreases the chances of having stepchildren (*H2*). Finally, we expect that partners living in educationally heterogamous couples are more likely to have stepchildren than partners living in highly educated homogamous couples (*H3*). This is going to be especially the case for heterogamous couples formed by low educated partners, since the partner with lower education, male or female, is more likely to bring children within the new union.

3. Data and analytical strategy

To answer our research question, we used GGS data of 14 countries (Belgium, Bulgaria, Czech Republic, Estonia, France, Georgia, Hungary, Italy, Lithuania, The Netherlands, Norway, Poland, Romania, and Sweden)[2]. The countries considered are very different in terms of their fertility level, prevalence of separations, divorces and second unions. For instance, Northern and Western European countries have on average higher levels of fertility, in combination with higher prevalence of separations and second unions, relatively to Central and Eastern European countries where fertility levels are lower, and divorces and separations less common. Still, our analysis is descriptive and due to small sample sizes, we cannot analyse countries' separately.

GGS are the most recent comparable cross-country data with available information about stepchildren. In GGS data, information on both partners' education and stepchildren is available if the respondent is coresiding with the partner. Thus, we kept in our sample only respondents coresiding with a partner at the time of interview. To make country-data homogenous, we chose an age-criterion, and we selected couples where the woman at the time of interview is between 38-45 years old, or 55-65 years old. Additionally, since we are studying fertility, we only considered heterosexual couples where the woman was younger than 46 years old at the time of union formation. Since GGS surveys took place in different years for different countries, in this study we focus on the comparison of two groups of couples, making sure that each group involves cohorts including the same pool of countries. The first group, the "old", comprises couples where the woman was born between 1940-1958 (N = 18,460), aged 55-65 years old at interview. The second one, the "young", comprises couples where the woman was born between 1960-1975 (N = 17,006), aged 38-45 years old.

Our main variable characterizing the couple is educational pairing. We have grouped respondents and their partners into three levels of education (low, medium, high) collapsing categories from the International Standard Classification of Education (ISCED 1997). The first group includes those who completed primary plus lower secondary school (ISCED 0, 1, and 2). The medium category consists of respondents who completed the upper-secondary and a post-secondary level (ISCED 3 and 4). Finally, highly educated respondents are those holding a bachelor/master/PhD degree (ISCED 5 and 6). The variable of educational pairing has seven categories: three categories for homogamous unions where both partners have the same level of education (both low, both medium, both high); two categories for female hypergamy (i.e., the man is more educated that the woman) and two

[2] The data collection took place between 2002 and 2013, see country-data details in Appendix, Table A1.

categories for female hypogamy (i.e., the woman is more educated than the man). For the categories of heterogamous couples, we distinguish between couples with a highly educated partner (the other having a low or medium education) and couples with only low-medium educated partners. Table A1 and A2 in Appendix show sample size by country and educational pairings for each group of cohorts.

Our dependent variables consist in the mean number of children and the proportion of childless couples. We specify the mean number of children in: (1) mean number of *all* children in the couple, including stepchildren; (2) mean number of *common* (or *shared*) children; (3) mean number of children from previous partnerships, distinguishing in children brought from the woman (3a), and children brought from the man (3b). Next, we specify the proportion childless in: (1) couples without any children at all, i.e., also without children from previous relationships; (2) proportion of couples without any common (or shared) child.

In short, our aim is to describe the growing importance of stepchildren and the growing complexity of family links in terms of couples' lives. We describe trends in "couples' fertility", considering stepchildren, we do not analyse these trends by applying explanatory models. In the following result section, we describe the relationships between our independent variables (i.e., women's birth cohort and educational pairing), and the dependent variables (i.e., mean number of children, and proportion childless), using different definitions of children, without including any additional control variable.

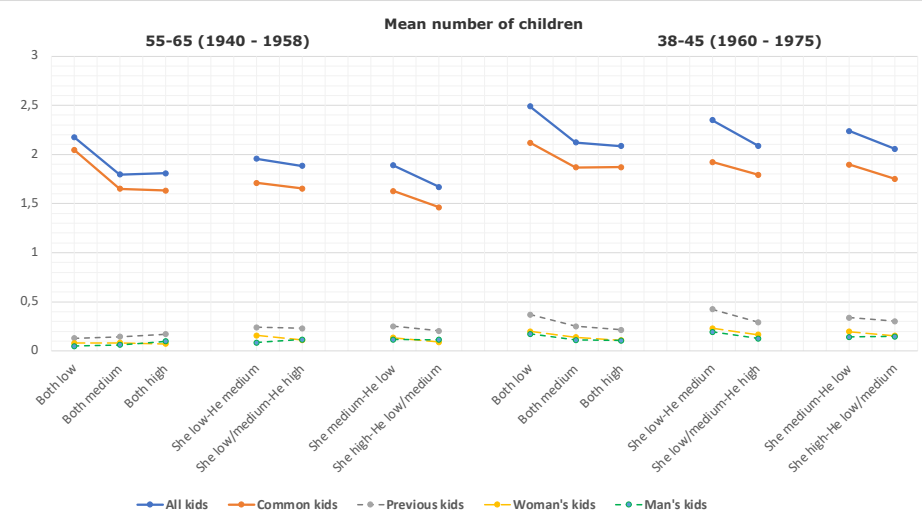
4. Results

From the older to the younger cohorts, fertility increased among all groups of couples (Figure 1, cohorts born in the 1940s and 1950s, left panel, compared to cohorts born in the 1960s and 1970s, right panel). Among homogamous couples, the low educated have more children than couples with medium or high education: 2.0 common children vs. 1.6 in the older cohorts; 2.1 vs. 1.9 in the younger cohorts. Fertility is also lower in heterogamous couples (hypergamous and hypogamous) with a partner having a high education: 1.8 vs. 1.9 in the recent cohorts. In the older cohorts, hypogamous couples had the lowest fertility: 1.6 and 1.5 common children respectively for women with a medium or high education. These patterns remain the same independently on the outcome we consider: the total number of children, including stepchildren, or counting only common children.

The mean number of stepchildren is increasing, especially for couples with low educated women, i.e., homogamous low educated couples, and hypergamous couples. For couples where the woman has a medium or a high education, higher fertility is mainly due to an increase in the number of common children.

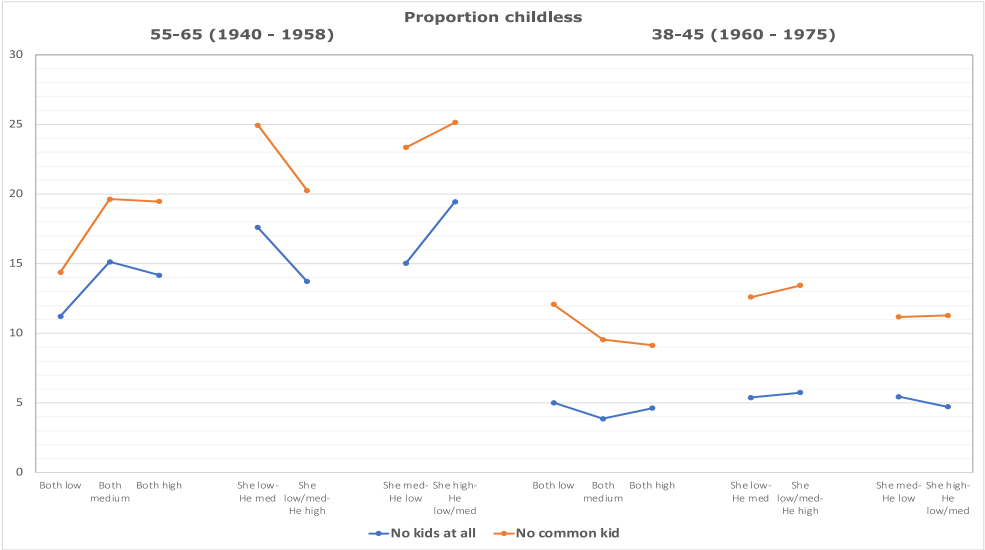
The picture is more complex when we look at childlessness among couples (Figure 2). In the older cohorts born in the 1940s and 1950s, childlessness (having no common child) was rare among low educated homogamous couples (14%), but more frequent among medium or highly educated couples (20%), and even more among heterogamous couples. When all children are considered, the proportion of childless couples is lower: 11% among low educated couples, between 14% and 20% among other couples, hypogamous couples with a highly educated woman is the group where "complete childlessness", i.e., also accounting for children born from previous unions, was the highest (19%).

Figure 1 – *Mean number of children by women's birth cohorts and educational pairings.*



Source: Own elaborations on GGS data (weighted).

Figure 2 – *Proportion childless, by women's birth cohorts and educational pairings.*

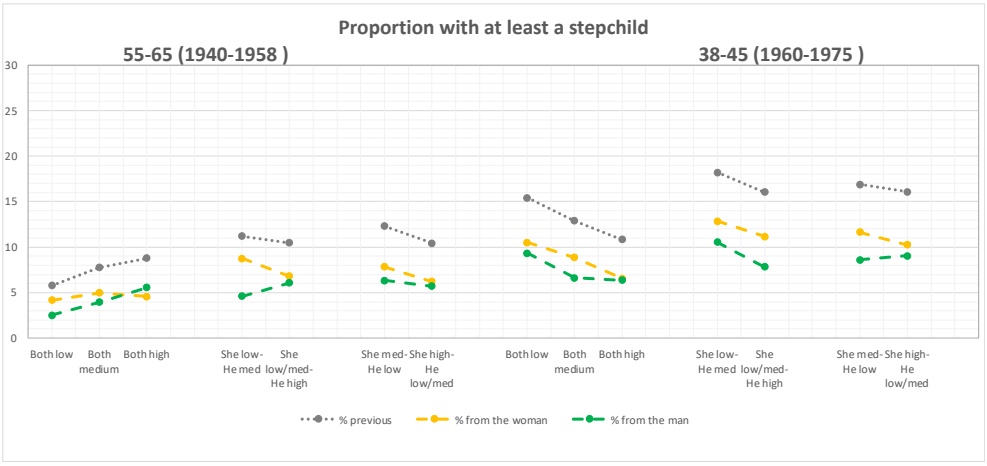


Source: Own elaborations on GGS data (weighted).

In the younger cohorts, when all children are considered, the proportion of childless couples has become very similar (about 5%) across all educational pairings. Medium and highly educated homogamous couples more often have at least one common child, while low educated couples, as well as heterogamous couples, especially hypergamous couples, remain more often without any common child, 13%, vs. 11% of hypogamous couples, and 12% for low educated homogamous couples.

The decline in childlessness comes from two different trends. First, the proportion of couples without any common child has dramatically decreased, for all groups except for low educated homogamous couples. Childlessness was less frequent in this category of couples; in the recent cohort, the relation has reversed, and childlessness has become most frequent for low educated homogamous couples. Second, the proportion of couples with at least one stepchild (a child born to one partner, with another parent, before the current union) has increased from 10% to 15% (Figure 3). The increase is by far the largest among low educated homogamous couples: +10%. This is probably because these couples make the transition to the first child rapidly after union formation, and most union disruptions involve at least one child; second unions are thus formed with a stepchild (Figure 3). Couples declare more often a stepchild, when the child is brought by the woman, but this could be due to omissions of children born to the man, because men less often coreside with their children born from a previous union.

Figure 3 – *Proportion of couples with at least a stepchild, by women's birth cohorts and educational pairings.*



Source: Own elaborations on GGS data (weighted).

5. Discussion and conclusions

The diffusion of stepfamilies calls for new ways to define fertility behaviour and analysing educational differentials in fertility. In this study, we aimed at analysing changes over time in the association between educational pairings and couples' fertility, accounting for stepchildren. We used GGS pooled data of 14 countries and examined differences among older (women born between 1940-1958) and younger (women born between 1960-1975) cohorts in couples' mean number of children and the proportion childless. Overall, we found that couples' fertility has increased of around 0.3 children, one third of this increase being attributable to stepchildren.

In line with our first hypothesis (*H1*), according to which highly educated homogamous couples less often have stepchildren, we found that the increase in the number of stepchildren is largest among couples where the woman has a low level of education, leading to a reversal in the educational gradient in stepparenting among homogamous couples. This is because union disruptions more often involve couples with children among the low educated, given that the educational gradient of union disruptions and second unions has reversed (Matysiak *et al.* 2014).

Stepfamilies are thus becoming more common among couples, especially among couples where the partners have a low level of education, a finding in line with our second hypothesis (*H2*), which focused on differences between heterogamous couples. When looking at differences between heterogamous couples and highly educated homogamous couples, in line with our third hypothesis (*H3*), we find that

the former have more often stepchildren than the latter. We did not find major differences according to the sex of the parent who had a child before the union, the overall level of education within the couple matters the most in stepfamily formation rather than the difference between partners' education. Considering only stepchildren living with the couple (or who had lived with the couple in the past), however, we found larger gender differences: most children live with their mother after a parental disruption and, in case of a stepfamily, children are most likely to live with a stepfather than with a stepmother. A finding which is probably related to gender differences in the declaration of children born from previous unions.

The constitution of stepfamilies is the consequence of a series of events, i.e., couples formed by partners who had already lived as a couple and had a child in a previous union. Fertility among these stepfamilies is related to the will to have at least a common child (Thomson et al. 1990). The mechanisms leading to differential "stepfertility" and its impact on family size are thus complex. The increase in couples' number of children over time is mainly due to common children for women with a medium or high education, and to stepchildren for couples where the woman has a low level of education. The recent decline in fertility can be put into perspective when we consider couples' fertility and include stepchildren, who have a stepparent in addition to their biological parents. The type of stepchild – stepparent relationship may differ with duration of coresidence, with the stability of the new couple, and (step)family relationships may last or not in the long term.

The current study has important limitations since the causal chain of mechanisms leading to the formation of a stepfamily were not investigated deeply, nor country-level variation could be analysed. Still, from this exploratory study emerges the importance to consider both stepchildren and common children when studying couples' fertility differentials by education. Family relations in stepfamilies and their cross-country variation remain a relevant avenue of research for the future, which is feasible thanks to the availability of harmonized survey data including questions about full partnerships and fertility histories.

Appendix

Table A1 – *Sample description by country and educational pairings (unweighted counts), cohorts 1940-1958, age-group 55-65 (survey year in parenthesis).*

Country	Both low	Both med	Both high	She low/med-He high	She low-He med	She high-He low/med	She med-He low	Mis-sing	Total
Bulgaria (2004)	377	356	167	74	77	91	87	5	1234
Georgia (2006)	69	388	179	85	55	57	68	2	903
France (2005)	301	182	119	111	213	46	116	10	1098

Table A1 (cont.) - *Sample description by country and educational pairings (unweighted counts), cohorts 1940-1958, age-group 55-65 (survey year in parenthesis).*

Hungary (2004-2005)	280	564	146	156	293	82	63	1	1585
Italy (2003)	920	115	37	56	126	34	73	0	1361
Netherlands (2002-2004)	231	39	113	169	123	19	27	62	783
Romania (2005)	680	382	84	72	282	27	32	0	1559
Norway (2007-2008)	126	579	345	248	201	175	177	73	1924
Estonia (2004-2005)	94	186	127	86	59	117	125	0	794
Belgium (2008-2010)	260	101	185	140	106	81	69	14	956
Lithuania (2006)	169	354	101	63	37	82	106	1	913
Poland (2010-2011)	360	1577	224	190	320	146	227	19	3063
Czech Rep. (2005)	81	548	58	107	148	40	37	37	1056
Sweden (2012-2013)	56	348	221	113	82	219	142	50	1231
Total	4004	5719	2106	1670	2122	1216	1349	274	18460

Table A2 – *Sample description by country and educational pairings (unweighted counts), cohorts 1960-1975, age-group 38-45.*

Country	Both low	Both med	Both high	She low/med-He high	She low-He med	She high-He low/med	She med-He low	Missing	Total
Bulgaria	219	723	208	67	92	172	82	5	1568
Georgia	15	728	229	138	42	113	31	0	1296
France	119	339	213	87	123	125	94	10	1110
Hungary	80	639	133	70	125	115	42	0	1204
Italy	384	298	64	82	132	69	165	0	1194
Netherlands	141	144	167	141	64	63	81	59	860
Romania	154	683	111	58	204	46	40	1	1297
Norway	74	477	485	173	164	346	118	184	2021
Estonia	11	352	124	59	17	172	47	0	782
Belgium	87	166	275	74	59	130	79	23	893
Lithuania	23	592	137	83	19	138	55	3	1050
Poland	38	987	193	79	80	211	71	16	1675
Czech Rep.	27	524	77	92	60	48	33	32	893
Sweden	9	355	282	116	32	265	48	56	1163
Total	1381	7007	2698	1319	1213	2013	986	389	17006

Acknowledgements

This paper uses data from the GGS wave 1 (DOIs: 10.17026/dans-z5z-xn8g, 10.17026/dans-xm6-a262), see Gauthier, A. H. et al. (2018) or visit the GGP website (https://www.ggp-i.org/) for methodological details. Laurent Toulemon benefited from funding by the *Agence Nationale de la Recherche* under the *France 2030* plan, with reference ANR-21-ESRE-0037.

References

ANDERSSON, L., 2023. The Role of Gender Differences in Partnering and Re-partnering for Gender Differences in Completed Fertility. *Population Research and Policy Review*, Vol. 42, No. 2, p. 17.

DI NALLO, A., 2019. Gender Gap in Repartnering: The Role of Parental Status and Custodial Arrangements. *Journal of Marriage and Family*, Vol. 81, No. 1, pp. 59–78.

GAUTHIER, A., CABAÇO, S., EMERY, T., 2018. Generations and gender survey study profile. *Longitudinal and Life Course Studies*, Vol. 9, No. 4, pp. 456–465.

GOLDSCHEIDER, F., KAUFMAN, G., SASSLER, S., 2009. Navigating the "new" marriage market. *Journal of Family Issues*, Vol. 30, No. 6, pp. 719–737.

HARKONEN, J., DRONKERS, J., 2006. Stability and Change in the Educational Gradient of Divorce. A Comparison of Seventeen Countries. *European Sociological Review*, Vol. 22, No. 5, pp. 501–517.

JALOVAARA, M., KREYENFELD, M., 2020. Childbearing Across Partnerships in Finland and Germany. *Divorce in Europe: New Insights in Trends, Causes and Consequences of Relation Break-ups*, Vol. 21, pp. 315–336.

KREYENFELD, M., GEISLER, E., CASTRO MARTÍN, T., HANNEMANN, T., HEINTZ-MARTIN, V., JALOVAARA, M., KULU, H., MEGGIOLARO, S., MORTELMANS, D., PASTEELS, I., SEIZ, M., SOLAZ, A., 2017. Social policies, separation, and second birth spacing in Western Europe. *Demographic Research*, Vol. 37, pp. 1245–1274.

MATYSIAK, A., STYRC, M., and VIGNOLI, D., 2014. The educational gradient in marital disruption: a meta-analysis of European research findings. *Population studies*, Vol. 68, No. 2, pp. 197–215.

MCLANAHAN, S. 2004. Diverging Destinies: How Children Are Faring Under the Second Demographic Transition. *Demography*, Vol. 41, No. 4, pp. 607–627.

MCLANAHAN, S., PERCHESKI, C., 2008. Family Structure and the Reproduction of Inequalities. *Annual Review of Sociology*, Vol. 34, No. 1, pp. 257–276.

MCNAMEE, C., RALEY, K., 2011. A note on race, ethnicity and nativity differentials in remarriage in the United States. *Demographic Research*, Vol. 24, pp. 293–312.

NITSCHE, N., MATYSIAK, A., BAVEL, J.V., VIGNOLI, D., 2018. Partners' Educational Pairings and Fertility across Europe. *Demography*, No. 38, pp. 1–36.

OSIEWALSKA, B., 2017. Childlessness and fertility by couples' educational gender (in)equality in Austria, Bulgaria, and France. *Demographic Research*, Vol. 37, No. August, pp. 325–362.

PERELLI-HARRIS, B., SIGLE-RUSHTON, W., KREYENFELD, M., LAPPEGÅRD, T., KEIZER, R., BERGHAMMER, C., 2010. The educational gradient of childbearing within cohabitation in Europe. *Population and development review*, Vol. 36, No. 4, pp. 775–801.

RALEY, R.K., SWEENEY, M.M., 2020. Divorce, Repartnering, and Stepfamilies: A Decade in Review. *Journal of Marriage and Family*, Vol. 82, No. 1, pp. 81–99.

SCHNOR, C., VANASSCHE, S., VAN BAVEL, J., 2017. Stepfather or biological father? Education-specific pathways of postdivorce fatherhood. *Demographic Research*, Vol. 37, No. 1, pp. 1659–1694.

SHAFER, K., JAMES, S.L., 2013. Gender and Socioeconomic Status Differences in First and Second Marriage Formation. *Journal of Marriage and Family*, Vol. 75, No. 3, pp. 544–564.

THOMSON, E. 2014. Family Complexity in Europe. *The ANNALS of the American Academy of Political and Social Science*, Vol. 654, No. 1, pp. 245–258.

THOMSON, E., MCDONALD, E., BUMPASS, L., 1990. Fertility desires and fertility: Hers, his, and theirs. *Demography*, Vol. 27, No. 4, pp. 579–588.

TOULEMON, L. 2014. Single low-educated women and growing female hypogamy. Presented at Quetelet Seminar : 40th edition Fertility, childlessness and the family : A pluri-disciplinary approach, Louvain-la-Neuve.

TRIMARCHI, A. VAN BAVEL, J., 2018. Pathways to marital and non-marital first birth: The role of his and her education. *Vienna Yearbook of Population Research*, Vol. 15, pp. 143–179.

TRIMARCHI, A., VAN BAVEL, J., 2020. Partners' Educational Characteristics and Fertility: Disentangling the Effects of Earning Potential and Unemployment Risk on Second Births. *European Journal of Population*, No. 0123456789.

VAN BAVEL, J. 2021. Partner choice and partner markets. *Research handbook on the sociology of the family*, pp. 219–231.

Alessandra TRIMARCHI, University of Messina, alessandra.trimarchi@unime.it
Laurent TOULEMON, Institut National d'Études Démographiques (Ined), toulemon@ined.fr

LOW FERTILITY IN CONTEXT: THE CASE OF ITALY

Chiara Ludovica Comolli, Antonella Bianchino, Agata Maria Madia Carucci

Abstract. Numerous studies have shown that fertility behaviour in Italy, as elsewhere, strongly depends on subnational contextual factors. With this study we add to that literature investigating how group-specific fertility rates differ across fine-graded local areas' characteristics. We utilize ISTAT micro-level vital registration records including all births registered in each Italian municipality for the year 2022 and information about parity and parents' age to construct age- and parity- specific fertility rates. We match fertility data with the municipalities' Inner Areas classification measuring the availability of (distance from) crucial services, and with other indicators of municipalities' socioeconomic conditions that may influence fertility, such as per capita income, the share of high-tech workers and the availability of public childcare services. First, we present a descriptive illustration of the most recent group-specific fertility rates in different kinds of local areas in terms of population size, geographical location, and inner area class. Second, we utilize linear regression models to investigate the association between local areas' characteristics and age- and parity-specific fertility rates.

1. Introduction

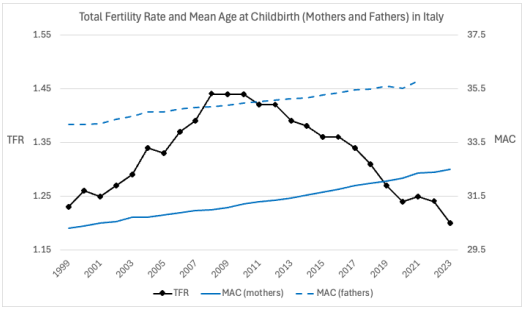
1.1. Recent developments in fertility in Italy

Persistently low birth rates contribute to population ageing and growing inequalities, posing a challenge to countries' social and financial sustainability. Italy registers a lowest-low level of fertility (Kohler *et al.*, 2002) with birth rates having decreased steadily in the past 50 years. In the more recent years, such long-term fertility descent produced a shrinking female population in reproductive age, which today contributes to the low number of births (Mencarini and Vignoli, 2018). Yet, part of the overall fertility decline is due to changes in childbearing behaviour, both in terms of *timing* and *quantity*. Figure 1 shows the Total Fertility Rate (TFR) and Mean Age at Childbearing (MAC) trends between 1999 to 2023. After an initial positive tendency at the beginning of the years 2000s due to the recuperation of births postponed during the 1990s, TFR declined sharply after 2010, going from 1.44 children per woman in 2010 to 1.20 children per woman in 2023. The MAC instead has been constantly increasing for both mothers and fathers since the 1990s, going from age 30.3 to 32.5 and from 34.2 to 35.8 respectively (Figure 1).

Several studies relate the onset of the fertility decline after 2010 to the consequences of the economic and financial crisis of 2008-09 (Sobotka *et al.*, 2011). It is worth noting

that the overall drop in TFRs was especially driven by declines in first order fertility rates and in the age-specific fertility rates of younger women, suggesting a delaying effect of the crisis on fertility schedules (Comolli, 2017). The postponement of the transition to parenthood to later ages was a consequence of the growing economic instability and labour market uncertainty during the first half of the 2010s. However, while usually as soon as the economy recovers couples recuperate births at least in part, many European countries including Italy, continued to register negative trends in the second half of the 2010s, long after the recovery. Nowadays, more than a decade after the Great Recession, fertility rates in Italy are getting closer to their historical minimum.

Figure 1 – *Total Fertility Rates and Mean Age at Childbearing in Italy 1999-2023.*



Source: Authors' elaboration based on ISTAT data.

1.2. Analysing low fertility with fine-graded municipality-level data

In Italy as in other contexts, territorial sociodemographic and socioeconomic characteristics strongly influence population change (Benassi *et al.*, 2023) and fertility (Campisi *et al.*, 2023; Vitali and Billari, 2017). Notable sub-regional differences are overlooked when focusing on higher-levels geographic entities only. As noted elsewhere, in fact, the Italian territorial and demographic heterogeneity increasingly requires shifting the focus of investigation to the local level (Salvati *et al.*, 2020; Tomassini *et al.*, 2024). Municipality-level data are useful because they allow leveraging a large variation in contextual characteristics. The availability of TFRs at the municipality-level is rare; to our knowledge, in fact, these data have not been used before. Previous studies on Italy have analysed fertility trends and their determinants with TFRs at regional or provincial level and Crude Birth Rates (CBRs) at the municipality level (Salvati *et al.*, 2020; Benassi and Carella, 2023). These studies have focused, and have been extremely successful, in modelling spatial dependency in fertility behaviour across territories (often large subnational entities), while our aim is to describe how fertility varies by fine-graded local areas characteristics.

This is especially important in relation to the presence (or lack) of services that makes a community attractive, or on the contrary, vulnerable to social and economic

marginalization (Reynaud and Miccoli, 2018). A useful indicator of the degree of presence or lack of such services is the classification of municipalities into inner areas provided by the Annual Reports on the National Strategy for Inner Areas of the Department for Cohesion Policies[1] (Barca and Carrosio, 2020). Municipalities that are disadvantaged by the absence of and distance from essential services are classified as Inner Areas[2].

While no study has addressed the fertility differential between central and inner areas, a rich literature has investigated the urban-rural divide in fertility behaviour. The urban socioeconomic context generates educational, wealth and occupational opportunities especially for women, and a cultural environment that, compared to rural areas, favours postponement of childbearing to later ages and disincentivise having large families (Kulu *et al.*, 2007; Riederer and Buber-Ennser, 2019). Yet, the contrast urban-rural (as much as the dichotomy central-inner areas) does not fully cover the multifaceted spectrum of intermediate levels of urbanization-rurality (as much as the centrality-marginalization in case of inner areas). Cities, and especially metropolitan regions and capitals, tend to have lower and later fertility compared to less densely populated urban areas, like suburban centres (Buelens, 2021).

The current study's objective is twofold. First, we aim at providing an empirical overview of the most recent estimates of group-specific fertility rates in the Italian context, measuring them at a very fine-graded local level. Second, we describe how the observed group-specific fertility rates vary depending on the contextual socioeconomic conditions.

2. Data and Method

2.1. Data and Variables

We utilize restricted-use micro-level vital registration records including all births registered by municipality for the year 2022 from the Italian Institute of Statistics. For each live birth we have information about the birth' parity and about the mother's nationality and age. We select native Italian women[3], and using live births and resident women by 5-years age groups in reproductive age 15-49, we construct total (TFR), and age- (ASFR) and parity- (PSFR) specific (first and second order) fertility rates at the municipality level.

[1] https://politichecoesione.governo.it/it/politica-di-coesione/strategie-tematiche-e-territoriali/strategie-territoriali/strategia-nazionale-aree-interne-snai/le-aree-interne-2021-2027/

[2] Essential services include: (i) complete education system up to upper secondary school, (ii) hospitals with first-level Emergency and Acceptance Departments, and (iii) at least "Silver" level railway stations.

[3] Identical analyses could be conducted on foreign-born women but in sake of conciseness, we decided to limit our study to native Italian women.

Then, we match this dataset with municipalities' characteristics that may be relevant for fertility behaviour. First, we merge the 2020's classification of the inner areas provided by the Annual Reports on the National Strategy for Inner Areas (Department for Cohesion Policies) distinguishing between Centres (hub, intra-municipality hub, and belt) and Inner Areas (intermediate, peripheral, and ultraperipheral). Inner areas are identified based on an accessibility indicator, calculated in terms of travel minutes by car to the nearest hub municipality (a municipality offering essential educational, health, and rail transport services). Municipalities that are more than 27 minutes away from the nearest hub are identified as inner areas. Centres include belts, i.e., peri-urban areas that are less than 27 minutes away from the nearest hub. Inner areas are further divided into three groups: intermediate (between 27 and 40 minutes away from the nearest hub), peripheral (between 40 and 67 minutes), ultraperipheral (more than 67 minutes).

Then, we consider municipalities' demographic and socioeconomic indicators. First, we merge municipalities' population and territorial extension (km^2), and calculate population density (ISTAT 2022[4]). Second, we add income and educational attainment which are well-documented determinants of fertility (Brand and Davis, 2012; Van Wijk and Billari, 2024). We measure income with the average before tax per capita income in the municipality (Minister for the Economy and Finance, MEF), and gender-specific education with the municipality's share of men and women with tertiary education (Permanent Census of Population, Housing and Enterprises ISTAT). Finally, we include measures of welfare provisions and labour market conditions that are also crucial factors influencing fertility (Neyer, 2013; Comolli, 2017). We operationalize welfare support in the municipality through (i) public expenditure on social services (2020, Euros per capita) and (ii) public childcare uptake (2021, percentage of children age 0-2 enrolled in public childcare) (ISTAT's Survey on interventions and social services of individual and associated municipalities). We measure labour market conditions with (i) unemployment rates by gender (ISTAT's 2021 Permanent Census of Population and Housing) and (ii) the relevance of the high-tech sector in the economy measured as percentage workers in high-technology manufacturing and knowledge-intensive high-technology services in the municipality (2021 ASIA-UL Statistical register of local units).

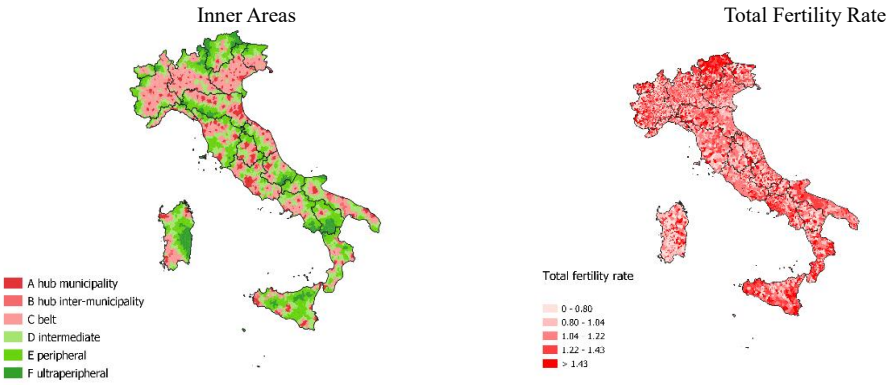
2.2. Method

TFRs have the fundamental advantage over CBRs of taking into consideration the age structure of the population for which we want to estimate fertility behaviour, providing a crucial indication of the intensity of the fertility behaviour in the population. Yet, estimating TFRs in very small areas can lead to unreliable measures insofar a small and exceptional number of live births in given years or age groups relative to a very small population of women in such age groups, can produce abnormal fertility levels. This is

[4] *Confini delle unità amministrative a fini statistici*, https://www.istat.it/it/archivio/222527.

the case for Italy which has a very large number of municipalities with overall population below 1,000 inhabitants (N=2,023) and with 1,001-5,000 inhabitants (N=3,507). To avoid such overestimation of fertility rates, we grouped together small municipalities (below 5,000 inhabitants) belonging to the same province and being of the same inner area class. After this regrouping we obtained 2,696 local areas, of which 2,374 are municipalities with a population of more than 5,000 individuals, and the remaining 322 are local areas obtained merging municipalities with a population of 5,000 individuals or less. This method is simple and transparent, although it has the limitation of reducing the number of local areas considered, and especially in those communities that are more strongly affected by the lack of services (small inner areas). We will address further this limitation in our conclusions. Table A.1 in the Appendix presents the distribution of all the variables considered in the analyses. We first show the territorial distribution of TFR, PSFR and ASFR by local area in 2022. Second, we present the results from a multivariate linear regression analysis which illustrates the differences in ASFR and PSFR by inner area class. Finally, we add the other socioeconomic indicators to investigate their role in the observed differences in fertility rates across types of local area.

Figure 2 – *Inner areas and Total Fertility Rates by municipality in Italy 2022.*



Source: Authors elaboration based on ISTAT data.

3. Results

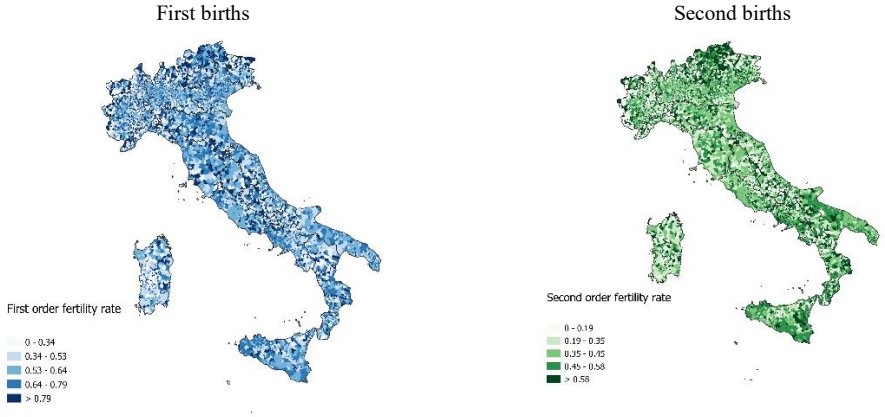
3.1. Descriptive analyses

Figure 2 (left panel) maps municipalities by their classification as Centres (hub, inter-municipality hub, and belt) or Inner Areas (intermediate, peripheral, and ultraperipheral). Almost half (48.5%, N=3,834) of Italian municipalities fall into at least one of the inner areas' typologies, mostly located along the Alpine and Apennine arc but also in plains

and coastal areas, from North to South. Among inner areas, 1,928 municipalities (24.4%) are classified as intermediate, 1,524 are peripheral (19.3%), and finally 382 are ultraperipheral (4.8%).

Figure 2 (right panel) presents a map of municipalities' TFRs in 2022, highlighting the territorial heterogeneity within regions and provinces in overall number of children per woman. We notice a strong territorial variation also in the fertility by parity (first and second order fertility rates) presented in Figure 3. These figures demonstrate the importance of exploiting municipality-level data and group-specific (e.g. parity) fertility rates. While there are some municipalities with similar first and second order fertility rates (e.g. high first and second order fertility rates, like in the province of Vercelli), we also see municipalities with relatively high first but low second order fertility rates (like in other provinces of Piemonte like Alessandria, Asti, or Cuneo).

Figure 3 – *First and Second Births Total Fertility Rates by municipality in Italy 2022.*



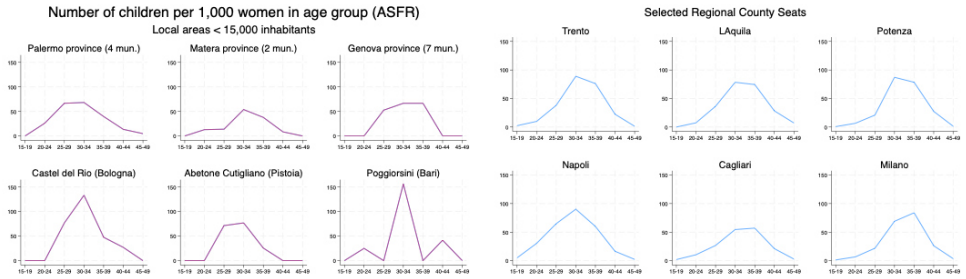
Source: Authors elaboration based on ISTAT data.

Figure 4 presents Age-specific Fertility Rates (ASFR) in 12 local areas selected because of different size, location and class of inner area. The left panel illustrates the age profile of fertility behaviour in six selected local areas with less then 15,000 inhabitants. The first three include very small municipalities grouped together due to their very low population. The three local areas include respectively four belt municipalities in the province of Palermo[5], two intermediate municipalities in the province of Matera[6], and seven ultraperipheral municipalities in the province of Genova[7]. The three areas include 11,799, 5,524, and 2,067 individuals. The other three

[5] Cerda 4,942, Lascari 3,674, Sciara 2,569 and Scillato 614 inhabitants.
[6] Calciano 667 and Tricarico 4,857 inhabitants.
[7] Fascia 73, Fontanigorda 244, Gorreto 94, Propata 113, Rondanina 60, Rovegno 492 and Santo Stefano d'Aveto 991 inhabitants.

graphs show ASFR in other small municipalities from Northern (Bologna), Central (Pistoia) and Southern (Bari) Italy. In terms of population, they range from 1,291 inhabitants in Poggiorsini to 1,902 in Castel del Rio. All three are inner areas but they are in different classes (Castel del Rio intermediate, Abetone Cutigliano ultraperipheral, Poggiorsini peripheral). Finally, the right panel illustrates the ASFRs in six selected regional county seats with populations ranging from 65,000 (Potenza) to 1.3 million (Milan) individuals. All county seats are classified as hubs.

Figure 4 – *Age specific fertility rates (ASFR) by local area in Italy 2022.*



Source: Authors elaboration based on ISTAT data.

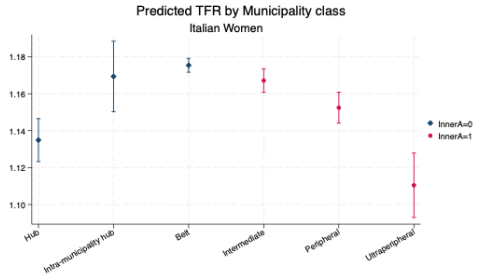
Figure 4 shows that fertility schedules differ quite substantially over local areas[8]. We do not observe strong patterns either along the more typical lines of enquiry (i.e. large vs. small area, southern vs. northern, etc) nor when distinguishing between centres and inner areas. The highest peak in ASFR is registered in the peripheral municipality of Poggiorsini in the province of Bari at 150 births per 1,000 women aged 30-34, while the lowest ASFR is observed in the intermediate municipalities in the province of Matera and in Cagliari, the county seat hub of Sardegna, at around 50 births per 1,000 women. The age at which the peak in fertility rates is observed varies from the late-twenties/early-thirties in the belt municipalities of the province of Palermo to the late-thirties in the big city hub of Milan.

3.2. Multivariate analyses

This section presents the results of multivariate linear regression models analysing the association between the inner area detailed class and the TFRs and group-specific fertility rates (ASFR and PSFR). Figures 5-6 present predicted fertility rates by local area class, net of population density in the area, share of men and women with tertiary education and regional dummies.

[8] The irregularities in small municipalities (e.g. Poggiorsini) are due to the low number of women in certain age groups and the year-to-year fluctuation in number of births to these women (see Method section).

Figure 5 – *Total Fertility Rates by local area class. Italy 2022.*



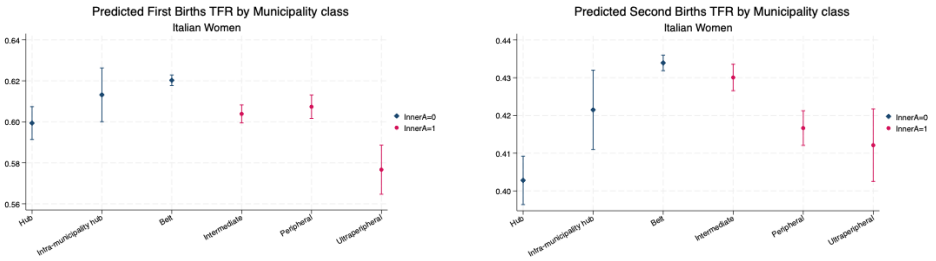
Source: Authors elaboration based on ISTAT data.

Figure 5 shows that the predicted TFR is higher in the belt municipalities and lowest in the ultraperipheral areas. The latter are only a few, so confidence intervals are quite large, and they overlap with the predicted TFR for the hubs so we cannot definitely conclude that in the ultraperipheral areas fertility is lower than in the large urban cities. Figure 6 shows, however, that first order fertility rates are significantly lower in ultraperipheral areas compared to any other kind of class of municipality while in the hubs what is lowest is the progression to a second child. The advantage of the belts also appears when we look at the predicted fertility rates for second births which are higher in those areas compared to any other.

If we look more specifically at the predicted ASFR in the different areas (Figure 7) we see childbearing delayed to later ages in hubs and belts relative to ultraperipheral areas, especially for first order fertility (Figure 8). Women living in ultraperipheral areas display similar fertility rates to belt areas until the late 20s (except for first order births to women 25-29 which are slightly higher in the ultraperipheral areas). However, from age 30 women display significantly lower fertility (early thirties in first order and late thirties in second order fertility rates).

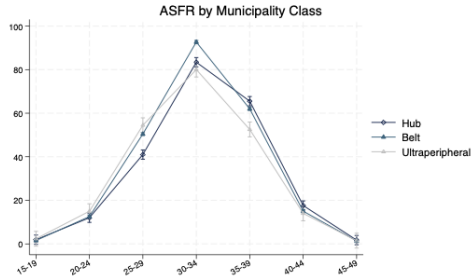
Finally, Table 1 shows that the differences in TFRs across class of local area do not entirely depend on socioeconomic municipality characteristics. Controlling for local areas' economic and labour market conditions (income levels, men's and women's unemployment rates, share of workers' high-tech specialization) and welfare support (share of children aged 0-2 in public childcare, expenditure on social policy) only modestly reduces TFRs' differences by local areas' class (Model 2).

Figure 6 – *First and Second order Fertility Rates by local area class. Italy 2022.*



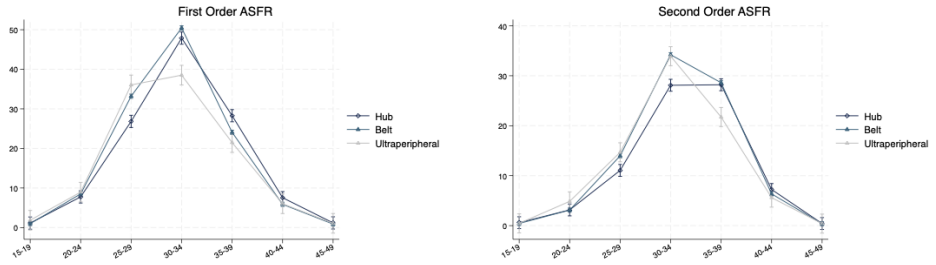
Source: Authors elaboration based on ISTAT data.

Figure 7 – *Age-specific Fertility Rates by Municipality Class. Italy 2022.*



Source: Authors elaboration based on ISTAT data.

Figure 8 – *Age- and Parity- specific Fertility Rates by Municipality Class. Italy 2022.*



Source: Authors elaboration based on ISTAT data.

Table 1 – *Total Fertility Rates, local area class and socioeconomic indicators. Coefficients from linear regression models. Italy 2022.*

	Model (1)		Model (2)	
	Coeff.	CI	Coeff.	CI
Local area classification - Reference category: Belt				
Hub	-0.040***	(-0.053 - -0.028)	-0.039***	(-0.051 - -0.026)
Intra-municipality hub	-0.006	(-0.026 - 0.013)	-0.005	(-0.024 - 0.015)
Intermediate	-0.034***	(-0.042 - -0.027)	-0.028***	(-0.036 - -0.020)
Peripheral	-0.049***	(-0.059 - -0.039)	-0.037***	(-0.047 - -0.027)
Ultraperipheral	-0.091***	(-0.109 - -0.073)	-0.072***	(-0.091 - -0.054)
Population Density	0.019***	(0.015 - 0.022)	0.015***	(0.012 - 0.019)
Men with tertiary education (%)	-0.001	(-0.003 - 0.000)	-0.003***	(-0.005 - -0.002)
Women with tertiary education (%)	-0.003***	(-0.004 - -0.002)	-0.004***	(-0.005 - -0.002)
Per-capita Income (per 10 thousand Euros)			0.008***	(0.007 - 0.010)
Male Unemployment Rate (%)			-0.006***	(-0.009 - -0.003)
Female Unemployment Rate (%)			0.004***	(0.002 - 0.007)
High-tech specialization (%)			0.002***	(0.002 - 0.003)
Children (age 0-2) in childcare (per 10% increase)			-0.003**	(-0.006 - -0.000)
Public Social Expenditure per capita (per 100 Euros)			0.009***	(0.003 - 0.016)
Regional dummies	YES		YES	
Constant	1.271***	(1.251 - 1.291)	1.122***	(1.078 - 1.167)
Observations	18,872		18,865	
R-squared	0.179		0.188	

Source: Authors elaboration based on ISTAT data.

4. Discussion

The study suffers from a few limitations. First, the aggregation of the smaller municipalities reduces the advantage of disposing of municipality-level data and the variability in the area characteristics. We prioritize the correct estimation of the fertility rates over the exploitation of the variety of the over 7.9 thousand Italian municipalities. Future research could elaborate on our analyses and, for instance, gathering more time points data, estimate fertility rates in very small municipalities by averaging live births and women by age groups over a few years to avoid small fluctuations in live births having a large and implausible impact on TFR estimates. Second, for the sake of conciseness here we did not look at variation over time although it would be important to investigate time trends in fertility in response to changes in the characteristics of municipalities, especially from before to after the Great Recession, when the fertility decline started. Finally, the R^2 in Table 2 is quite low, suggesting that other variables not included in our models also contribute to the variation in TFRs across local areas. Our aim was not to fully explain the territorial variation in TFR, however, we acknowledge the low overall explanatory power of the model. Despite these limitations, our study points to the relevance of local area characteristics and especially the accessibility as an

important determinant of fertility behaviour. We stress the importance of investigating group-specific behaviour which informs not only about the levels of fertility but also about the timing and parity progression of fertility. Building on our findings, possible avenues for future studies concerns the role of the interaction between local area type and socioeconomic structures and its change over time in shaping group-specific fertility.

Appendix

Table 1 – *Descriptive statistics.*

Variable	Obs	Mean	Std. Dev.	Min	Max
TFR	2,696	1.167	.217	0	2.5
First Order FR	2,696	.613	.138	0	2.5
Second Order FR	2,696	.428	.117	0	.934
Km^2	2,696	784.306	1,650.848	10.795	18,296.012
TOT popres2021	2,696	153,825.07	498,652.9	1071	19,391,582
Population density	2,696	622.94	965.86	1.006	11,766.31
Public childcare 0-2	2,696	13.898	12.297	0	100
High-tech Specialization	2,696	2.581	3.749	0	48.368
% Men Tertiary Edu	2,696	18.709	5.206	5.769	47.097
% Women Tertiary Edu	2,696	28.94	6.152	3.448	56.075
Public Social Expenditure	2,696	20,206.135	3,776.182	11,274.424	48,385.574
Income PC	2,695	111.151	78.343	0	618.802
Men Unemployment (%)	2,696	7.552	3.563	1.029	24.939
Women Unemployment (%)	2,696	10.832	4.314	1.613	25.189
Inner Area Class					
Hub	182	6.75%			
Intermunicipal Hub	59	2.19%			
Belt	1,546	57.34%			
Intermediate	529	19.62%			
Peripheric	309	11.46%			
Ultraperipheral	71	2.63%			

Source: Authors elaboration based on ISTAT data.

References

BARCA, F., CARROSIO, G. 2020. Un modello di policy place-based: la Strategia Nazionale per le Aree Interne. In G. OSTI, JACHIA E. (Eds.) *Un disegno di rinascita delle aree interne*, Bologna: Il Mulino.

BENASSI, F., CARELLA, M. 2023. Modelling geographical variations in fertility and population density of Italian and foreign populations at the local scale: a spatial Durbin approach for Italy (2002–2018),*Quality Quantity*, Vol.57, No.3, pp. 2147-2164.

BENASSI, F., MUCCIARDI, M., CARELLA, M., NACCARATO, A., SALVATI, L. 2023. Spatial Segregation in Action? An Empirical Assessment of Population Concentration of Foreigners and Nationals in Italy, 2002–2018. *International Migration Review*.

BRAND, J. E., DAVIS, D. 2012. The impact of college education on fertility: Evidence for heterogeneous effects. *Demography*, Vol. 48, pp. 863–887.

BUELENS, M. 2021. Subnational spatial variations of fertility timing in Europe since 1990. *Cybergeo: European Journal of Geography*, No. 1000.

CAMPISI, N., KULU, H., MIKOLAI, J., KLÜSENER, S., MYRSKYLÄ, M. 2020. Spatial variation in fertility across Europe: Patterns and determinants. *Population, Space and Place*, Vol. 26, No. 4.

COMOLLI, C. L. 2017. The fertility response to the Great Recession in Europe and the United States: Structural economic conditions and perceived economic uncertainty. *Demographic research*, Vol. 36, pp. 1549-1600.

KOHLER, H. P., BILLARI, F. C., ORTEGA, J. A. 2002. The emergence of lowest-low fertility in Europe during the 1990s. *Population and development review*,Vol. 28, No. 4, pp. 641-680.

KULU, H., VIKAT, A., ANDERSSON, G. 2007. Settlement size and fertility in the Nordic countries. *Population Studies*, Vol. 61, No. 3, pp. 265–285.

MENCARINI, L., VIGNOLI, D. 2018. *Genitori cercasi. L'Italia nella trappola demografica*. Egea, Università Bocconi Editore.

NEYER, G. 2013. *Welfare states, family policies, and fertility in Europe. In The demography of Europe*. Dordrecht: Springer Netherlands.

REYNAUD, C., MICCOLI, S. 2018. Depopulation and the aging population: The relationship in Italian municipalities. *Sustainability*, Vol. 10, No.4, pp. 1004-1016.

RIEDERER, B., BUBER-ENNSER, I. 2019. Regional context and realization of fertility intentions: The role of the urban context. *Regional Studies*, Vol. 53, No.12, pp. 1669–1679.

SALVATI, L., BENASSI, F., MICCOLI, S., RABIEI-DASTJERDI, H., MATTHEWS, S.A. 2020. Spatial variability of total fertility rate and crude birth rate in a low-fertility country: Patterns and trends in regional and local scale heterogeneity across Italy, 2002–2018. *Applied Geography*. Vol. 124.

SOBOTKA, T., SKIRBEKK, V., PHILIPOV, D. 2011. Economic Recession and Fertility in the Developed World. *Population and Development Review*, Vol. 37, No. 2, 267–306.

TOMASSINI, C., ALBERTINI, M., C. LALLO 2024. *Avanzare insieme nella società anziana. Considerazioni multidisciplinari sulla domanda di assistenza agli anziani in Italia*. Bologna: Il Mulino.

VAN WIJK, D., BILLARI, F. C. 2024. Fertility postponement, economic uncertainty, and the increasing income prerequisites of parenthood. *Population and Development Review*, Vol. 50, No. 2, pp. 287-322.

VITALI, A., BILLARI, F.C. 2017. Changing determinants of low fertility and diffusion: a spatial analysis for Italy. *Population Space and Place*, Vol. 23, No.2.

Chiara Ludovica COMOLLI, University of Bologna, chiara.comolli@unibo.it
Antonella BIANCHINO, Istat, bianchin@istat.it
Agata Maria Madia CARUCCI, Istat, carucci@istat.it

IMMIGRATION, ECONOMIC HARDSHIP AND POPULISM IN 2018 ELECTORAL OUTCOME IN ITALY. WAS GROUP THREAT THEORY RIGHT?[1]

Gabriele Ruiu, Pietro Mazzette

Abstract. According to Group Threat Theory, when the relative size of the immigrant population increases compared to natives, immigrants are increasingly perceived as a menace. This paper argues that right-wing populist parties have effectively exploited threat mechanisms to achieve considerable success in the 2018 Italian National Electoral Competition. The study uses a dataset at the electoral district level obtained by combining data from the Ministry of Interior data and the Italian National Institute of Statistics. In particular, the incidence of foreigners in the total population and socio-economic indicators are examined in relation to the 2018 electoral outcome of right-wing populist parties. We use spatial autoregressive models to account for spatial dependencies and spatial heterogeneity in the data. We find that the higher the share of foreigners and the lower the employment rate, the greater the electoral preference for the platform proposed by right-wing populist parties.

1. Introduction

Populism and immigration are intertwined phenomena that significantly shape contemporary political landscapes worldwide. Their relationship is complex, often marked by rhetoric and societal polarization. Populist movements frequently exploit immigration issues to mobilize support, tapping into public anxieties and grievances.

According to Group Threat Theory (GTT), when the relative size of the immigrant population increases compared to natives, immigrants are increasingly perceived as a threat to: i) cultural identity; ii) the security of the insider group; iii) economic security due to the competition for jobs and welfare support (see, among others, Blumer, 1958; Stephan and Stephan, 2000). These perceived threats are stronger when the condition of the local economy is harsher.

As well-known among Italian political scientists, immigration was one of the hot topics in the 2018 electoral campaign (Combei *et al.,* 2020). This electoral competition represented a turning point in recent Italian politics. Indeed, the electoral outcome was characterized by the decline of two mainstream parties (Partito

[1] The authors have contributed equally to each section of the paper.

Democratico and Forza Italia) and the striking success of two populist parties (Lega and Five Stars Movement). Populist leaders capitalized on fears of cultural dilution, economic competition, and security risks associated with immigration to rally their base. By framing immigration as a crisis, they reinforced the idea that only they can protect the interests of the native population (see Faggian *et al* 2021).

Moreover, populism tends to simplify complex issues, offering straightforward solutions to intricate problems. In the case of immigration, this translates into calls for tighter borders, deportation, or even outright bans on certain groups.

Conversely, immigration can also fuel the rise of populism. Rapid demographic changes and the perceived failure of mainstream parties to address immigration concerns can breed resentment and disillusionment among segments of the native population. This discontent provides fertile ground for populist rhetoric, promising to restore control and the prominence of the native population. In other words, populist parties exploit the threat mechanisms the Group Threat Theory describes. Dennison and Geddes (2018) showed that migration crises, such as the one in 2015-16 in Europe, significantly raised the salience of immigration in political and media discourse, activating pre-existing anxieties. According to Dennison and Geddes, this shift allowed populist parties to frame immigration as a threat to national identity and sovereignty.

Following Guiso et al. (2017), we consider a party populist when it promotes short-term policies that may have an immediate high electoral payoff (such as trade protection, resistance to foreign immigration, citizenship income, or other safe employment policies) with no regard for their long-term costs. Using this criterion and analyzing the electoral programs and public statements made by political party leaders, we identify Lega, Fratelli d'Italia (FDI), and CasaPound as "right-wing populist parties exploiting the threat mechanism".[2] It should be remarked that another populist, anti-establishment party, the Five Star Movement (5SM), never used the same crude anti-immigrant rhetoric as Lega. In particular, 5SM expressed solidarity for refugees in their electoral programme and proposed limiting their stay

[2] See Livi Bacci on Neodemos (2018) for a detailed analysis of the electoral program of each party on the immigration issue. See also Emanuele *at al* (2019) for a qualitative analysis of the 2018 political campaign. Some examples of the rhetoric used by Lega, FDI and CasaPound are the following (our translation): "Since illegal immigration has doubled in the last 5 years, sexual crimes have increased by 40%. The PD's policy is mainly responsible for all this. The left and the M5S prioritize ideology over reality, which is clear to everyone. For them, the immigrant, even if illegal, is more important than anything else, even if he rapes, kills, or dismembers a body" (Rossano Sasso, Lega, Elected); "Thousands of homeless people, many of them Italians, risk dying from frostbite. The Government should implement an extraordinary assistance plan: they should be treated like asylum seekers who landed illegally and be guaranteed food and shelter." (Giorgia Meloni, Leader of FDI); "Here they chase the children of Italians to vaccinate them, while those who land illegally can bring any type of disease." (Matteo Salvini, Leader of Lega).

in Italy by negotiating redistribution quotas within European Union. Other social issues were focal to their electoral platform (e.g., introducing a basic income law, the so-called "*reddito di cittadinanza*", establishing stricter limits on holding elective offices for those convicted of criminal offences, etc.). Therefore, our focus is on right-wing populism. See also Faggian *et al.* (2021), who remarked that immigration was not the driver of the 2018 electoral success of 5SM.

Building on GTT predictions, our research questions may be formulated as follows:

RQ1: Have right-wing populist parties obtained more success where the presence of immigrants is higher?

RQ2: Have right-wing populist parties obtained more success where there are more significant socio-economic vulnerabilities?

The answers are not trivial, given that in some regional contexts where Lega has its historical roots, immigrants are crucial for the local economy (e.g., the agricultural workforce in Veneto).

Furthermore, resident foreigners can vote only for local representatives but not for national elections. Obviously, we may have territories where the presence of immigrants is consolidated to have second-generation immigrants who have acquired Italian citizenship. Therefore, we may have that these territories are, simultaneously, those with the highest presence of foreigners who are not entitled to vote and those with the highest incidence of second-generation immigrants who presumably will not vote for an anti-immigrant platform. Indeed, according to Istat data, 87% of foreign citizenship acquisitions happened in the Center-North of Italy in 2022. In the same year, 83.7% of foreigners were residents in the Center-North.[3]

The current study uses a dataset at the electoral district level obtained by combining data from the Ministry of Interior and the Italian National Institute of Statistics (Istat). In particular, the incidence of foreigners in the total population and socio-economic indicators are examined in relation to the 2018 electoral outcome of right-wing populist parties. We use spatial autoregressive models to account for spatial dependencies and spatial heterogeneity in the data.

Using a completely different approach compared to the one proposed in this paper, Faggian *et al.* (2021) have already pointed out a positive correlation between the presence of immigrants and votes for the Lega. Our empirical exercise confirms this finding, solving some empirical issues left open by this previous contribution. To the best of our knowledge, this is the first work to propose a spatial approach to investigate the success of right-wing populism.

[3] In 2022, there were about 5 millions of resident foreigners and about 200,000 citizenship acquisitions.

The paper is structured as follows: the second section describes data and method; the third section presents the results; and in the fourth section, we delineate some general conclusions and trace the future development of this research.

2. Data and method

The Istat disseminated the 2018 Geographical shapefiles on electoral districts aggregating municipalities based on the Legislative Decree No. 189 dispositions of December 12, 2017. Istat also integrated a set of socio-economic indicators that includes information on the demographic structure of the population, the economic situation (the employment rate), the structure of the local economy (share of people employed in primary, secondary and tertiary sectors), local human capital (incidence of both low educated people and high educated people). These indicators are calculated by exploiting the 2011 Italian Population Census. However, these databases do not contain information about electoral outcomes. We integrated Istat's data with publicly available data on election results from the Ministry of the Interior, creating a joint database containing each district's electoral outcome and socio-economic and demographic statistics. Note that the fact that socio-economic indicators are not precisely simultaneous to electoral data is not necessarily a drawback in our integrated databases. Indeed, this may attenuate a potential reverse causality problem in our analysis. That is, the presence of foreigners could be higher where they are more tolerated. Thus, one might argue that tolerance, which is related to rejecting anti-immigrant rhetoric, determines the presence of foreigners and not *vice versa*. However, this potential mechanism should be partially mitigated using data from a few years before the electoral outcomes. In addition, as argued by Dennison and Geddes (2018), the salience of migration at time t is only reactivated by contemporaneous episodes, but the reasons behind the success of anti-immigration rhetoric derive from issues that arose well before time t. Concerning the immigrant stock size, we also ran another empirical exercise using the presence of immigrants averaged for the years 2015-2017. The Pearson correlation between the 2011 immigrant stock size and the more recent measure is 0.978, suggesting that being a persistent cumulative process (Massey, 1990) also, using the 2011 data we can still characterize territorial differences in the presence of immigrants.

2.1. Method

As anticipated, We define as «right-wing populist parties exploiting the threat mechanism» the following three lists: Lega, Fratelli di Italia and CasaPound. We

sum the share of votes going to right-wing populist parties on the total number of votes as the dependent variable in the following spatial Durbin regression model:

$$Y_i = X\beta + \bar{Y}\lambda + \bar{X}\gamma + u \quad (1)$$

Our territorial units i are the 63 plurinominal electoral districts for the election of the Chamber of Deputies.[4] The choice of the plurinominal electoral district is because the electoral rules imply that only in this type of district is possible to distinguish the share of votes going to each party. We focus our analysis on Chamber of Deputies because the composition of the electoral districts for Senate is more geographically aggregated, leading to only 33 districts.

X is an n × K matrix of exogenous covariates. $\bar{Y} = W_p Y_p$ is a n × P matrix of spatial lags for the dependent variable. $\bar{X} = W_p X_p$ is n × P matrix of spatial lags for the exogenous covariates. Our main explicative variable is the incidence of foreigners over 1,000 residents.

We assume that spatial spillovers are captured not only by the spatial lagged dependent variable but also by the spatial lag of the variable capturing foreigner presence. In other words, people from a district i may be feel threated (and vote for parties sustaining anti-immigrants stances) even by the presence of immigrants in neighboring communities. We will call this indirect effect as "encirclement effect".

We define contiguity weights using the Queen criterion. The weights are normalized using the row sum.

The employment rate and the Mazziotta-Pareto Index are also used as regressors in equation (1) to capture the difference in the economic situation and in the socio-material vulnerabilities among territories.[5] Among other control variables, we include: the incidence of people with a post-secondary education title, the aging index, and the share of employment in the secondary sector.

The level of education is defined as the percentage of highly educated people for two reasons: i) education fosters analytical skills and critical thinking, enabling individuals to assess political and social issues with greater nuance. This should reduce the likelihood of accepting simplistic populist narratives; ii) Highly educated people do not suffer competition for jobs from immigrants (Bratti and Conti, 2018).

[4] The 2018 Italian elections were based on the so-called *Rosatellum* Law. The structure of the law is configured as a mixed electoral system. In particular, part of the deputies (the 37% of the total) are elected in *collegi uninominali* where the candidate that collect most preference is elected, while the 61% are assigned through *collegi plurinominali,* that is, seats are proportionally distributed among the coalitions and individual lists that surpassed a required national threshold. The remaining 2% of seats is allocated to Italians residing abroad. Thus, in uninominal district a coalition support a single candidate, making it impossible to distinguish among the shares of each party.

[5] The Mazziotta-Pareto Index of social and material vulnerability is a synthetic non-compensatory index that combines seven socio-demographic indicators. For details see Mazziotta and Pareto (2015).

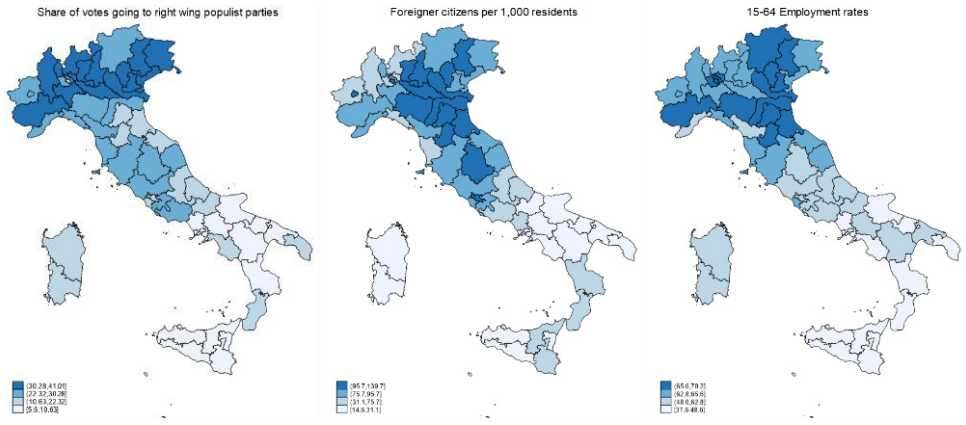
Hence, they should be less susceptible to threat mechanisms. The percentage of employment in the secondary sector is introduced to grasp better the difference in the economic structure between Northern and Southern Italy. The aging index is introduced because older population might be less accustomed to cultural diversity and more inclined to support policies that favor restrictionism and populist rhetoric, seeing these as protective measures.

It is important to note that an earlier work on the topic, carried out by Faggian *et al.* (2021), used municipal-level data to analyze the success of Lega and 5SM in the 2018 Senate elections. They focus on the Senate to consider the results associated with more mature electors and to avoid the over-representation of very young electors that characterize the 5SM (only individuals older than 25 can vote to elect the Senate). Another critical difference between their empirical exercise and the present one is that they did not use spatial regression to model the long-lasting political tradition that characterizes some Italian territories. These traditional zones of influence are generally more extended than a single municipality (see Fitzgerald, 2018). This becomes even more relevant in the case of Lega, which derived from Lega Nord, a party, in turn, born from the merger of three main regional autonomist movements (Lega Lombarda in Lombardy, Liga Veneta in Veneto, Piemont Autonomista in Piedmont). The same considerations about the historical territorial ties could be applied to the so-called Red-Belt (Diamanti, 2009). On the empirical ground, ignoring spatial autocorrelation can lead to errors in estimates and inferences (Anselin, 1988). The fact that the origins of Lega are strongly related to some Northern Italian autonomist movements leads us to run a separate analysis only using the share of votes going to this party. This approach enables a clearer understanding of the role played by threat mechanisms. Lega Nord has historically expressed similarly harsh opinions about people from Southern Italy. With the transition from the regional party Lega Nord to the national party Lega, under the leadership of Matteo Salvini, the focus shifted from internal to international immigrants. Thus, conducting separate analyses allows us to confirm the party's ability to exploit threat mechanisms even among those previously targets of its populist campaigns.

In addition, we believe that local factors too heavily influence results at the municipal level. The electoral rules for the plurinominal district impose the so-called blocked list, i.e. people are allowed to express a vote on the list but not to express a preference for a candidate within the list. Albertazzi and Zulianello (2021) have shown that local context variations are related to variations in the electoral success of populist parties. Despite the elections being at the national level, these results seem to suggest that people are not allowed to choose directly the representative they trust, so they tend to vote for the platform that better addresses local necessities. By aggregating data at the district level, we believe these local distortions may compensate for each other and thus become less relevant. In any case, our empirical

exercise may also be considered a corroboration of Faggian *et al*'s findings using different statistical techniques and more aggregate data.[6] Figure 1 reports the distribution of the share of votes going to right-wing populist parties, the incidence of foreigners in the total population, and the employment rate. Colors are attributed based on quartiles.

Figure 1 – *Electoral outcome for right-wing populist parties, foreigner incidence and employment rate. 2018 Italian electoral districts*



Note:Data refer to 2018 for electoral outcome while to 2011 for foreign presence and 15-64 employment rates.

In general, it seems that where the presence of foreigners is higher, the share of votes going to populist parties tends to be higher. Interestingly, especially in some southern districts, the share of votes to populists seems to be higher where employment rates are lower. This preliminary evidence is in line with GTT's prediction.

3. Empirical results

Table 1 shows the estimation results of equation 1 for the total share of right-wing populist parties (columns 1 and 1-B) and Lega (columns 2 and 2 -B)[7].

[6] Note also that using electoral districts allows us to solve the problem of giving equal importance to each municipality, even in cases of large differences in population size. Plurinominal districts are constructed by aggregating contiguous municipalities under the constraint that the district population does not deviate (in absolute terms) by more than 20% from the total average population of the districts.
[7] Before running the spatial durbin model, we tested spatial dependency though the Global Moran test. We strongly reject the hypothesis of independence (chi2=81.57, p-value: 0.000)

Table 1 – *Explaining the share of votes going to right-wing populism.*

	(1)	(1 - B)	(2)	(2-B)
	Share Pop	Share Pop	Share Lega	Share Lega
Foreigners	0.109 (0.025)***	.	0.089 (0.022)***	.
Foreigners_15_17	.	0.083 (0.024)***		0.068 (0.022)***
Employment rate	-1.176 (0.205)***	-1.025 (0.205)***	-0.868 (0.183)***	-0.749 (0.189)***
Ageing Index	0.018 (0.014)	0.010 (0.014)	0.015 (0.013)	0.007 (0.013)
Inc_High_Educated	-0.203 (0.153)	-0.181 (0.162)	-0.329 (0.140)**	-0.280 (0.149)*
Employed_Sec	0.272 (0.070)***	0.339 (0.075)***	0.294 (0.064)***	0.326 (0.068)***
MazziottaPareto	-0.630 (0.104)***	-0.502 (0.098)***	-0.472 (0.087)***	-0.379 (0.085)***
Spatial lags				
Foreigners	0.124 (0.038)***		0.083 (0.033)**	
Foreigners_15_17		0.107 (0.033)***		0.075 (0.029)**
Share Pop	0.472 (0.128)***	0.749 (0.107)***		
Share Lega			0.510 (0.105)***	
N	63	63	63	63
pseudo R^2	0.936	0.876	0.950	0.907

The intercepts of each models have been omitted to save space.
Standard errors in parentheses
** p < 0.10, ** p < 0.05, *** p < 0.01*
Wald test of spatial terms (MODEL 1): chi2(2) = 34.95 Prob > chi2 = 0.0000
Wald test of spatial terms (MODEL 1-B): chi2(2) = 69.47 Prob > chi2 = 0.000
Wald test of spatial terms (MODEL 2): chi2(2) = 34.53 Prob > chi2 = 0.0000
Wald test of spatial terms (MODEL 2-B): chi2(2) = 93.95 Prob > chi2 = 0.0000

First of all, note that in both columns, the spatial lags are highly statistically significant, which suggests that the necessity of considering the spatial dimension is confirmed on empirical grounds. In particular, we have a positive spatial correlation for both the foreign presence and the share of populist parties (Lega, respectively). This means that our hypothesis of the existence of an encirclement effect seems to be supported by data. In addition, the results support the idea of the existence of a territorial model of voting. The only relevant difference between the first two models and models 2 and 2-B, is that the incidence of highly educated people is significant only in the latter.

A negative relation between votes going to the Lega and educational levels was already reported by Faggian *et al.* (2020). Threat mechanisms are particularly effective among those who may suffer the competition for jobs or for the welfare system from immigrants. Given that immigrants in Italy are mainly employed in low-skill positions (Bratti and Conti 2018), we believe that the negative correlation

between the incidence of highly educated people and the vote share going to Lega is very reasonable. It is more difficult to explain why this result is not confirmed when the analysis is extended to other right-wing populist parties. One hypothesis could be that Lega has historically attracted votes from middle and lower social classes, as well as from less urbanized areas, particularly in the North. As education levels rise, part of this electorate might lose interest in this more regional approach and turn to parties offering a more national or global vision.

It should be noted that the beta coefficients reported in Table 1 could not be interpreted as partial derivates of the dependent variable with respect to each regressor as in the traditional linear regression model since the effect of an explanatory variable also passes through the spatial spillover. More specifically, the spatial lag of the dependent variable modifies the covariate effects. A change in the presence of foreigners in district i changes the conditional mean of y in that i (direct effect), and that change in y_i changes the conditional mean of y in all contiguous districts. Hence, the indirect (or spillover) effect is the effect of foreigner presence on the conditional mean of the share of votes in other districts. Because a spatial lag of foreigner presence is included in the model, there is also a second indirect effect. The latter implies that a change in foreigners in i changes the conditional mean of y in neighbouring districts. The sum of the direct and two indirect effects is our total effect. In Table 2, we then report the direct, the indirect, and total effects of our main variables of interest in model 1: the presence of foreigners, the employment rate, and the Mazziotta-Pareto Index.

Table 2 – *Direct, Indirect and Total effect of the presence of foreigners, the employment rate and the Mazziotta-Pareto Index.*

direct		Coeff	SE	Z	P-Value	Confidence Int	
	Foreign presence	0.139	0.026	5.28	0.000	0.088	0.191
	Employment rate	-1.271	0.226	-5.63	0.000	-1.715	-0.829
	Mazziotta_Pareto_Index	-0.681	0.099	-6.87	0.000	-0.876	-0.487
indirect							
	foreign	0.301	0.090	3.31	0.001	0.123	0.479
	employment rate	-0.954	0.491	-1.94	0.052	-1.916	0.007
	Mazziotta_Pareto_Index	-0.511	0.223	-2.29	0.022	-0.949	-0.073
total							
	foreign	0.440	0.104	4.22	0.000	0.236	0.645
	employment rate	-2.226	0.638	-3.49	0.000	-3.477	-0.975
	Mazziotta_Pareto_Index	-1.193	0.253	-4.72	0.000	-1.688	-0.697

Confirming GTT's predictions, the results indicate that a one-point increase in the incidence of foreigners leads to an increase of 0.44 points in the share of right-wing populist parties, and interestingly, the effect seems to be mainly driven by what

we have called the encirclement effect. Also confirming GTT, for one-point increase in the employment rate, we have a substantial decrease in the share of votes going to right-wing populists. This result also aligns with that obtained by Faggian et al. 2020. On the contrary, with respect to the presence of foreigners, the effect of employment is mainly driven by direct effects. Finally, an increase in socio-material vulnerability leads to a decrease in the consensus for right-wing populism. In this case, it is reasonable to assume that preferences went to 5MS, given that the introduction of basic income was its warhorse during the electoral campaign.

As a check for our model, we test for an eventual auto-correlation in model residuals. We firmly reject the auto-correlation, indicating a good model specification.

By way of final empirical exercises, keeping all the other variables constant, we predicted our dependent variable, assuming a low value for the foreigner presence (the mean minus one standard deviation) and alternatively assuming a high value (mean plus one standard deviation). We repeat the same exercise for the employment rate. Results are reported in Table 3 and refer to model 1.

Table 3 – *Predicting the share of populists for low and high values of foreigner presence and employment rate.*

	Margin	Delta-method st. errors.
Foreign mean - 1 sd	6.817	3.58*
Foreign mean	21.793	0.571***
Foreign + 1 sd	36.77	3.605***
Emplo rate mean - 1 sd	43.166	6.159***
Emplo rate mean	21.793	0.571***
Emplo rate + 1 sd	0.42	6.149

Interestingly, the effect of foreigners' presence seems symmetric around the mean. When their presence is below/above one standard deviation of the mean, we have a share of fifteen percentage points lower/higher than the baseline. The same cannot be said for the employment rates. When employment is high, one of the mechanisms suggested by GTT, the competition for jobs, becomes weaker. Accordingly, our results indicate that the preference for right-wing populist parties is high, especially with a low employment rate.

Note also from table 1 that the results seem to not change significantly when we use the more recent measure of immigrant presence.

4. Conclusions

Our empirical analysis supports the idea that right-wing populism has successfully exploited threat mechanisms. Indeed, GTT predicts that the anti-immigrant rhetoric will be particularly effective where the share of foreigners is relatively high, and economic difficulties may generate competition between insiders and outsiders.

The authors do not want to express any value judgment on exploiting this type of mechanism. In any case, a note of caution should emerge from the present work. When fears are fueled, the consequences can be unpredictable. Research has shown that manipulating perceived threats can have profound and often unintended impacts on social cohesion and political stability (Allport, 1954; Stephan and Stephan, 2000).

When political entities capitalize on and amplify existing anxieties, it may lead to increased social polarization, heightened intergroup tensions, and even spur radicalization among certain population segments. Thus, while the socio-economic context drives the initial sense of insecurity, instrumentalizing these fears for political gain can destabilize communities and erode democratic norms. It is crucial for policymakers and stakeholders to recognize the potential for such outcomes and to strive for approaches that address the root causes of these fears without exacerbating divisions within society.

A possible future direction for our analysis is to reconstruct a crime rate at the electoral section level and include it among predictors. Indeed, crimes involving immigrants were frequently exploited by populist leaders to ride the threat mechanism. We also plan to extend the analysis to the 2023 election, which has seen FDI eroding Lega's consensus.

Acknowledgements

This paper recognizes the funding of the project "APEGADRUS"- DM737/2021-22. DM737 is an initiative financed in the ambit of the EU -NEXT GEN FUNDS. Pietro Mazzette recognizes that he worked on this paper in the ambit of the Ph.D. in Economics, Management and Quantitative Methods, University of Sassari.

References

ALBERTAZZI D., ZULIANELLO M. 2021. Populist electoral competition in Italy: the impact of sub-national contextual factors. *Contemporary Italian Politics*, Vol. 13, No. 1, pp. 4–30.

ALLPORT G.W. 1954. *The Nature of Prejudice*. Cambridge (MA): Addison-Wesley Publishing.

ANSELIN L. 1988. *Spatial Econometrics: Methods and Models*, Dordrecht(NL): Kluwer Academic Publishers.

BRATTI M., CONTI C. 2018. The effect of immigration on innovation in Italy. *Regional Studies*, Vol. 52, No 7, 934-947.

BLUMER H. 1958. Race Prejudice as a Sense of Group Position. *Pacific Sociological Review*, Vol. 1, No. 1, pp. 3-7.

COMBEI C.R., FARNÈ M., PINTO L., GIANNETTI D. 2020. Populism and Policy Issues: Examining Political Communication on Twitter in Italy 2018-2019. *Italian Political Science*, Vol. 15, No. 2, pp. 223-241.

DENNISON J., GEDDES A. 2019. A Rising Tide? The Salience of Immigration and the Rise of Anti-Immigration Political Parties in Western Europe. *The Political quarterly*. Doi: https://doi.org/10.1111/1467-923X.12620.

DIAMANTI I. 2009. *Mappe dell'Italia politica. Bianco, rosso, verde, azzurro … e tricolore*. Bologna: Il Mulino.

EMANUELE V., MAGGINI N., PAPARO A. 2019. The times they are a-changin': party campaign strategies in the 2018 Italian election. *West European Politics*, Vol. 43, No. 3, pp. 665–687.

FAGGIAN A., MODICA M., MODREGO F., URSO G. 2021. One country, two populist parties: Voting patterns of the 2018 Italian elections and their determinants. *Regional Science Policy & Practice*, Vol. 13, No. 2, pp. 397-414.

FITZGERALD J. 2018. *Close to home. Local ties and voting radical right in Europe*. Cambridge, UK: Cambridge University Press.

GUISO L., HERRERA H., MORELLI M., SONNO T. 2017. Demand and supply of populism. *Working paper* 1703, Einaudi Institute for Economics and Finance (EIEF).

LIVI BACCI M. 2018. Elezioni politiche 2018. L'immigrazione nei programmi elettorali. *Neodemos*, 02/20/2018.

MASSEY, D. S. 1990. Social structure, household strategies, and the cumulative causation of migration. *Population Index*, Vol. 56, pp. 3–26.

MAZZIOTTA M., PARETO A. 2015. On a generalized non-compensatory composite index for measuring socio-economic phenomena. *Social Indicators Research*, Vol. 127, No. 3, pp. 983–1003.

STEPHAN W.G., STEPHAN C.W. 2000. An integrated threat theory of prejudice. In S. OSKAMP (Ed.), *Reducing prejudice and discrimination*, Mahwah, NJ: Lawrence Erlbaum Associates Publishers, pp. 23-45.

Gabriele RUIU, University of Sassari, gruiu@uniss.it
Pietro MAZZETTE, University of Sassari, p.mazzette@phd.uniss.it

A LOOK AT THE LOCAL GROWTH OF THE MAIN FOREIGN COMMUNITIES IN ITALY OVER THE LAST TWENTY YEARS[1]

Marina Amante, Massimo Mucciardi

Abstract. In the field of population studies, the process of spatial concentration tends to raise multiple and diverse issues. For this reason, the spatial settlement of foreign nationals in Italy has gained increasing interest because of its effects on social cohesion, integration programs and urban dynamics. In this context, the main purpose of this study is to assess the spatial distribution of foreign nationals in Italy in the years 2003 to 2023. Following new insights and using a population growth model, the study aims to discover patterns of concentration or dispersion of the foreign population within the different municipalities and metropolitan cities of Italy, making a temporal comparison. The analysis strategy used to achieve the aim is based on a local spatial analysis applied to the exponential development patterns, in addition to the global spatial analysis. The analysis conducted in this study finds spatial inequalities in the pattern of foreign population growth, highlighting areas with different levels of urbanization in relation to the different foreign citizenships considered.

1. Introduction

The extraordinary peculiarity of the more general migration phenomenon in Italy has been found in the country's gradual transition from an ancient place of emigration to a country of immigration since the late 1970s (Casacchia *et al.* 2022). This transition can also be explained with reference to the increasing globalization of migratory phenomena (Strozza 2019). In recent years, Italy has witnessed a significant increase in the presence of foreign communities throughout the country. Each of these profoundly nomadic and globalized-looking communities has different characteristics on multiple levels. For example, the groups are differentiated by the gender factor: some groups are formed mostly of women while in others the male component is stronger (Casacchia *et al.* 2022). That is, these are social as well as demographic differences. These foreign communities play a fundamental role both in absolute terms and in terms of impact on the total resident population. Although the complex distribution related to country of origin remains constant, significant

[1] This article is the result of a joint effort by the authors. In particular: Marina Amante has written sub-paragraph 1.1; Massimo Mucciardi has written paragraphs 2 and 3. Paragraph 1 was written jointly.

updates have occurred with reference to the size and weight of the various foreign communities, and which have been influenced by the emergence of new migration routes, the weakening of some migration flows and the steady strengthening of others. For this reason, the spatial distribution of foreign population in host societies has always attracted the interest of scholars, arousing particular attention when it takes the form of residential segregation. As pointed out in multiple research (Casacchia *et al.* 1999), attention must be paid to the so-called "push factors" with regard to the dynamics of immigration within developing countries; and regarding Italy considerable importance is attributed to "pull factors", which influence the characteristics and size of foreign communities in Italy. In fact, to discuss push and pull factors is to analyse the main motivations that drive these multiethnic groups of individuals to leave their places of origin for new destination routes. A study by Bonifazi *et al.* (2018) provides a good example. In this report, the authors highlighted what appeared to be the main motivations for several foreign communities to leave their home countries. The picture outlines a situation that seems to involve about 85 percent of migrants, for whom movement to host countries stems from: affective, familiar purposes, such as family reunification (22.7 percent); individual or family social ascent (22.1 percent); work difficulties in the country of origin (20.8 percent); and the desire for higher earnings (20.3 percent). With reference to the phenomenon of territorial settlement, a dynamic has been highlighted whereby not all foreign groups tend to settle or concentrate in the same way: in fact, as pointed out by Mucciardi *et al*. (2024), the territorial settlement of foreign nationals is related to both the length of stay and the different migration patterns followed by the groups. The diverse territorial distribution of foreign communities surely constitutes one of the characteristics that marks the various processes in which immigrant groups adapt to the host country. It can be supposed that in the initial phase the choice of where to stay depends strongly on the presence of a support network and/or an employment opportunity that allows for the realization of the initial migration project (Ferrara *et al.* 2010). At a later stage, other multiple factors eventually modify the initial choice. With reference to spatial distribution, a trend toward increasingly stable settlement was also noted (ISTAT 2022). Even so, the concept of spatial distribution with reference to foreign communities residing in countries of adoption has been repeatedly analysed in connection with the concept of spatial segregation of foreign groups. In this sense, the mechanism of spatial clustering within a host country is manifested among multiethnic groups of individuals and is an inverse process to what is assumed by a causal spatial distribution mechanism (Freeman *et al.* 1971). An interesting contribution is made, for example, by the study conducted by Massey and Denton (1988), who discuss residential segregation as a true degree in which two or more groups live separately from each other in different parts of the urban environment. What causes can this concept be traced to? Well, as urged by multiple

studies, members of foreign collectives are sometimes distributed in such a way that they are overrepresented in some areas and underrepresented in others: sometimes foreigners are spatially concentrated within a very small area, occupying a much smaller space than members of the host nation. For this reason, the two variables "space" and "population" are closely related (Voss 2007) and compose a question of interest for statistics, geography and demography. Considering the context just outlined, the main purpose of this research is to assess and map the geographical pattern (concentration or diffusion) of the ten principal foreign populations residing in Italy over a 20-year period from 2003 to 2023. The main foreign communities analysed include: Romania, Albania, Morocco, China, Ukraine, India, Philippines, Bangladesh, Egypt, Pakistan. Assuming possible peaks of concentration or spatial dispersion along the Italian territory and over the time frame examined, the study uses two exponential growth models of the national data, calculated over 10 years and 20 years, respectively. For this reason, a further aim of the research is to identify specific clusters and spatial patterns of foreign communities considering the Exponential Growth Model (EGM). To accomplish the specific objective of the research, the study proposes peculiar spatial statistical techniques for identify clusters and spatial patterns in the growth models of foreign population settlements at the municipal level. The analysis conducted in this study detects spatial inequalities in the growth model of foreign populations, highlighting areas with different levels of urbanization in relation to the different foreign citizenships considered.

1.1 The issue of settlement patterns: the Italian case

Further back in time, the Chicago school theorized a process of urban growth (Amico *et al.* 2016) characterized by pockets and waves of foreign communities taking on structures of spatial segregation. The emergence of a metropolitan-type model occurred through a phenomenon of gradual assimilation, which caused immigrant communities to lose their cultural connotations. Subsequently, several studies have shown that even in the European territorial space, the relationship between settlement patterns and indicators of social integration reveals contexts in which clustered pockets of settlement and low levels of cultural integration do not affect the entire migration movement but only certain foreign nationalities (Musterd 2003). In this landscape, studies on residential segregation have extended from the extensive literature of the Chicago School (Bailey 2012) to classical studies (Duncan and Duncan 1955 and Massey and Denton 1988). With reference to the Italian case, a growing foreign presence has developed in recent years both in absolute terms and in terms of its impact on the total population distributed from North to South

throughout the country (Benassi *et al*. 2020), as a consequence of new economic and globalizing dynamics (Borja and Castells 2002). In Italy, the main determinants of the spatial displacements of different foreign communities represents a hotly debated topic in the scientific community, where the presence among multiple settlement patterns of foreigners along the territory is discussed. In particular, the presence of three types of settlement patterns was highlighted, namely: the "metropolitan" pattern, typical of foreign communities with imbalances in their gender structure; the "diffuse" pattern typical of communities with a greater degree of dispersion among Italian provinces; the "frontier" pattern typical of communities bordering with Italy (Ferrara *et al.* 2010; Forcellati and Strozza 2006). Other research pointed out in the late 1990s that the border-type model was typical of communities of Slavic and Tunisian origin; while the diffuse model characterized Moroccans and Albanians, and finally the metropolitan model was typical of Filipino and Peruvian communities (Casacchia *et al.* 1999). Italy, then, represents one of the multiple OECD International Organization countries characterized over the past decade by the presence of an interesting share of highly skilled and tertiary educated migrants. This process, which is typical of states of recent immigration such as also Spain, Ireland and Norway, clashes with the stable quantitative consistency of low-skilled migrants (Brezzi *et al.* 2010) who then stabilize in the typical labour sectors of agriculture, construction, domestic and care services. Over 5 million foreigners live in Italy nowadays (ISTAT 2023), however they are not distributed equally across the country (Benassi *et al.* 2023). Thus, there has been a shift from an initial phase in which the greatest concentrations of foreigners were visible in areas of the North-Centre (Strozza *et al.* 2016) to a phase defined as "redistribution" (Benassi *et al.* 2023) and spatial dispersion. However, the main reason correlated with a continued and increased presence of foreign groups in Northern Italy is due to the greater availability of jobs (Mucciardi *et al.* 2021), even though the economic crisis of 2008 (Colombo and Dalla Zuanna 2019) and, more recently, the Covid-19 pandemic (Bonifazi *et al.* 2021) have led to a slow reduction of the migration phenomenon in the national territory; it is also true that despite these complex dynamics, there has been a growth in push factors for some foreign communities (Conti *et al.* 2023). Considering the present overview, the next section will be devoted to the exposition and observation of the collected data and the methodologies used for an accurate reading of the grouped data at both the national and municipal levels.

2. Data source and methodology

2.1 Data Source

As previously mentioned, when we are concerned with observing and studying the causative and derivative processes of migration phenomena, for example in Italy, the scientific community identifies as a significant trait of the foreign population its ability to move and settle outside its borders according to different spatially unequal settlement patterns in a multifactorial function. For this reason, and before analysing the obtained results, it seems necessary to examine the source and extrapolated data in order to achieve the objectives mentioned earlier. Data were provided by the institutional website of the National Institute of Statistics (ISTAT 2023), and they refer to the foreign population habitually residing in the municipalities of Italy by the years 2003, 2013 and 2023. Indeed, the use of microdata-level units of analysis, such as that relating to municipalities, allows the extrapolation of key information about the local heterogeneity of the foreign settlement phenomenon. It should be noted that the total foreign and resident population in the Italian territory, analysed with reference to Italian municipalities, was divided according to the country of citizenship. What is more, it is important to emphasize here a clear methodological note, valid, however, with respect to the application related to spatial analysis only: in fact, the territorial boundaries of the 7903 municipalities present throughout Italy as of the year 2023 were reconstructed, so that they remain stable over time and ensure proper comparison within the specified time periods. Ultimately, the present research uses the variable "foreign population" by making use of the criterion of citizenship held by subjects at the individual level: "foreigner" is, in fact, the one who has not acquired the Italian citizenship. However, we are unable to separate the impact of obtaining citizenship from the data. This may lead to slightly underestimating some foreign groups.

2.2 Methodology

Indeed, the cross-cutting character of the approach used to analyse the data is what distinguishes this study. The purpose is to provide a logical framework that identifies spatial concentration both as a demographic phenomenon and as an effectively geographic phenomenon, related to the concept of polarization (Benassi *et al.* 2023) of foreign groups. The approach that we employed is that of the Exponential Growth Model (EGM) applied to the foreign population resident in Italy in the years considered. The model explains how a population expands, given infinite supplies and no growth restrictions, at a constant rate all over time. Well, to gain a

deeper comprehension of the patterns of residential settlement an EGM is used. The model is so calculated on ten years ($EGM10$) and twenty years ($EGM20$). As a result, using the well-known formula for the exponential growth of a population, we may write for the generic foreigner population g and the generic municipality i:

$$EGM(k) = \frac{1}{\Delta t} LN\left(\frac{P_t}{P_{t-k}}\right) \quad (1)$$

where P_t is the foreigner population at time t; P_{t-k} is the foreigner population at time $t-k$ in the time span $\Delta t = P_t - P_{t-k}$. So, when we set: $t = 2023$ and $k = 10$ we obtain $EGM10$; when we set: $t = 2023$ and $k = 20$ we obtain $EGM20$.
In the first step, considering the formula (1), we calculate $EGM10$ and $EGM20$ for the entire national territory for the top ten foreign nationalities present in the year 2023 in Italy (Table 1). It is important to note that the model assumes constant growth rates, so it does not allow us to consider the effects of shocks caused by policy changes, economic conditions, and social issues. These shocks can cause significant and immediate declines in population size, which the model fails to predict. We will highlight in the next section the most important results we obtained by considering spatial cluster and outliers analysis.

Table 1 – *Total population[2], EGM20 and EGM10 for the top ten foreign communities.*

Citizenship	Total 2023	Total 2013	Total 2003	EGM20	EGM10
Romania	1081836	996526	143738	0.101	0.008
Albania	416829	483131	185933	0.040	-0.015
Morocco	415088	456890	188123	0.040	-0.010
China	307038	223405	57616	0.084	0.032
Ukraine	249613	203595	25700	0.114	0.020
Bangladesh	174058	92446	19566	0.109	0.063
India	167333	129977	33760	0.080	0.025
Philippines	158926	142607	56479	0.052	0.011
Egypt	147797	89768	29717	0.080	0.050
Pakistan	144129	84182	20232	0.098	0.054

It's important to keep in mind that, whereas geographic global statistics can determine whether spatial structure exists, they are unable to identify the locations of clusters or measure the degree to which spatial dependence changes between locations (Mucciardi 2012). Instead in a local index of spatial autocorrelation, each unit is characterized by one value of the index; it gives the individual contribution of that location in the global spatial autocorrelation measured on all n locations.

[2] The reader is warned that the data does not take into account the acquisition of citizenships.

Thus, after computing the $EGM10$ and $EGM20$, we start to calculate the Local Moran indices. Figure 1 shows the mapping of the significance of the indices by type of spatial autocorrelation, High-High (HH), Low-Low (LL), Low-High (LH) and High-Low (HL), for the main foreign communities applied to $EGM20$[3]. HH and LL mean that geographically close EGM values tend to be similar on a map: geographically close municipalities have high (H) EGM values, just as geographically close municipalities have low (L) EGM values. HL and LH mean that geographically close EGM values tend to be dissimilar on a map: in geographically close municipalities, high (H) EGM values tend to be close to low (L) EGM values and vice versa (LH)[4]. The figure shows the Local Moran indices and the spatial autocorrelation of four foreign communities habitually residing in Italy over the time considered: Romania, Albania, Egypt and China. Well, after outlining the methodological framework behind the present research, the next section will be devoted to a final discussion concerning the main implications obtained from a careful investigation about the main patterns of territorial settlement of foreigners.

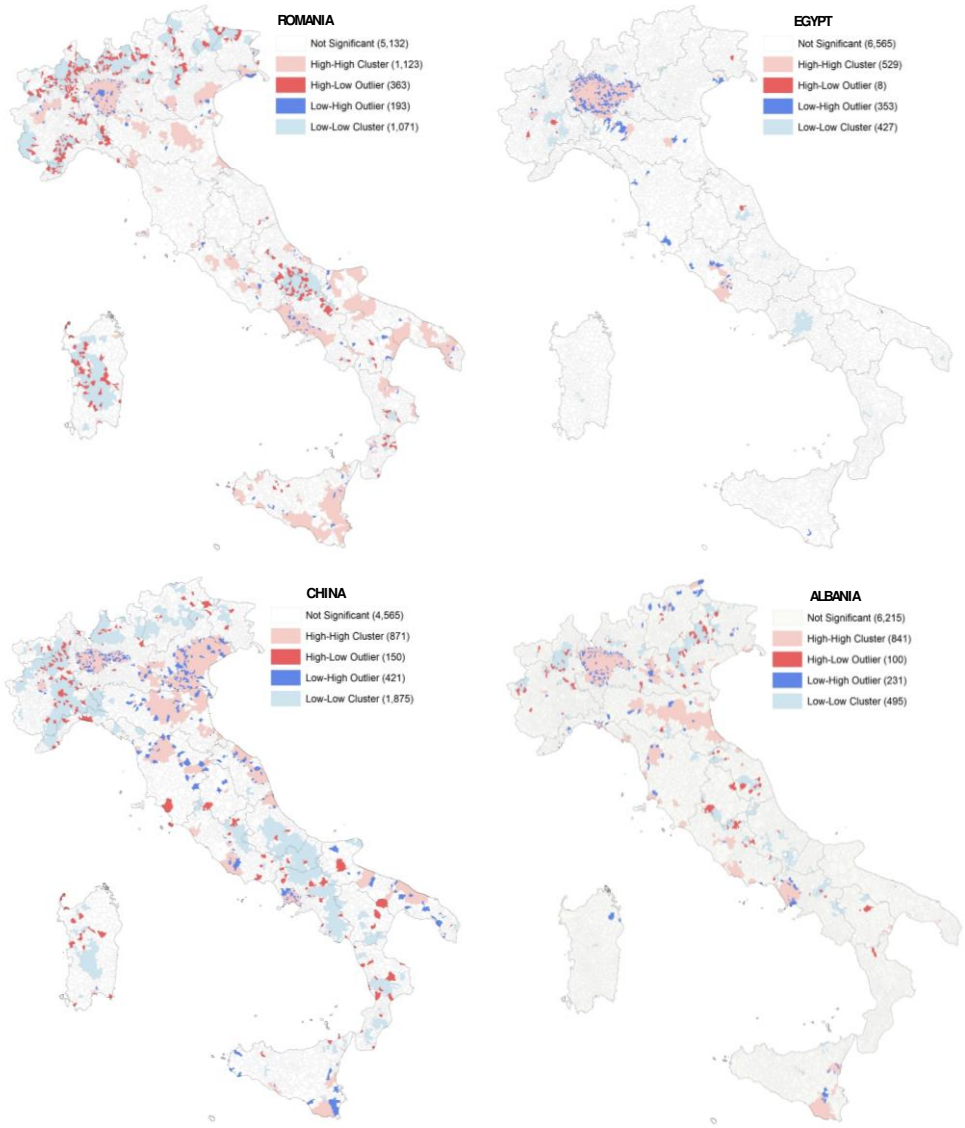
3 Discussion and final remarks

The aim of this research is to provide new keys that can explain the recent spatial settlement patterns of the main foreign communities, habitually residing in Italy in the years 2003, 2013 and 2023. Indeed, the originality of the present research, rests on the need to investigate and verify whether the increases in growth rates are spatially concentrated or whether they are uniformly distributed. The EGM approach is an excellent tool in demographic analysis, especially for examining foreign populations that are undergoing fast development as a result of immigration. Indeed, this research has deemed it appropriate to use the spatial approach to determine geographical pattern and clusters of the primary foreign communities in terms of growth. In fact, on the one hand, a global type of analysis is carried out using nationwide summary data (in absolute data): the global indices provided a summary of the temporal evolution of the main foreign communities in Italy. On the other hand, local spatial analysis techniques and the mentioned EGM are applied to municipal data. Specifically, the local spatial analysis applied to $EGM20$ at the municipal level indicates trends of stability over time for some resident foreign populations. Although the present research expands its field of observation to ten resident foreign groups, four were specifically chosen to be highlighted as representative of two precise and opposite patterns of foreign community settlement.

[3] We show maps for a select few communities due to space constraints. You may get all of the $EGM20$ maps for the ten foreign communities are available at the following link: supplementary file. $EGM10$ are available on request.
[4] See Anselin (1995) for further technical information on the four types of spatial autocorrelation.

In detail, Romania and Albania manifest a trend toward a "diffuse" type of pattern, which is also characteristic of the Moroccan population. On an opposite side, Egyptian and Chinese communities seem to reiterate a more concentrated type of territorial settlement pattern over time. In particular, for certain groups, such as the Chinese, spatial concentration becomes a real strategy positively linked to economic activities. As is also evident from the maps observed in Figure 1, the result of this mechanism is that of a phenomenon of real spatial clustering in strategic points of the nation, noting a strong concentration in Lombardy, Tuscany, Veneto, Emilia Romagna and Latium. This migration pattern, which takes into account the familiar and ethnic nature of the settlement process, results in a stable increase in the distribution of Chinese in the mentioned areas, thus reinforcing spatial patterns with Chinese enclaves (Zhou 1998). Whats more, cluster and outlier analysis shows different trends in exponential growth rates, especially in the four communities analysed. In fact, the results show that at the local level, High-High and Low-Low situations grow, so there is a growth in the level of positive spatial autocorrelation resulting in the growth of clustering phenomena along the national territory. The maps emphasize high values of the exponential growth phenomenon and high similarity (hot spots), that is, calculated at the municipal level, more inhomogeneous for the Albanian, Romanian and Chinese foreign communities. In opposition, for the Egyptian community the positive autocorrelation phenomena (hot spots) seems to link to a markedly metropolitan trend. It is true, in fact, that the maps related to clusters and outliers captured on the Italian territory, show an exponential growth especially in the North-Central regions compared to the Southern regions, probably due to attractiveness factors related to greater economic development. Despite the presence of a picture that seems to confirm what the most recent literature (Mucciardi *et al*. 2024) has already pointed out with reference to relative stable growth over time, two interesting results seem to emerge. First, there was a conspicuous increase in the exponential growth rate ($EGM20$) between 2013 and 2023 of the Ukrainian population (0,114), which is shown in Table 1. Second, the exponential growth rate calculated over 10 years ($EGM10$) is negative for the Albanian population (-0.015) and the Moroccan population (-0.010): results that are probably related to the citizenship acquisition process. However, the findings obtained must also be evaluated in light of the growth model chosen, which does not consider shock effects. Nonetheless, the results of this study provide a good starting point for investigating the drivers of the foreign settlement. Understanding the spatial positioning of the foreign groups in the Italian territory can be an excellent tool in building a policy agenda that takes into account different nationalities and related occupations. Therefore, the integration of EGM with spatial analysis has the potential to assist policy-makers in comprehending and forecasting demographic patterns.

Figure 1 – *EGM20*: *Significative maps of Local Moran (p<0.05) and cluster type (Municipalities that have non-significant statistical value are labelled as not significant).*



Spatial contiguity = inverse distance; Spatial weights = standardized; False Discovery Rate = yes

This, in turn, would facilitate more efficient planning and allocation of resources for urban areas experiencing growth. In conclusion, the present analysis is not limited to an annual snapshot of the investigated phenomenon, and the very peculiarity of the proposed methodology lies in the fact that it takes into account the 20-year growth and its spatial distribution. This crossover between the methodologies used and the different time intervals, thus enables future forecasts to be made, which are especially useful in terms of decision making in the policy sphere.

References

AMICO A., D'ALESSANDRO G., DI BENEDETTO A., NERLI BALLATI E. 2016. Lo sviluppo dei modelli insediativi: rumeni, filippini e cinesi residenti a Roma. *Cambio. Rivista Sulle Trasformazioni Sociali*, Vol. 3, No. 6, pp. 123–146. https://doi.org/10.13128/cambio-19262.

ANSELIN L. 1995. Local Indicator of Spatial Association. *Geographical Analysis*, Vol. 27, No. 2, pp. 93-115.

BAILEY N. 2012. How spatial segregation changes over time: Sorting out the sorting processes. *Environment and Planning A: Economy and Space*, Vol. 44, No. 3, pp. 705–722. https:// doi.org/10.1068/a44330.

BENASSI F., CARELLA M., MUCCIARDI M., NACCARATO A., SALVATI L. 2023. Spatial segregation in action? An empirical assessment of population concentration of foreigners and nationals in Italy, 2002-2018. *International Migration Review*. Vol. 0, No. 0. https://doi.org/10.1177/01979183231170808

BENASSI F., IGLESIAS-PASCUAL R., SALVATI L. 2020. Residential segregation and social diversification: Exploring spatial settlement patterns of foreign population in Southern European cities. *Habitat International*, Vol. 101.

BONIFAZI C., CARUSO M. G., HEINS F., PAPARUSSO A., PANACCIONE D. 2018. Le traiettorie di mobilità: arrivare e muoversi in Italia. In PEREZ M. (Ed.), *Vita e percorsi di integrazione degli immigrati in Italia*, Rome: ISTAT, pp. 35-51.

BONIFAZI C., HEINS F., LICARI F., TUCCI E. 2021. The regional dynamics of internal migration intensities in Italy. *Population, Space and Place*, Vol. 27, No. 7, pp. 1-15.

BORJA J., CASTELLS M. 2002. *La città globale.* Novara: De Agostini.

BREZZI M., DUMONT J-C., PIACENTINI M., THOREAU C. 2010. Determinants of Localization of Recent Immigrants Across OECD Regions. *Paper for OECD workshop on Migration and Regional Development*, 7 June 2010.

CASACCHIA O., DIANA P., STROZZA S. 1999. La distribuzione territoriale di alcune collettività straniere immigrate in Italia: caratteristiche e determinanti. In

BRUSA C. (Ed.), *Immigrazione straniera e multicultura nell'Italia di oggi*, Vol. 2, Milan: Franco Angeli, pp. 75-103.

CASACCHIA O., REYNAUD C., STROZZA S., TUCCI E. 2022. Internal migration patterns of foreign citizens in Italy. *International Migration*, Vol. 60, No. 5, pp. 183–197. https://doi.org/10.1111/ imig.12946.

COLOMBO A. D., DALLA ZUANNA G. 2019. Immigration Italian Style, 1977-2018. *Population and Development Review*, Vol. 45, No. 3, pp. 585-615.

CONTI C., MUCCIARDI M., SIMONE M. 2023. Exploring the Settlement Models of the Main Foreign Communities Residing in Italy (2003-2021). *Social* Sciences, Vol. 12, No. 9, p. 524. https://doi.org/ 10.3390/socsci12090524.

DUNCAN O. D., DUNCAN B. 1955. A methodological analysis of segregation indexes. *American Sociological Review*, Vol. 20, No. 2, pp. 210-217.

FERRARA R., FORCELLATI L., STROZZA S. 2010. Modelli insediativi degli immigrati stranieri in Italia. *Bollettino Della Società Geografica Italiana*, Vol. 13, No. 3, pp. 619–639.

FORCELLATI L., STROZZA S. 2006. Modelli insediativi delle comunità straniere in Italia: un quadro di sintesi. *Rivista Italiana di Economia Demografia e Statistica*, Vol. 60, No. 1-2, pp. 127-150.

FREEMAN L., PILGER J., ALEXANDER W. 1971. A Measure of Segregation Based Upon Spatial Arrangements, Honolulu, Mimeo, University of Hawaii.

ISTAT 2023. Retrieved from: http://dati.istat.it/Index.aspx?QueryId=19125

ISTAT 2022. Retrieved from: https://www.istat.it/storage/rapporto-annuale/2022/Rapporto_Annuale_2022.pdf

MASSEY D. S., DENTON N. A. 1988. The Dimensions of Residential Segregation. *Social Forces*, Vol. 67, No. 2, pp. 281–315. https://doi.org/10.1093/sf/67.2.281.

MUCCIARDI M., BERTUCCELLI P. 2012. The impact of the weight matrix on the local indicators of spatial association: an application to per-capita value added in Italy. *International Journal of Trade and Global Markets*, Vol. 5, No. 2, pp. 133-141.

MUCCIARDI M., PIRROTTA G., TOFFLE M. E. 2024. Measuring the spatial concentration of the main foreign communities residing in Italy using a new approach. *Rivista Italiana di Economia Demografia e Statistica*, Vol. 78, No. 1, pp. 142-154.

MUCCIARDI M., TOFFLE M. E., LUCCHESE G. 2021. Social inclusion of Sicily immigrants: an attempt to measure social integration. In GÜLERCE H., GIRASELLA E., SKOUFI M. (Eds.), *Migration, Social Entrepreneurship and Social Inclusion*, Naples: Editoriale Scientifica, pp. 135-155.

MUSTERD S. 2003. Segregation and Integration: A Contested Relationship. *Journal of Ethnic and Migration Studies*, Vol. 29, No. 4, pp. 623-641.

STROZZA S. 2019. Immigration and foreign nationals in Italy: Evolution, characteristics and current and future challenges. In FRIGERI D., ZUPI M. (Eds.), *From Africa to Europe. The political challenge of migrations*, Rome: Donzelli Editore, pp. 297–330.

STROZZA S., BENASSI F., FERRARI R., GALLO G. 2016. Recent demographic trends in the major Italian urban agglomerations: the role of foreigners. *Spatial Demography*, Vol. 4, No. 1, pp. 39-70.

VOSS P. R. 2007. Demography as a Spatial Social Science. *Population Research and Policy Review*, Vol. 26, No. 5, pp. 457–76.

ZHOU Y. 1998. Beyond Ethnic Enclaves: Location Strategies of Chinese Producer Service Firms in Los Angeles. *Economic Geography*, Vol. 74, No. 3, pp. 228-251.

Marina AMANTE, University of Messina, marina.amante1997@gmail.com.
Massimo MUCCIARDI, University of Messina, massimo.mucciardi@unime.it.

EXPLORING MORTALITY PATTERNS ACROSS TERRITORIAL AREAS IN ITALY

Sara Miccoli, Marco Battaglini, Francesca Licari, Marco Marsili, Cecilia Reynaud

Abstract. Mortality differentials across rural-urban gradients have long been of interest due to their implications for public health policies. Traditionally, rural areas were believed to have lower mortality thanks to healthier lifestyles and environmental factors. Recent trends suggest a narrowing gap, raising questions about the factors driving these disparities. The differences in mortality rates between territories depend on the different geographical characteristics, and this underscores the need of an analysis at a more detailed territorial level. This study aims to investigate mortality disparities and patterns across central and peripheral gradients and territorial areas in Italy, using the National Strategy for Inner Areas (SNAI) classification. We seek to understand how mortality rates vary between Inner and Central settings, focusing on individuals aged 60 and older. Our analysis includes standardized mortality rates and life expectancy at age 60, examined through spatial analysis techniques, such as Moran's Index and Local Indicators of Spatial Association (LISA). Additionally, we analyse mortality by major causes of death to identify the most significant contributors to the disparities.

The findings reveal considerable heterogeneity in mortality rates, with a notable U-shaped pattern indicating higher survival in Ultra-peripheral areas. While overall improvements in mortality levels are observed in recent years, these are less pronounced in the South. The reduction in mortality from circulatory system diseases and cancers contributes significantly to the improvements in both Poles and Inner Areas. However, higher mortality from respiratory and nervous system diseases is more prevalent in Inner Areas, especially in Ultra-peripheral municipalities.

By outlining mortality differentials and patterns in Italy, this study provides valuable insights for informing targeted health policies aimed at reducing disparities and improving overall public health outcomes for the older population.

1. Introduction

Mortality, the principal negative component of population dynamics, influences both the growth and age structure of populations. Mortality rates represent a critical measure of population health, shaping demographic patterns and informing public health policies. Variations in mortality rates are evident not only between countries but also within countries, across different regions, and between urban and rural

settings. Research has shown that rural areas often exhibit lower mortality compared to urban areas, largely due to healthier lifestyles and better environmental conditions (Woods, 2003). Conversely, urban areas benefit from advanced healthcare services, which can mitigate some of the health risks associated with urban living (Domínguez-Berjón *et al.*, 2005; Haynes and Gale, 1999). Mortality differentials across rural-urban gradients have long intrigued public health researchers due to their implications for health policy. However, recent trends suggest a narrowing gap, raising questions about the underlying factors driving these disparities (Woods, 2003; Allan *et al.*, 2017). These patterns are evident in several European countries where studies have been conducted, such as in England (Allan *et al.*, 2017) and Germany (Ebeling *et al.*, 2022).

Italy, known for its high life expectancy, has made significant strides in improving public health over the past few decades. The well-documented North-South gradient reveals notable regional disparities, with mortality rates varying significantly. In 2022, the North reported a mortality rate of 12.2 per thousand, while the South recorded 11.9 per thousand. Further disaggregation reveals even greater variability, with mortality rates ranging from 9.7 in Barletta to 16.4 in Alessandria at the provincial level. This variability is influenced by population structure, healthcare systems, and other factors (Domínguez-Berjón *et al.*, 2005). In Italy, the National Health Service is a regionally based public system and aims to offer free or low-cost health services, ensuring all people equal access to health care regardless of income. The historical socio-economic divide between North and South also influences health outcomes: Northern regions generally have better health infrastructure and higher income levels than Southern regions.

At a broader territorial level, Caselli et al. (2021) analysed mortality trends in Italian regions, providing a comprehensive picture of the vulnerabilities that emerged during the recent pandemic and outlining possible developments for the near future. Cavalieri and Ferrante (2020), using a Beta-convergence approach, analysed the influence of fiscal decentralisation on the convergence dynamics of infant mortality rates and life expectancy at birth among Italian regions over the period 1996-2016.

Despite the progress in the analysis of mortality in understanding its territorial determinants, there is a gap in the research on the spatial variability of mortality within the country, especially at lower territorial level.

The differences in mortality rates depend closely on the different geographical characteristics of the areas, and this underscores the necessity of an analysis at a more detailed territorial level to understand the different patterns of mortality. This study aims to fill the gap in the existing literature by providing a detailed analysis of mortality disparities across different territorial areas in Italy, considering the

National Strategy for Inner Areas (SNAI) classification[1]. In Italy, this classification of municipalities provides a valuable framework for the analysis of the spatial variability in mortality. The classification distinguishes areas based on their distance to essential services (in terms of mobility, health, and education), which are critical factors influencing mortality (Istat, 2024).

Our research hypothesizes that the differences between the types of areas defined by the SNAI classification are as significant as those between broader geographical macro-areas (North, Centre, South[2]). In particular, we hypothesize that living in a place more or less distant from essential services is more influential in determining mortality rates than the traditional advantages associated with the rural lifestyle. To explore this hypothesis, we focus on mortality among individuals aged 60 and older, a population group that records higher mortality and is most affected by variations in service availability and quality.

The aim of this research is to analyse mortality among these individuals ‘at the municipal level, investigating differences between Inner and Central areas over time. We aim to determine how mortality differs between these areas, and whether this gradient is confirmed at a broader geographical level (North, Centre, South). Additionally, we seek to identify which gradient is stronger: that between Inner and Central areas, or that between Central-North and South regions. Furthermore, in order to identify the underlying factors contributing to the differences, we will explore the specific causes of death that are most responsible for survival differences at the Central/Inner level. In doing so, we want to offer insights that can inform targeted public health policies aimed at reducing mortality disparities and improving health outcomes for the old population in Italy.

2. Data and Methods

This study utilizes mortality data during the last decade sourced from the Italian National Institute of Statistics (Istat). The analysis covers Italian municipalities, divided into broader macro geographical areas (North, Centre, South). The municipalities are also classified according to the SNAI classification, which identifies areas based on the distance to essential services such as health, education, and mobility. This classification divides municipalities into: Central Areas, which in turn are divided into six categories: *Poles* (A) and *Intermunicipal Poles* (B), which

[1] According to the greater distance to essential services, SNAI classifies municipalities into Central (Poles, Intermunicipal poles and Belts) and Inner Areas (Intermediate, Peripheral and Ultra-peripheral municipalities). See also Data and Methods paragraph.

[2] These areas are obtained by grouping the five major socio-economic regions (NUTS1 under Eurostat classification): North-West with North-East, and South with Islands.

offer all three essential services; *Belt municipalities* (C), with good access to Poles; and Inner Areas, further categorized into *Intermediate* (D), *Peripheral* (E), and *Ultra-peripheral* (F) areas, based on their increasing distance from essential services.

The primary focus is on individuals aged 60 and older, a demographic group that typically exhibits higher mortality rates and more significant health disparities.

The first indicator we consider is the standardized mortality rate for individuals aged 60 and above from 2011 to 2013 and from 2021 to 2023, calculated for each municipality using the age structure of the average Italian population in 2017 (central year to the periods considered) as the standard population. To mitigate issues related to small data sets or anomalous years, we consider three-year periods together, ensuring more robust and reliable results. To analyse spatial variability, we apply geographical analysis techniques: we calculate the Moran's Index to measure spatial autocorrelation and the local Moran's I statistic, a Local Indicator of Spatial Association (LISA), using a second-order Queen contiguity matrix. This helps us understand the degree to which mortality rates are clustered or dispersed across different areas, providing insight into spatial patterns and the potential influence of local factors.

The second indicator we analyse is life expectancy[3], calculated using the official methodology (Istat, 2001). In order to overcome the problem of possible underreporting of deaths in the smaller parts of the country (particularly in the Inner Areas of central Italy), we develop mortality tables using a two-year database. We compare life expectancy over two-years periods (2011-2012 and 2022-2023, to avoid the influence of pandemic COVID-19 on survival function estimates) by SNAI classification and major socio-economic regions.

Finally, we consider the standardized mortality rates for individuals aged 60 and more distinguished by major causes of death[4] in the years 2011 and 2019. We focus on the 60+ age group, because of the insignificance of mortality due to accidental causes from the age of 60 onwards, and because, from this age, the weight and time course of certain chronic degenerative diseases (e.g. cancer and cardiovascular diseases) are relevant. Analyses were conducted using annual age classes. Applying the composition method introduced by Pollard (1982), we identify which group of causes of death contributes most significantly to the differences in mortality between areas.

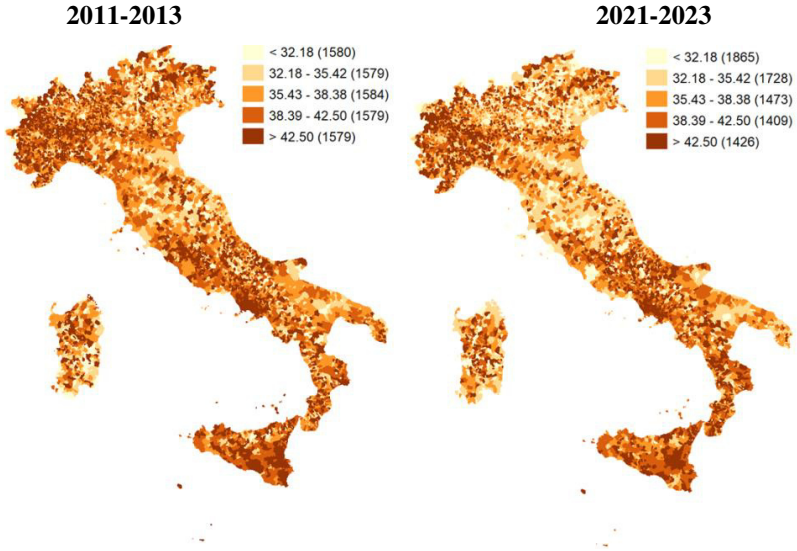
[3] Life expectancy is a statistical measure of the average time someone is expected to live, based on the year of birth, current age and various demographic factors.

[4] In our analisys we grouped the main causes of deaths into: Circulatory system disease; Diseases of the digestive and genitourinary systems; Cancers; Infective and respiratory tract diseases; Endocrine and blood diseases; Degenerative diseases of the nervous system; Other.

3. Results

The standardized mortality rates for individuals aged 60 and older, as shown in Figure 1, exhibit considerable heterogeneity both in 2011-2013 and 2021-2023. Higher mortality rates are observed in the North-West and in some Southern regions, particularly in Campania, in both periods. However, there are also notable high mortality areas in the Alpine regions of the North-East and in the Apennines. In some areas, a decrease in mortality levels seems to appear between the two periods.

Figure 1 – *Standardized mortality rates (60+), Italian municipalities, periods 2011-2013 and 2021-2023.*

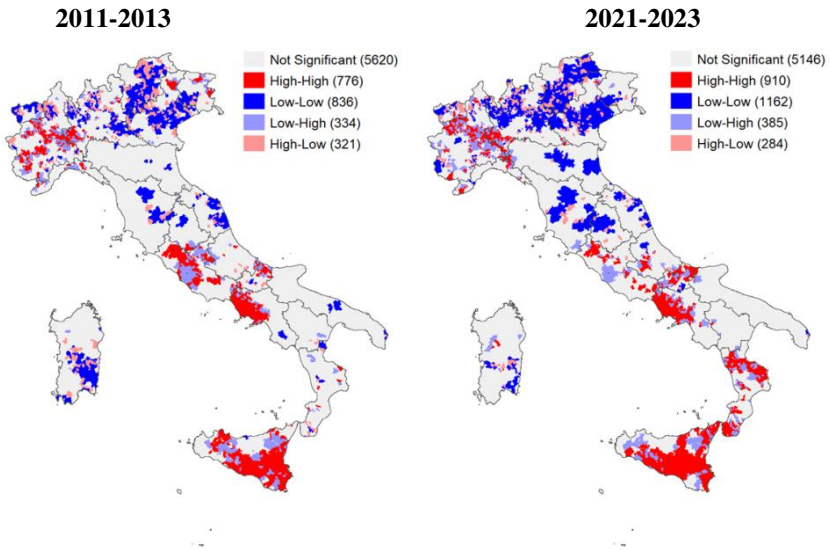


Source: our elaboration on Istat data.

The spatial analysis allowed exploring the geographical distribution of mortality rates. The Moran's Index, computed for both periods, is positive and statistically significant: 0.10 for the first period, and 0.13 for the second, with pseudo p-value smaller than 0.001. This indicates a positive global spatial autocorrelation in mortality rates, suggesting that municipalities with high mortality rates are often located near areas with similarly high mortality rates, and confirming a certain homogeneity of the phenomenon across the territory. The existence of homogeneous spatial clusters of municipalities is further confirmed by the Local Indicators of Spatial Association (LISA), which allow to see where exactly the clusters are located and to check for other types of clusters. The LISA results (showed in the LISA cluster maps in Figure 2) reveal high-high clusters in the South, particularly in Campania

and Sicily, as well as some smaller clusters in the North-West. Interestingly, there is also a high-high cluster in Lazio during the first period that disappears in the second period, suggesting that some changes over the years are consistent with improved health systems and living conditions. Conversely, low-low clusters are observed in the North-East and Central regions emphasizing the lower impact of mortality in these territories. These spatial patterns highlight the existence of significant regional disparities while high-high clusters in the South suggest persistent health disadvantages in these regions.

Figure 2 – *Local indicators of spatial autocorrelation (LISA) cluster map: standardized mortality rates (60+), Italian municipalities, periods 2011-2013 and 2021-2023.*



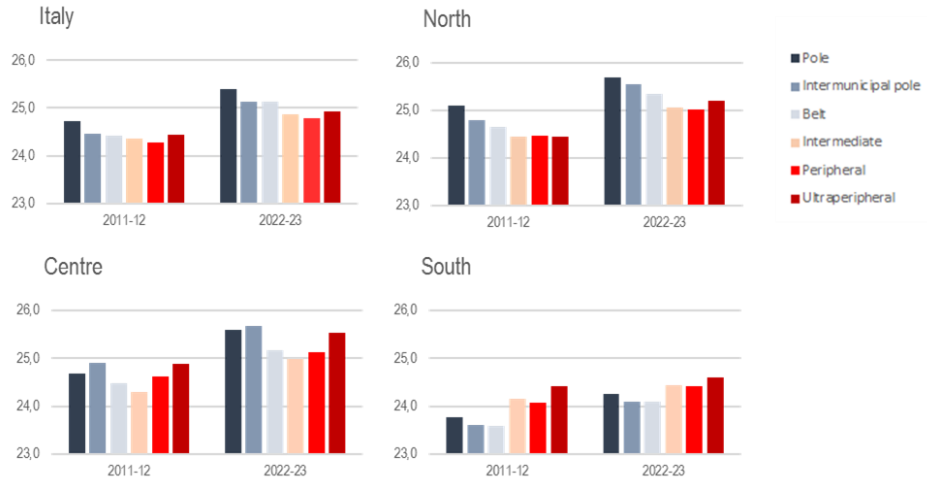
Source: our elaboration on Istat data.

Furthermore, we examined differences in life expectancy at age 60 across the different SNAI classifications (Figure 3), confirming an overall improvement in mortality levels between the period 2011-2012 and 2022-2023[5]. However, this improvement is less pronounced in the South. Life expectancy at age 60 shows a decreasing trend moving from Central to peripheral areas, with a notable U-shaped pattern indicating higher survival in Ultra-peripheral areas. This pattern is consistent across geographical divisions, although the contrast between North and South is particularly striking. In Italy as a whole and in the North, life expectancy at age 60

[5] In the life expectancy analysis, we consider two-year periods (2011-2012 and 2022-2023) to avoid the influence of pandemic COVID-19 on survival function estimates.

is highest in Poles and decreases progressively as the distance from these urban centres increases. Notably, life expectancy in Ultra-peripheral areas is higher than in Intermediate and Peripheral areas. In the Centre, Intermunicipal Poles exhibit the highest life expectancy, closely followed by Poles. Similarly to the North, Ultra-peripheral areas in the Centre have higher life expectancy than Intermediate and Peripheral areas. In contrast, in the South, all municipalities in the Inner Areas have higher life expectancy than Poles. This suggests that Inner areas may offer a better quality of life. In the North, this advantage could be offset by the availability of more efficient services in Poles and Belt municipalities. In the Centre, the same pattern holds for the first two types of areas. In the South, however, the limited availability or lower quality of services in Central areas may be insufficient to counterbalance the lower quality of life in cities.

Figure 3 – *Life expectancy at 60 years by SNAI classification, years 2011-2012 and 2022-2023.*



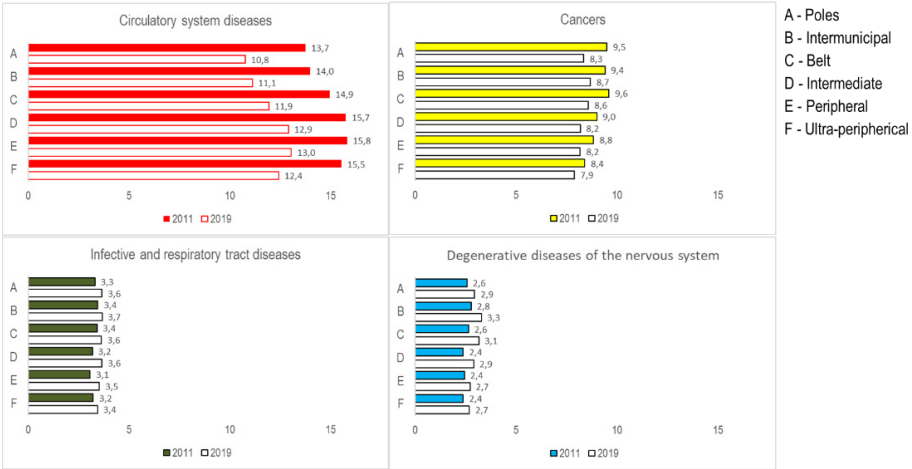
Source: our elaboration on Istat data.

To further understand the differences in mortality, we analysed the standardized mortality rates for individuals aged 60 and more, focusing on four primary groups of causes: circulatory system diseases, cancer, infectious and respiratory tract diseases, and degenerative diseases of the nervous system. This analysis allowed to identify the causes of death contributing most significantly to the mortality disparities between different types of areas within the SNAI classifications.

Diseases of the circulatory system show higher mortality rates in the Inner Areas, consistent with the importance of the immediate availability of health services for

these conditions. Notably, Ultra-peripheral areas exhibit slightly lower mortality rates from circulatory system diseases compared to Intermediate and Peripheral areas (Figure 4). This suggests that healthier lifestyles, potentially related to types of work and other factors, may prevail in these remote areas. Despite a general decrease in mortality rates from circulatory system diseases between the two periods, the relative differences between the types of municipalities remain consistent. For cancer, mortality rates decrease progressively as municipalities are located farther from the Poles (Figure 4). This trend may be attributed to the fact that cancer is heavily influenced by lifestyle and external conditions. Central areas, which tend to have higher exposure to risk factors such as pollution and stress, exhibit higher cancer mortality rates compared to more peripherical areas. Between the two periods, a decrease in cancer mortality rates is observed, yet the pattern of higher rates in Central areas compared to the Inner ones remains evident. Infective and respiratory tract diseases, as well as degenerative diseases of the nervous system, show an increase in mortality rates between the two periods. This rise is consistent with the emergence of new diseases that have become significant causes of death in recent years. For these types of mortality, lower rates are recorded in the three types of Inner Areas, particularly in Ultra-peripheral areas. This pattern may suggest that the more remote areas are less affected by the conditions that exacerbate these diseases or benefit from healthier environments.

Figure 4 – *Standardized mortality rates for some main groups of causes of death at 60 years and more, years 2011 and 2019.*

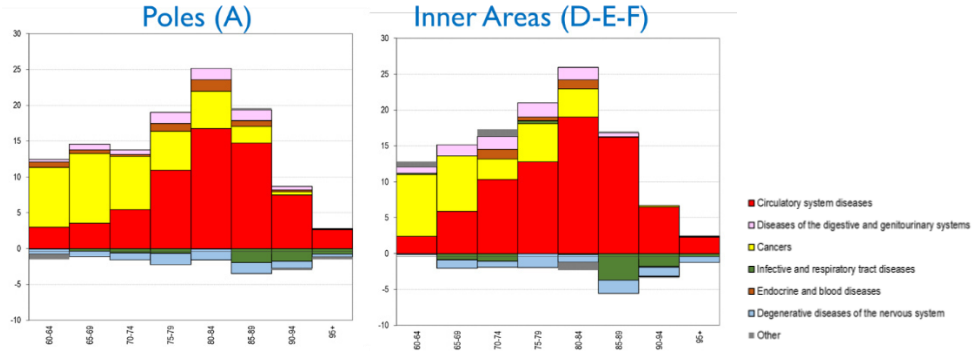


Source: our elaboration on Istat data.

These findings underscore the significant role of healthcare access in managing circulatory system diseases and suggest that healthier lifestyles in Ultra-peripheral areas may contribute to slightly lower mortality rates for these diseases. The persistence of these differences over time highlights the need for targeted public health interventions to address the needs of each area type. Additionally, the trends in cancer mortality emphasize the importance of lifestyle and environmental factors in influencing health outcomes. The rising mortality rates from infectious and respiratory tract diseases and degenerative diseases of the nervous system point to the need for ongoing public health vigilance and targeted strategies to mitigate these emerging health threats.

As observed in Figure 5, the decomposition of the change in life expectancy at age 60 by main groups of causes of death and age classes, from 2011 to 2019[6], shows that most causes of death contribute to an increase in life expectancy, especially up to age 90. The reduction in mortality from circulatory system diseases and cancers accounts for the most significant improvements in both Poles and Inner Areas. Specifically, 65% of the change in Poles comes from improvements in circulatory system diseases, and 39% from cancers. In Inner Areas, these figures are 75% and 29%, respectively. This is another confirmation of the importance of healthcare facilities, on one hand, and of healthy lifestyle, on the other hand. Additionally, it is notable that mortality from respiratory and nervous system diseases has a greater impact in Inner Areas, reflecting higher vulnerability to these conditions in more remote regions, where certain health services are absent and distant.

Figure 5 – *Decomposition of the change in life expectancy at 60 years by main groups of causes of death and age classes, from 2011 to 2019.*



Source: our elaboration on Istat data.

[6] Also in the decomposition of the change in life expectancy by groups of causes of death, we do not consider the last years (2020-2021) to avoid the effect of the pandemic on causes of death.

4. Discussion and conclusion

From this descriptive analysis, it emerges that Italy, one of the longest-lived countries in the world, is characterized by significant heterogeneity at a territorial level. The SNAI classification proves suitable for studying these territorial differences, suggesting that lifestyles, environmental factors and the presence of social and health services play a fundamental role in shaping mortality patterns.

In the general context of lower mortality in the Centre-North compared to the South, Central and Inner Areas exhibit different survival levels. At a national level, survival decreases moving from the Centres to the periphery and then rises in the Ultra-peripheral areas, highlighting a U-shaped pattern of life expectancy. This U-shaped pattern is preserved across geographical divisions, though it accentuates the contrasting situation between the North and South. Indeed, in the latter, all the Inner Areas show a more favorable situation.

The analysis by cause of death further emphasizes these contrasts. Circulatory system diseases have higher mortality rates in Inner Areas, underscoring the critical role of healthcare services in managing these conditions. However, Ultra-peripheral areas exhibit slightly lower rates, suggesting healthier lifestyles may mitigate some risks. Despite a general decrease in mortality rates from circulatory diseases over time, the differences between municipal types remain consistent. Cancer mortality rates decrease progressively with distance from Central areas, suggesting the influence of lifestyle and environmental factors on cancer. Poles, with higher exposure to risk factors such as pollution and stress, show higher cancer mortality rates compared to Inner areas, where the rhythms of life can be less stressful and environmental conditions tend to be better. The general decrease in cancer mortality over time does not alter the pattern of higher rates in the Central Areas compared to peripheral ones. The rise in mortality rates from infectious and respiratory tract diseases and degenerative diseases of the nervous system between the two periods seems to reflect the emergence of new significant health threats. With continued advancement in life expectancy and a growing old population, these threats need to be addressed urgently, including taking into account the different access to health care facilities in different areas. Lower rates in the three types of Inner Areas, particularly Ultra-peripheral areas, suggest that these areas are less affected by conditions exacerbating these diseases or benefit from healthier environments. Analysis of the living and environmental conditions in different territories can be an important piece in a framework aimed at understanding the mechanisms behind these diseases and their treatment.

The findings suggest that while Inner areas in the North and Centre benefit from better services in Poles and Belt municipalities, the South does not have sufficient services to counterbalance the lower quality of life in Central areas. The significant

role of healthcare access in managing circulatory diseases and the healthier lifestyles in Ultra-peripheral areas highlight the complexity of public health interventions. Indeed, the analysis would seem to suggest the necessity to invest in public health and ensure easy access to health care facilities as much as to dedicate proper attention to promoting healthy lifestyles. Addressing the territorial disparities requires targeted strategies that consider the needs of each area type, aiming to reduce mortality disparities and improve health outcomes for the older population across Italy.

In summary, this study underscores the necessity of analysing spatial variability in mortality at a detailed territorial level. Such an approach provides insights for informing public health policies aimed at reducing mortality disparities and enhancing health outcomes, particularly for the older population in Italy. The persistence of differences in mortality rates and the impact of emerging health threats highlight the ongoing need for vigilant public health strategies and interventions tailored to the specific characteristics of different regions.

References

ALLAN, A., et al. 2017. Unravelling urban-rural health disparities in England, *Population Space and Place*, Vol. 23, No. 4, pp. 201-217.

ASTROM, S., et al. 2018. Exploring Spatial Disparities in Health Outcomes, *Demographic Research*, Vol. 38, No. 10, pp. 291-318.

CASELLI, G., EGIDI, V. and STROZZA, C. 2021. *L'Italia longeva: dinamiche e diseguaglianze della sopravvivenza a cavallo di due secoli*. Il mulino.

CAVALIERI, M., FERRANTE, L. 2020. Convergence, decentralization and spatial effects: An analysis of Italian regional health outcomes, *Health Policy* 124.2, pp. 164–173.

DOMÍNGUEZ-BERJÓN, M. F., BORRELL, C., LÓPEZ, R., PASTOR, V. 2005. Mortality and socioeconomic deprivation in census tracts of an urban setting in Southern Europe, *Journal of Urban Health*, Vol. 82, No. 2, pp. 225-235.

EBELING, M., et al. 2022. Socioeconomic Deprivation and Health: Mapping the Neighbourhood, *Journal of Epidemiology and Community Health*, Vol. 76, No. 5, pp. 512-520.

GARTNER, A., FAREWELL, D., ROACH, P., DUNSTAN, F. 2011. Rural/urban mortality differences in England and Wales and the effect of deprivation adjustment, *Social Science & Medicine*, Vol. 72, No. 10, pp. 1685-1694.

GHEASI, M., et al. 2019. Spatial Inequality in Health and Access to Healthcare, *Letters in Spatial Resource Science*, Vol. 12, No. 3, pp. 329-349.

HAYNES, R., GALE, N. 1999. Socioeconomic Deprivation, Health Status, and Lay Knowledge, *Health & Place*, Vol. 5, No. 4, pp. 279-290.
ISTAT 2001 Tavole di mortalità della popolazione italiana per provinciale regione di residenza. Anno 1998.
ISTAT 2024. La demografia delle Aree Interne: dinamiche recenti e prospettive future, *Statistiche Focus 29 luglio 2024*.
LAW, M., MORRIS, R. 1998. Socioeconomic Deprivation and Health: Mapping the Neighbourhood, *Journal of Epidemiology & Community Health*, Vol. 52, No. 3, pp. 225-231.
POLLARD, J., 1982. The expectation of life and its relationship to mortality, *Journal of the Institute of Actuaries* , Volume 109 , Issue 2 , September 1982, pp. 225 - 240.
WOODS, R. 2003. Urban-Rural Mortality Differentials: An Unresolved Debate, *Population and Development Review*, Vol. 29, No. 3, pp. 405-421.

Sara MICCOLI, Istat, sara.miccoli@istat.it
Marco BATTAGLINI, Istat, battagli@istat.it
Francesca LICARI, Istat, licari@istat.it
Marco MARSILI, Istat, marsili@istat.it
Cecilia REYNAUD, Università Roma Tre, cecilia.reynaud@uniroma3.it

FRAGILITY AND ANTI-FRAGILITY OF ITALIAN VOLUNTEERING IN THE POLYCRITICAL CONTEXT

Umberto Di Maggio, Stefania Della Queva, Manuela Nicosia, Giuseppe Notarstefano, Sabrina Stoppiello, Giuseppe Terzo

Abstract. Volunteering is a multifaceted phenomenon of great relevance to the social sciences. Several studies confirm how it significantly contributes to mitigating and partially alleviating the effects of crises (Guo et al., 2021; Della Porta, 2020) and, more generally, the distorting effects of capitalism and the market (Rago & Venturi, 2020). Despite the fact that crises increase the range of social needs to which institutional welfare alone cannot respond, it is noted that Italian volunteering has changed during the Covid pandemic, experiencing a crisis due to a reduced capacity to involve and activate individuals willing to engage in activities of general interest (Eurispes, 2022).
In 2021, there were about 360,000 non-profit institutions (NPIs) active in Italy and about 261,000 with volunteers (76.1% of the total NPIs). The aim of this paper is to describe the different forms of "organised volunteering" and to analyse specific contexts in order to identify the factors that, at territorial level, could influence the traditional way of volunteering and allow new forms of engagement and civic participation to emerge.
Specifically, we analyse the data collected by the permanent censuses of non-profit institutions in order to classify NPIs according to structural characteristics, human resources, activities carried out, and other qualitative information such as scope (mutual versus public benefit), target of mission (individuals versus community), network of stakeholders, and category of vulnerability addressed. This analysis makes it possible to identify different types of volunteering and their changes over time, providing the empirical basis for new reflections on the fragility (or not) of organised volunteering in Italy.

1. Introduction

Volunteering is an activity undertaken spontaneously by individuals without monetary remuneration (Fiorillo & Nappo, 2015). In our work, we focus on volunteering in nonprofit institutions (NPIs), which we define as organized volunteering. It is particularly important to analyse this type of volunteering because it serves as a key resource of labour and initiative for the nonprofit sector (Beatton & Torgler, 2018). Organized volunteering significantly contributes to the functioning of NPIs, enabling them to expand their service offerings and reach more beneficiaries. The literature has long addressed this topic using both qualitative and quantitative approaches to examine its evolution in Italy over the past decades.

Building on sociological and political studies (Ranci, 2006; Biorcio & Vitale, 2016; Ambrosini & Erminio, 2020; Caltabiano et al., 2024), we also aim to investigate this phenomenon from a statistical and economic perspective to better understand the internal dynamics of NPIs and underscore their critical role in promoting territorial well-being, as highlighted by a growing body of literature (Terzo, 2021, 2022; Terzo et al., 2023, 2024).

Based on these premises, the aim of our work is to investigate the evolution of organised volunteering between 2015 and 2021, using data from the ISTAT census of non-profit institutions. Through this analysis, we intend to identify and describe emerging patterns, paying particular attention to the dynamics of growth, changes in the types of activities carried out, and the geographical distribution of volunteers. This study will enable us to better understand the changes taking place in the non-profit sector and to provide useful recommendations for improving policies to support volunteering, with the ultimate aim of promoting the well-being and sustainable development of communities and regions.

The rest of the paper is structured as follows. The next section presents a descriptive analysis of volunteering in NPIs using the 2021 data from the Istat permanent censuses of non-profit institutions, highlighting the main changes compared to the previous census in 2015. Section 3 presents the results of a shift-share analysis to examine changes in volunteering across regions and sectors between 2015 and 2021. Finally, section 4 concludes.

2. The Nonprofit Institutions Census and data on volunteering

2.1 The Nonprofit Institutions Census

Istat carried out the censuses of non-profit institutions according to a definition of NPIs contained in the System of National Accounts (1993 and 2009) and used by the Handbook on Nonprofit Institutions (United Nations, 2003), which complies with the basic requirement of prohibiting the distribution of profits (SNA 1993, par. 4.54). The strategy for the permanent census of non-profit institutions is based on two key elements: the Statistical Register and the sample survey, conducted periodically, to gather information also on thematic issues and to conduct time series analysis. The statistical register, updated annually, is the sampling frame for the survey. The sample survey ensures: a) an integration of the information content of the statistical register through thematic focus; and b) validation of models to estimate eligibility and main variables of the units included into the statistical register.

In 2022, Istat carried out the second permanent census of NPIs, with a sample frame, which include 110,000 units (30% of units gathered by the Statistical Register of NPIs). In April 2024 Istat disseminated the first results of the Permanent Census of non-profit Institutions, based on NPIs Register updated on 31 December 2021, reference period of sample survey.

Collected data allow to improve the measurement of the contribution of organised volunteering to the enlargement and strengthening of civic society participation. Census surveys have always collected information on volunteers operating in NPIs and the definitions adopted are coherent with those adopted by the ILO, by the Handbook on Nonprofit institutions, and with Third Sector Reform legislation, as well.

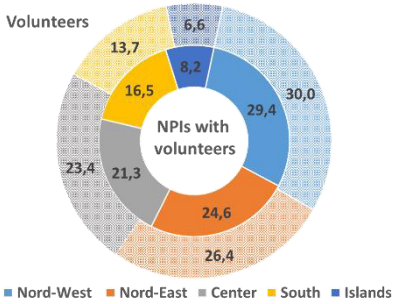
2.2 The organised volunteering in 2021: main characteristics

The Second Permanent Census of NPIs allowed to update the data on non-profit volunteering: in 2021, the 71.3% of NPIs operated with the contribution of 4.617 million of volunteers. Italian volunteers represent a crucial resource for the non-profit sector, contributing significantly to addressing vulnerability and mitigating the effects of the Covid-19 pandemic emergency. The 71.3% of total active NPIs in Italy operated with 4.662 million of volunteers. More than half of total number of volunteers (58.3%) involved in non-profit sector are males, the 41.7% are females. The presence of volunteers is more consistent in areas of Northern Italy (56.4%), compared to the Centre (23.4%), the South (13.7%) and the Islands (6.6%) (Figure 1).

Even considering the presence of volunteers in relation to the resident population, the Northern and Central regions show higher figures than the national average (equal to 790 per 10,000 inhabitants), with 1,165 volunteers per 10,000 inhabitants in the Northeast, 892 in the Centre, and 887 in the Northwest. The South and the Islands show 492 and 509 volunteers per 10,000 inhabitants, respectively.

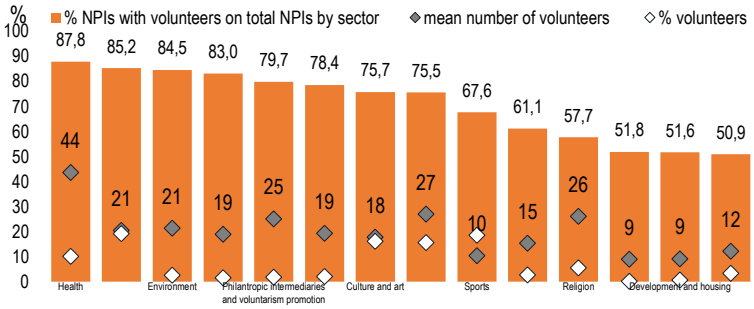
The first three activity sectors, Culture, sport, and recreation, gather the 53.8% of total number of volunteers (Figure 2). NPIs active in Social Assistance and Civil Protection share the 15.6 % of volunteers, Health sector the 10.1%, Religion the 5.6%. NPIs with the largest organisational structure are those active in the Health sector, with 44 volunteers on average per institution (the national figure is 18 volunteers per NPI). The Social Assistance and Civil Protection follow this sector, with 27 volunteers per institution, the Religion sector (26), and the Philanthropy and Promotion of Volunteerism (25).

Figure 1 - *NPIs with volunteers and volunteers by macro regions. Year 2021, percentage values.*



Source: Our data processing from Istat – Non-profit Institutions Census.

Figure 2 - *NPIs with volunteers on total NPIs of sector, volunteers by activity sector, mean number of volunteers by activity sector. Year 2021, percentage values.*

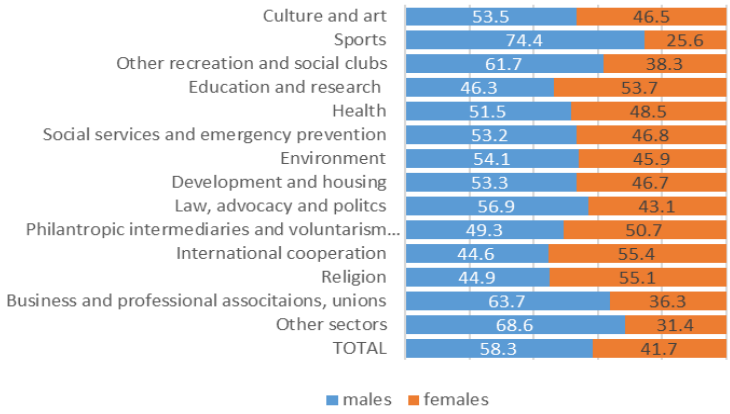


Source: Our data processing from Istat – Nonprofit Institutions Census.

The activity sectors with the highest incidence of female volunteers are Cooperation and International Solidarity and Religion (with 55 volunteers per 100 volunteers); Education and Research (with 53.7%), Philanthropy and Promotion of Volunteering (50.7%), and Health (48.5%) (Figure 3).

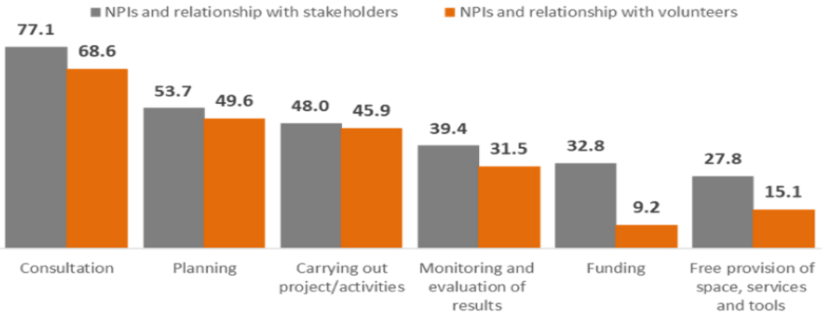
The 56.6% of NPIs with volunteers have significant relationships with them, compared with 46.8% of total NPIs (Figure 4). More than two-thirds of NPIs consult their volunteers in defining activities (68.6 %). Under half involve them in planning (49.7%) and implementing projects (45.9%). One-third of NPIs involve their volunteers in monitoring and evaluation of their activities.

Figure 3 - *Volunteers by activity sector and gender. Year 2021, percentage values.*



Source: Our data processing from Istat – Nonprofit Institutions Census.

Figure 4 - *NPIs and NPIs with volunteers by type of stakeholder involvement. Year 2021, percentage values.*



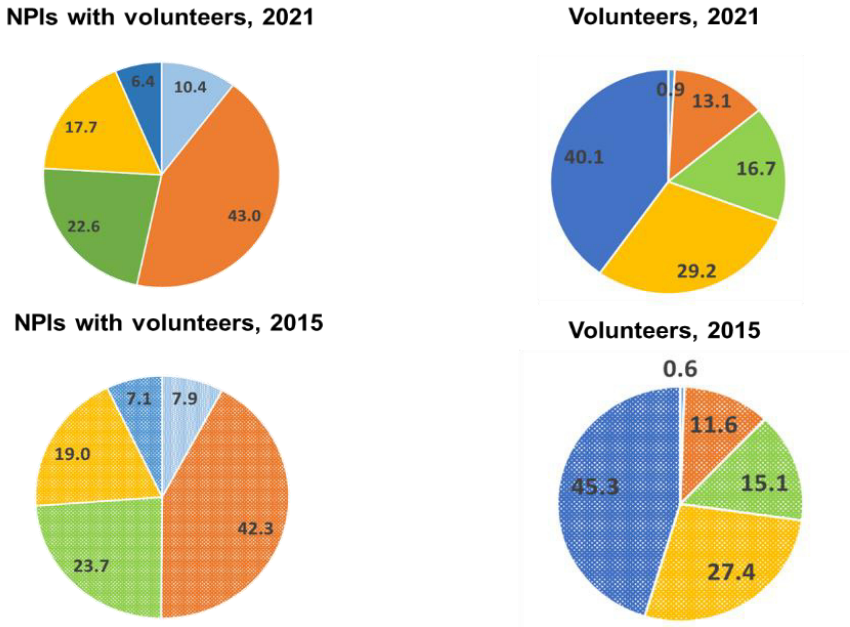
Source: Our data processing from Istat – Nonprofit Institutions Census.

2.3 The dynamics of nonprofit volunteerism through time

The analysis we carried out highlights a consistent decrease of the number of volunteers active in the non-profit sector. As manner of fact, the comparison with the last census (which refers to year 2015) shows a 15.7% of decrease (in 2015 active volunteers were 5.528 million). Inevitably, the number of NPIs with volunteers decrease as well: in 2015, the quota was equal to 79.6%, reporting a decrease of 16.5 %.

Looking at the structure of NPIs, which rely on the contribution of volunteers, allows to better describe how non-profit volunteering is changing. NPIs with volunteers range in size from extremely large to extremely small. The first evidence concerns a downsizing of the sector, which seems to be taking place. More specifically, in 2021, the share of medium and large NPIs (units with more than 20 volunteers) decreases (from the 26.1% in 2015 to 24.1% in 2021), while volunteering small-sized NPIs grows (Figure 5). Compared to 2015, NPIs with 1 or 2 volunteers increase from 7.9% to 10.4%, and units with 3-9 volunteers grow by 0.7% (from 42.3% to 43%). Moreover, the share of volunteers of medium-sized institutions grows slightly (29.7% volunteers compared to 27.4% in 2015). Large NPIs also downsize organizationally, from 130 volunteers on average in 2015 to 111 volunteers in 2021.

Figure 5 - *Nonprofit institutions with volunteers and Volunteers by size. Year 2021 and 2015 (percentage values).*



Source: Our data processing from Istat – Nonprofit Institutions Census.

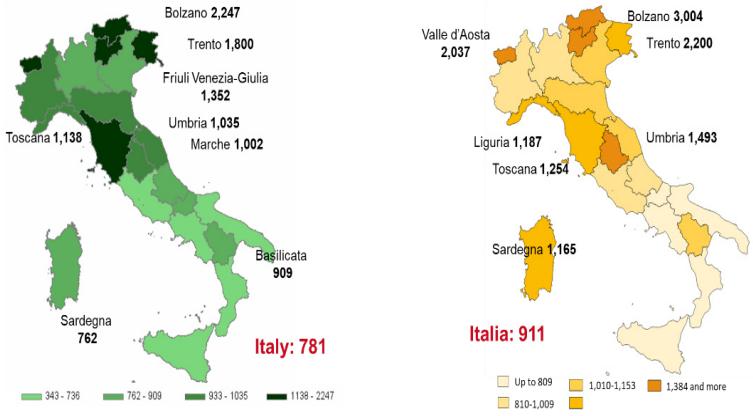
The trend just described can be explained by considering the period which the data collected refer to, which is when Covid-19 emergency was still ongoing. Therefore, is important to consider that the restrictions and the social distancing imposed affected the activities of both organisations and individuals. If some organisations certainly suffered from the impossibility of carrying out their activities

(which in some cases were completely suspended), some others have intensified and have reshaped themselves in order to cope with the crisis and to ensure the support for the weakest and most vulnerable segments of society. In this context, the contribution of volunteers has been relevant in addressing the vulnerabilities and hardships that have arisen because of the health emergency.

By looking at the gender of volunteers, it can be noted that the composition of female volunteers and males engaged in the non-profit sector is similar to the one described in 2015. In fact, the decrease is evident for both categories, although lower on females (-17.3% for men, -15.3% for women).

As for the regional distribution of volunteering, the analysis highlights that, although we can observe the decrease of volunteering in all areas of Italy, the share of volunteers in the Southern and North-Western regions is slightly higher if compared to 2015 figures. Nevertheless, by comparing the data through the two years, the ratio of volunteers to local population (which the figure 6 refers to) shows the same picture in terms of distribution of volunteers in Italian regions.

Figure 6 - *Non-profit volunteers per 10,000 inhabitants. Year 2021 and 2015.*



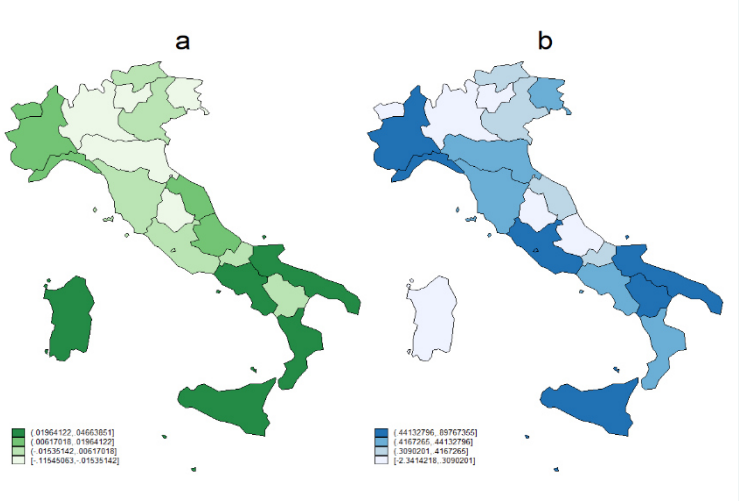
Source: Our data processing from Istat – Nonprofit Institutions Census.

3. Understanding changes in organised volunteering between 2015 and 2021: some empirical insights

In this section, we present the results of an analysis using the standard shift-share methodology (Dunn, 1960; Esteban-Marquillas, 1972). This approach is key to understanding the dynamics of volunteer distribution within the non-profit sector between 2015 and 2021, disaggregated by region (NUTS-2) and sector of activity per the International Classification of Non-Profit Organizations (ICNPO). Originally

developed for regional economic analysis, shift-share has been adapted to identify variations in volunteering patterns due to regional and sectoral characteristics. Shift-share decomposes changes in volunteering into three components: national, structural, and local. The national component reflects trends affecting volunteering countrywide. The structural component captures how sector-specific characteristics influence volunteering, based on public awareness, funding, and needs. Finally, the local component examines how local factors, such as politics and socio-economic conditions, affect volunteering levels, highlighting how regions deviate from national trends. To present the results clearly, two maps are shown in Figure 7, offering a comprehensive overview of volunteer distribution dynamics. The first map (Figure 7a) shows the structural component, with positive growth mainly in Southern regions, while most Central and Northern regions show negative variation. The local component (Figure 7b) reveals a more heterogeneous distribution, with high values in Southern regions like Sicily and Puglia, as well as Central-Northern regions like Lazio, Piedmont, and Liguria.

Figure 7 - *Shift-share analysis results.*



Source: Our elaboration on Istat data (Nonprofit Institutions Census).

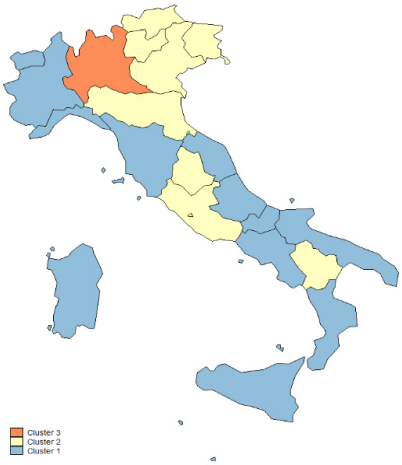
In order to analyse the behaviour of the two structural and local components together, we carry out a sign analysis. This analysis consists of classifying the regions into four clusters according to the joint direction of the two components, summarised as positive or negative, as can be seen in Table 1. It is clear that the sign + indicates a positive trend, while the sign - indicates a negative trend.

Table 1 - *Analysis of signs: description of clusters.*

Cluster	Structural component	Local component
1	+	+
2	+	-
3	-	-
4	-	+

Figure 8 shows the results of the sign analysis in a quantile map. We report three clusters as there are no regions that show a negative sign of the structural component and a positive sign of the local component (cluster 4). The majority of regions have a positive sign for both components and therefore fall into cluster 1. In this cluster we find all the Southern regions with the exception of Basilicata. Cluster 2 includes all the regions that make up the North-East macro-region according to the NUTS 1 classification. It is interesting to note the case of Lombardy, the only region to show a negative trend in both components.

Figure 8 - *Shift-share analysis: analysis of signs.*



Source: Our elaboration on Istat data (Nonprofit Institutions Census).

The results presented above show significant regional differences. The most interesting data undoubtedly concern the South of Italy (the Mezzogiorno), which shows positive results for both the local and structural components in almost all its regions. This may reflect a community response to the shortcomings of the welfare state, exacerbated by the pandemic. Local characteristics, such as certain traditions of solidarity and community networks, together with the expansion of sectors particularly affected by the crisis, may have contributed to the strengthening of

volunteering. In addition, local initiatives, infrastructure improvements, and new forms of volunteering were able to respond effectively to emerging needs. This strengthening of volunteering could be seen as a measure of resilience and adaptation to socio-economic difficulties, with a significant difference compared to other regions of the country.

The Mezzogiorno, with its socio-economic fragility, has experienced a crisis that has highlighted and accentuated its structural vulnerabilities. However, the community's response through volunteering represents a dimension of antifragility. In other words, while the multiple crises has exposed the weaknesses of this area of the country, the active response and strengthening of volunteering show how communities can turn difficulties into opportunities for development. It is interesting to note the results for Piedmont and Liguria, which show similar dynamics to the Southern regions. Both regions are experiencing a marked demographic decline, particularly in the inner areas. The decline in population, due to low birth rates and migration to other areas, reduces the number of young people and increases the proportion of elderly people. This phenomenon leads to greater territorial vulnerability, as resources and services are often concentrated in urban areas, leaving rural areas more isolated and less well served. As in the Southern regions, volunteering can be a response to structural deficiencies and the limited capacity of the welfare state to address new social risks. Finally, the results from Lombardy can be interpreted in terms of a kind of crowding-out effect. The progressive professionalisation observed in the non-profit sector could indeed lead to a gradual replacement of volunteers by paid workers with specific skills. This phenomenon reflects the role of Lombardy, which is characterised by a high level of employment in the non-profit sector, in attracting resources and investment, thereby promoting greater structuring and specialisation of the services offered.

4. Concluding remarks

The results of this exploratory study offer important insights into the state and evolution of organised volunteering in Italy. Our findings indicate that analysing volunteer work in the non-profit sector requires considering the framework of changes it has undergone and will continue to face (Ranci, 2006; Licursi et al., 2022). The pandemic led NPIs to suspend or reschedule their activities, making volunteer contributions more crucial than ever in addressing emerging vulnerabilities and difficulties (Ambrosini & Erminio, 2020). The Third Sector Reform has also impacted NPI organisational practices, the effects of which will become clearer with upcoming census surveys (Caltabiano et al., 2024). This analysis represents a first step towards understanding the factors driving change in organised volunteering, with territory playing a significant role (Biorcio & Vitale, 2016). The forthcoming

data will provide a clearer picture of the characteristics and peculiarities of organised volunteering, its activities, and its role in supporting vulnerable social categories, the community, the environment, and the care of common goods (Citroni, 2022; Moro, 2022).

References

AMBROSINI M., ERMINIO D. 2020. *Volontari inattesi. L'impegno sociale delle persone di origine immigrata*. Trento.

BEATTON T., TORGLER B. 2018. Volunteering and life or financial shocks: does income and wealth matter?. *Applied Economics*, Vol. 50, No. 19, pp. 2190-2209. https://doi.org/10.1080/00036846.2017.1392003

BIORCIO R., VITALE T. 2016. *Italia civile: associazionismo, partecipazione e politica*. Donzelli Editore.

CALTABIANO C., VITALE T., ZUCCA G. 2024. *La prospettiva civica. L'Italia vista da chi si mette assieme per cambiarla. Decimo rapporto Iref sull'associazionismo sociale*. Feltrinelli.

CITRONI S. 2022. *L'associarsi quotidiano*. Meltemi.

DELLA PORTA D. 2020. Building bridges: Social movements and civil society in times of crisis. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, Vol. 31, No. 5, pp. 938-948. https://doi.org/10.1007/s11266-020-00199-5

DUNN E. 1960. A Statistical and Analytical Technique for Regional Science. *Papers, Regional Science Association*, Vol. 6, pp. 97-112.

ESTEBAN-MARQUILLAS J. 1972. A reinterpretation of shift-share analysis. *Regional and Urban Economics,* Vol. 2, No. 3, pp. 249-255. https://doi.org/10.1016/0034-3331(72)90033-4

EURISPES 2022. *Volontariato in pandemia. Le risposte delle organizzazioni e degli enti del Terzo Settore*. https://eurispes.eu/news/volontariato-in-pandemia-le-risposte-delle-organizzazioni-e-degli-enti-del-terzo-settore/.

FIORILLO D., NAPPO N. 2015. Formal and informal volunteering and health in Mediterranean Europe. *Athens Journal of Mediterranean Studies*, Vol. 1, No. 4,pp. 297-309.

GUO J., LIU N., WU Y., ZHANG, C. 2021. Why do citizens participate on government social media accounts during crises? A civic voluntarism perspective. *Information & Management*, Vol. 58, No. 1. https://doi.org/10.1016/j.im.2020.103286

LICURSI S., SABINA ET AL. 2022. Dopo la pandemia, cosa? Adattamenti e innovazioni nel volontariato organizzato in Italia. *La Rivista delle Politiche Sociali*, Vol. 1,pp. 57-71.

MORO G. 2022. *La cittadinanza in Italia, una mappa*. Carocci.

RAGO S., VENTURI P. 2020. *Evoluzioni e sfide trasformative del volontariato*. Bologna: Aiccon.

RANCI C. 2006. *Il volontariato*. Il Mulino.

TERZO G. 2021. Social capital, social economy and economic resilience of Italian provinces. *Papers in Regional Science*, Vol. 100, No. 5, pp. 1113-1135. https://doi.org/10.1111/pirs.12618

TERZO G. 2022. Investigating the link between social cooperation sector and economic well-being of Italian provinces through the lens of social capital. *Annals of Public and Cooperative Economics*, Vol. 93, No. 4, pp. 1041-1062.

TERZO G., NOTARSTEFANO G., DI MAGGIO U. 2023. Non-profit sector and regional well-being in Italy. *Regional Studies*, Vol. 57, No. 3, pp. 462-477. https://doi.org/10.1080/00343404.2022.2070146

TERZO, G., NOTARSTEFANO, G., DI MAGGIO, U. 2024. Strengthening local economies through social capital: The role of cooperative firms in driving economic growth in Italy. *Annals of Public and Cooperative Economics*, Vol. *95, No.* 3, pp. 863-882. https://doi.org/10.1111/apce.12463

UNITED NATIONS. 2003. *Handbook on Non-profit Institutions in the System of National Accounts*, Studies in Methods, Series F., No. 91, New York.

WILSON J., MUSICK M. 1997. Who cares? Toward an integrated theory of volunteer work. *American Sociological Review*, pp. 694-713. https://doi.org/10.2307/2657355

Umberto DI MAGGIO, Lumsa University, u.dimaggio@lumsa.it
Stefania DELLA QUEVA, Istat, dellaqueva@istat.it
Manuela NICOSIA, Istat, mnicosia@istat.it
Giuseppe NOTARSTEFANO, Lumsa University, g.notarstefano@lumsa.it
Sabrina STOPPIELLO, Istat, stoppiel@istat.it
Giuseppe TERZO, Luiss Guido Carli and LUMSA University, gterzo@luiss.it

STEEL INDUSTRY IN ITALY: WHICH TRADE-OFF?[1]

Annamaria Fiore, Lucia Mongelli

Abstract. The steel industry is one of the most strategic economic sectors in any country and it is crucial in Italy's economy. However, it is among the most polluting industries, posing serious risks to workers and residents near industrial plants.
The paper aims to contribute to providing evidence on which to base more informed and conscious policy choices and private decisions to balance economic growth with environmental sustainability and public health connected to the steel industry, comparing economic, environmental, and health indicators for Italian provinces, depending on the presence of significant iron and steel activities.
Using official data from the Italian Institute of Statistics, we investigate economic benefits, environmental impacts, and health risks associated with the presence of steel plants in territories where they operate. The Wilcoxon (Mann-Whitney) method, a non-parametric test based on rank order, is used.
Our results suggest that while the immediate economic impact is evident, the environmental and health outcomes – found in the data but not statistically significant - may be directly correlated only in the longer term, or other factors might be influencing these results. Therefore, policy decisions should anticipate and incorporate the future effects of current actions using increasingly refined and contextualised analytical tools.

1. Introduction

The EU Court of Justice has recently established in a ruling (June 25, 2024) that steelworks in Taranto should be shut down if the plant poses a threat to human health and the environment: "In the event of serious and significant threats to the integrity of the environment and human health, the operation of the installation must be suspended", in its own words.

The ongoing debate concerning the steel industry's trade-offs—balancing workers, production, and profit against citizens, health, and the environment—is now more critical than ever. The steel sector is not only pivotal due to its direct influence on output and jobs but also for its foundational support to other industries

[1] The paper is the result of the common work of the authors. In particular, Annamaria Fiore has written Sections 1, 2 and 5; Lucia Mongelli has written Sections 3 and 4, the Discussion and conclusions are from both authors.

reliant on steel, particularly in Europe. According to Weinel et al. (2024), this sector is exceptionally strategic in Europe, with over 2,600 companies and 315,000 workers per Eurostat's 2021 data.

In the European Union, Italy ranks second among European steel producers in terms of value-added and turnover, just behind Germany and ahead of France. The latest official statistics indicate the relevance of the industry's size: more than 450 local units, nearly 40 thousand employees, 38 billion euros in turnover, more than 5 billion euros in added value (Istat data 2021), and 11,6 billion euros in exports (Coeweb data 2023).

However, the steel industry is among the most polluting and poses serious risks to workers and residents near industrial plants, as highlighted by the situation with the former ILVA plant in Puglia. For this reason, it is necessary to have updated and disaggregated statistics at the most appropriate territorial level to have a structured knowledge base for policy decisions.

In this paper, we focus on comparing Italian provinces in terms of economic, environmental, and health performance, depending on the presence of significant steel industry activities, based on the official statistical information currently available.

2. Previous literature

The literature about the steel industry in Italy is quite copious: many articles focus on the diachronic evolution of the industry in the country, often from a comparative perspective with other European states (Sáez-García, 2016; Ranieri, 2019). Until a few years ago, the focus was mainly on privatisation (Brambilla and Lavista, 2020; Mollona and Pareschi, 2020). However, in recent years, the interest has predominantly shifted to environmental issues, with particular attention to the case of Taranto.

In Lai et al. (2019a), the authors investigate the role of states in governing the sustainability trajectories and decisions of companies and their local communities. According to the authors, the Italian government made its decisions on ILVA in the name of relevant risks related to economic dimensions (unemployment, economic development, and territorial competitiveness) and silenced the environmental and health risks.

In Bellantuono et al. (2021), the authors used two different aggregate indexes (i.e., the Adjusted Mazziotta-Pareto Index and the Adjusted Differences Mean Index) to analyse data: BESdT does detect many problems affecting the examined area of Taranto, but it seems not able to frame the crisis adequately. The critical situation does not always reflect lower territorial performance, neither at the level of

single indicators nor at the level of entire domains. Such discrepancy appears to be particularly evident within the economic domain.

More recently, the World Health Organization has evaluated the health impact of the Taranto steel plan (2023). It highlights the plant's negative environmental and health consequences, including air pollution and related diseases. The WHO has updated health impact estimates, suggesting 27 annual deaths in the worst-case scenario, reduced to 5 in the best-case scenario.

The report emphasizes the need for comprehensive assessments of the plant's policies on quality of life and the environment to align with sustainable development goals.

3. Context

The steel industry is one of the main productive sectors on which the national economy is based. Steel is a material used by many other sectors of the economy as one of the most versatile and recyclable materials.

The perimeter of the steel sector can be outlined through official statistics. According to Eurostat data from 2021, there are over 2,600 steel companies in Europe with 315,000 employees.

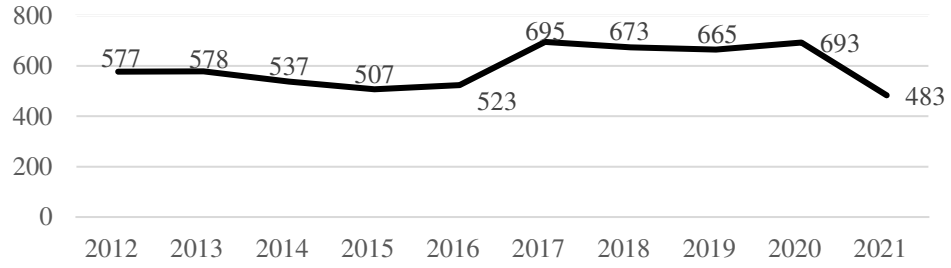
Italy maintains a key role at the European level, ranking second among European steel producers immediately after Germany and before France (Eurostat data, 2021).

In Italy, this industry also represents one of the primary sectors of national industrial production. In 2021, the international steel sector was affected by the Covid-19 pandemic, both directly in terms of production and indirectly, because the steel-using sectors experienced significant declines in activity compared to the previous year, such as the automotive sector, the production of the metal products sector, etc.

Moreover, the growth prospects for production are threatened by the rise in energy costs, the consequent increase in production costs, and the shortage of raw materials, and the consequent increase in the prices of ore, coal and iron scrap.

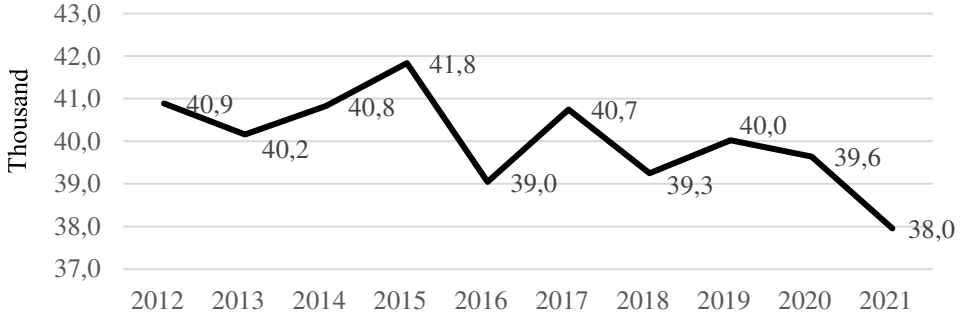
Based on Istat data as of December 31, 2021, there were 483 local units of enterprises in the steel sector in Italy (Istat ATECO 2007, NACE Rev.2, identified by code C24.1) with direct employment of about 38,000 people, with a variation over the previous year of -30.3 per cent in the number of units and -4.3 per cent in the number of employees, respectively (Figure 1 and 2).

Figure 1 - *Number of Local Units of active enterprises ATECO C24.1. Italy. Years 2012-2021. Absolute values.*



Source: Istat, ASIA UL

Figure 2 - *Number of Local Units employees of ATECO C24.1. Italy. Years 2012-2021. Average annual values.*

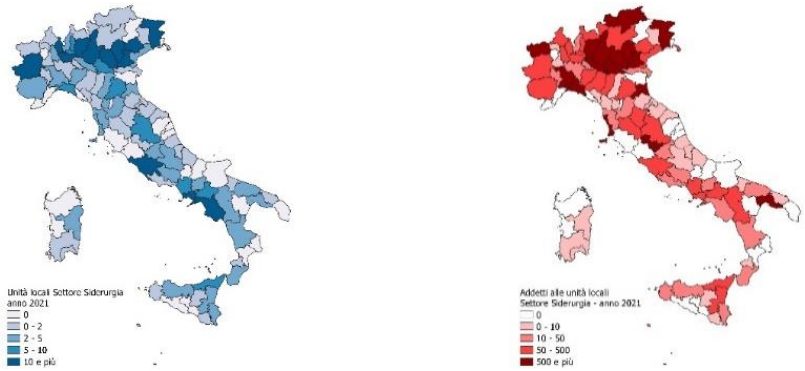


Source: Istat, ASIA UL

According to Istat, about 34% of the workforce of companies in the steel sector is concentrated in the North-West of the country, about 28% in the South and slightly more than 25% in the North-East (Figure 3); this is slightly more than 1% of the Italian manufacturing workforce.

The figures also show the industrial composition of the steel industry, which is heavily concentrated on medium-sized large firms, with 80% of workers employed in firms with over 50 employees. The bulk of the industry is made up of medium-sized steel companies.

Figure 3 - *Provincial distribution of Local Units (left panel) and of Local Units employees (right panel) of the steel industry in Italy (ATECO C24.1). Year 2021. Absolute values.*



Source: Istat, DATABASE di Indicatori Strutturali. Settore 24, https\\www.istat.it\settori-produttivi

4. Data and Method

Our analysis aimed to identify differences between territories based on a specific condition. Specifically, the differences between the territories had to concern socioeconomic, environmental aspects, and residents' health; the distinguishing characteristic between the two groups of territories was the presence or absence of large steel plants.

Considering the available data for all these dimensions, the lower territorial level at which this analysis could be conducted was the provincial level. At this point, the condition to divide the dataset into two subsamples was the number of employees in the sector: at least 250 for all the years in the considered time interval, namely from 2012 to 2021.

Out of the 109 provinces examined, they were classified into two categories: 19 provinces were identified as having a high concentration of steel industry employment (categorized as *steel-concentration*), while the rest were categorized as *no steel-concentration*.

We present the complete dataset description below (Table 1).

Table 1 – *Dataset.*

Variable	Source	Year
Employees	Istat – Archive ASIA	2012-2021
Dichotomous	1 = at least 250 employees (19 provinces)	Cumulative 2012-2021
DIMENSION: ECONOMIC		
Employment, unemployment, inactivity rate	Istat - Permanent Census of Population and Housing	2021
Low income	Our elaboration on MEF data	2021
Entrepreneurship rate	Istat - Statistical Registry of Local Units (ASIA)	2021
Specialization in high-technology sectors	Our elaboration on Istat data (ASIA)	2021
Density of local units	Istat - Statistical Registry of Local Units (ASIA)	2021
DIMENSION: ENVIRONMENTAL		
Synthetic indicator of atmospheric pollution	Istat - Environmental data in cities *(data referred to capital cities)*	2019-2020 2020-2021
DIMENSION: HEALTH		
Mortality due to malignant tumor	Istat - Survey of deaths and causes of death	2021
Mortality due to liver tumor*		2021
Mortality due to lung tumor*		2021
Mortality due to prostate tumor*		2021
Mortality due to bladder tumor*		2021

** Selected based on the study by Cazzolla Gatti and Velichevskaya (2022)*

As a first step of the analysis, we show the descriptive statistics of the variables of our dataset (Table 2).

The last decade has seen a decline in employment and local units within Italy's steel industry. The average number of employees has decreased from 397 in 2012 to 368.5 in 2021, and the number of local units has dropped from 5.6 to 4.7 per province. It is worth mentioning that the province with the highest number of employees is Taranto.

To test if there are significant differences between provinces with a high concentration of steel processing and those without, we use Wilcoxon's sum-of-ranks test (Mann-Whitney), a nonparametric test based on character rankings (Wilcoxon 1945; Mann and Whitney 1947). It tests the hypothesis that two independent samples or unmatched data are from populations with the same distribution. It is appropriate when dealing with small sample sizes or when the assumptions of the parametric tests are not met, as in our case. Results are presented in the next section.

Table 2 – *Descriptive statistics*.

Variable	Obs	Mean	Std. Dev.	Min	Max
Steel employees 2012	103	396.988	1305.267	0.000	11486.23
……					
Steel employees 2021	103	368.509	1126.859	0.000	9758.840
Avg. steel empl. 12-21	107	374.200	1204.800	0.000	10,794.600
Steel local units 2012	103	5.592	8.874	0.000	60.000
……					
Steel local units 2021	103	4.689	7.602	0.000	54.000
Employment rate 2019	107	45.262	5.949	33.418	58.773
Employment rate 2021	107	45.347	5.717	33.884	58.589
Unemploym. rate 2019	107	13.336	5.710	3.999	27.669
Unemploym. rate 2021	107	9.423	3.264	2.881	17.118
Inactivity rate 2019	107	47.990	3.804	38.786	55.348
Inactivity rate 2021	107	50.085	4.720	39.673	59.965
Low income 2021	107	28.513	7.100	18.971	44.379
Low work intens. 2019	107	48.875	5.258	40.057	60.619
Entrepreunership 2021	107	73.564	11.295	50.767	107.082
Density Local Units 2021	107	22.046	32.956	2.581	223.152
Hi-tech 2021	107	3.275	1.635	1.210	9.359
Air pollution 2019-20	101	23.205	20.201	0.000	66.667
Air pollution 2020 21	101	22.102	20.710	0.000	66.667
Tumor 2021	107	28.638	3.546	21.69	38.440
Liver 2021	107	1.420	.317	.740	2.340
Lung 2021	107	5.300	.931	3.280	7.100
Prostate 2021	107	1.398	.276	.840	2.110
Bladder 2021	107	1.017	.253	.500	1.660

5. Results

We tested whether there are significant differences in the data on employment, economic well-being, settled economy, air pollution, and health status of the population between the two categories of provinces: 19 provinces categorized as *steel-concentration*, while the rest as *no steel-concentration*, depending on the number of employees constantly equal to or greater than 250 throughout the examined decade.

Table 3 highlights the main results of this analysis.

In summary, we underline that:

- The results for economic indicators related to labour (employment, unemployment, inactivity) and income distribution are statistically significant ($p < 0.05$), suggesting that there is a significant difference between the two categories of provinces. In all cases, the indicators show better performance in *steel-concentration* provinces.

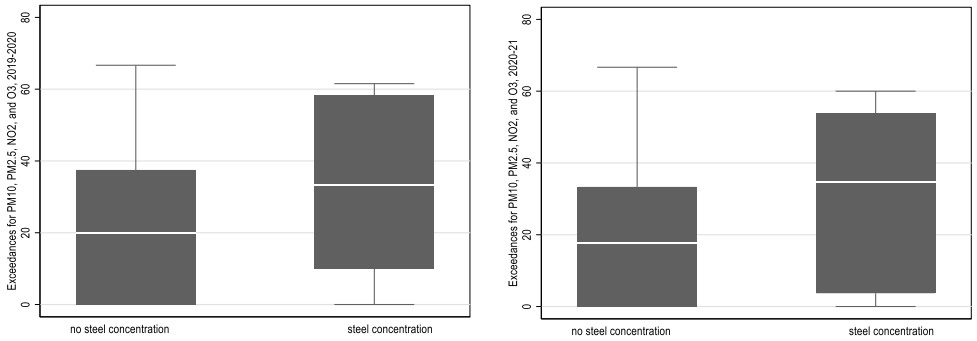
- Although the performance for settled economy (entrepreneurship rate and LU density) and innovation (specialisation in high-tech sectors) is on average higher for *steel-concentration* provinces, the differences are not statistically significant ($p > 0.05$).
- On average, environmental and health indicators present worst performance for *steel-concentration* provinces (except for bladder cancer). However, the Wilcoxon test highlights significant differences only in the case of mortality due to liver and bladder cancer.

Table 3 – *Results of the Wilcoxon rank sum test*.

Indicator	Test value	*p-value*
ECONOMIC INDICATORS		
Employment rate, 2021	-2.902	0.0037
Unemployment rate, 2021	3.318	0.0009
Inactivity rate, 2021	2.739	0.0062
Irpef taxpayers with total income less than 10k (incidence on total taxpayers), 2021	2.739	0.0062
Entrepreneurship rate, 2021	-1.663	0.0963
Specialization in high-tech sectors, 2021	-1.704	0.0884
Density of local units, 2021	-1.492	0.1358
ENVIRONMENTAL INDICATORS		
Summary air pollution indicator (exceedances of threshold values for PM10, PM2.5, NO2, and O3 concentrations), years 2019-2020	-1.697	0.0897
Summary indicator of air pollution (exceedances of threshold values for PM10, PM2.5, NO2, and O3 concentrations), years 2020-2021	-1.844	0.0651
HEALTH INDICATORS		
Mortality rate due to malignant cancer, 2021	-0.554	0.5794
Mortality rate due to liver cancer, 2021	-2.067	0.0388
Death rate due to lung cancer, 2021	-0.041	0.9675
Mortality rate due to prostate cancer, 2021	0.188	0.8512
Mortality rate due to bladder cancer, 2021	2.026	0.0427

Considering that environmental indicators refer to capital cities and not to the entire province, we consider it appropriate to conduct a further investigation by presenting an exploratory graphical analysis. Observing the box plots, we could conclude that in the *steel-concentration* provinces, there is a greater exceedance of threshold values for PM10, PM2.5, NO2 and O3 concentrations (Figure 4). However, this conclusion is not supported by the results of parametric tests (Table 3), possibly because official statistics provide data with a level of territorial disaggregation different from that adopted in our work.

Figure 4. *Summary indicator of environmental pollution. Years 2019-2020 and 2020-2021. Per 100 valid measurements, two-year moving averages. Values for capital cities.*



Source: Istat - Environmental data in cities

On the other hand, in the case of health indicators a comparison was made between the Italian average of health indicators with the top four steel-concentration provinces: Taranto, Brescia, Udine, Terni (Table 4).

Table 4 - *Cancer mortality rates* by selected steel-concentration provinces. Year 2021.*

Mortality rate	Italia	Taranto	Brescia	Udine	Terni
Malignant tumor	27.86	26.40	**28.10**	**32.49**	**32.35**
Liver cancer	1.36	**1.37**	**2.03**	**1.81**	**1.87**
Lung cancer	5.36	4.69	**5.49**	**5.77**	**6.57**
Prostate cancer	1.34	**1.39**	**1.45**	**1.48**	**1.83**
Bladder cancer	1.00	**1.18**	0.83	0.65	**1.19**

** Selected based on the study by Cazzolla Gatti and Velichevskaya (2022)*
Source: Istat. Survey of deaths and causes of death. https:\\dati.istat.it

Among those analysed, Terni is the only province where cancer mortality rates are higher than the national averages for all cases, while Brescia and Udine have higher mortality rates except for bladder cancer.

Taranto, on the other hand, has total cancer mortality rates lower than the Italian average and it has slightly higher mortality rates than the Italian average only for liver and prostate cancer. A larger difference is observed, however, in the case of the bladder cancer. These indicators, however, do not perfectly align with what can be inferred from the news reports, ongoing judicial investigations for the province of Taranto, or previous research (Cazzolla Gatti and Velichevskaya, 2022).

Anyway, our results are in track of what has already been expressed by Bellantuono et al. (2021), who noted, albeit with a different approach, that statistical indicators (in their case, those of the BESdT) are not sufficient to detect the overt

criticality of an area such as Taranto, and the same reasons can be cited there (data not available or not adequate or lack of historical series and/or of granularity).

6. Discussion and conclusions

In an era of growing global interconnectedness, it is crucial to simultaneously address emerging challenges related to health, environmental sustainability, and social inequalities. The decisions made by institutions and the behaviours adopted by businesses significantly impact the health of communities and the surrounding ecosystem.

Understanding phenomena in their socioeconomic-environmental dimension is essential for policymakers and stakeholders in the steel sector: public and private decisions should balance economic growth with environmental sustainability and public health, and require increasingly detailed data (Lai et al., 2019b).

In this context, the presence of steel plants, while bringing wealth to the areas where they are established, primarily by ensuring jobs and income, also raises increasingly serious questions about the destine of the populations living there. These populations continuously and inexorably suffer devastating effects on health and the environment, as exemplified by the case of the steel plant in Taranto.

This is the circumstance to delve deeper into the complex trade-off between economic growth and the deterioration of health and environmental conditions, using the available official statistical data.

Our results confirm increased employment and economic welfare for the territories hosting steel plants of some importance, at the same time risks related to the preservation of the environment and human health are non-negligible. If the economic impact is clear and immediate, environmental and health outcomes may be directly related in the long run.

To better understand and balance the costs and benefits, it is essential to have additional indicators with the same spatial granularity for environmental and sustainable development metrics or to address the lack of adequate historical data series (Bellantuono et al., 2021). To cite a few examples, we would need more detailed data on the contribution of steel plants to local economies, including job creation, investment, and technological innovation (economic aspect). Very recently, even the World Health Organization stated that important factors, such as contamination of soil, water, waste, food, the urban environment and green spaces are influenced by industrial policies in the steel industry, although these cannot be reliably quantified at present, but should be thoroughly evaluated as part of the sustainable development agenda. It is also important to consider confounding factors that might influence the results, such as socioeconomic status, access to healthcare, and other industrial activities in the areas investigated.

This research has some significant limitations. Firstly, the lack of accurate historical data prevents a full assessment of the long-term effects of steel plants. In addition, the lack of data with the spatial granularity needed to analyse local impacts limits the accuracy of conclusions. Confusing factors such as socioeconomic status and access to health care may affect outcomes, making a more advanced methodological approach necessary to isolate the specific effects of facilities. Finally, the evolution of industrial policies and technologies could alter impacts over time, requiring continuous updating of data for a more accurate and complete assessment.

In conclusion, future research directions foresee the inclusion of additional factors in the analysis to provide a more complete overview of risks, such as morbidity rates or specific production process characteristics of steel plants (e.g., continuous cycle, full cycle, and separated cycles as in Olmez et al., 2016). In the same way, it is imperative to develop methodologies capable of forecasting future scenarios; decision-makers need to anticipate and integrate the prospective impacts of current actions by employing increasingly sophisticated and context-specific analytical tools.

Acknowledgements

The authors would like to thank Rossana Mancarella for her helpful assistance with the dataset preparation and the participants of the LX Scientific Meeting of the Italian Society of Economics Demography and Statistics for their insightful comments.

References

BELLANTUONO N., LAGRASTA F.P., PONTRANDOLFO P., SCOZZI B. 2021. Well-Being and Sustainability in Crisis Areas: The Case of Taranto. *Sustainability*, Vol. 13, No. 3, pp. 1-23.

BRAMBILLA C., LAVISTA F. 2020. Privatizations and efficiency. Evidences from the Italian iron and steel industry, 1979–2016, *Industrial and Corporate Change*, Vol. 29, No. 3, pp. 757–778.

CAZZOLLA GATTI R., VELICHEVSKAYA A. 2022, Taranto's Long Shadow? Cancer Mortality Is Higher for People Living Closer to One of the Most Polluted City of Italy, *Sustainability*. Vol. 14, No. 5, Article No. 266.

LAI A., PANFILO S., STACCHEZZINI R. 2019a. The governmentality of corporate (un)sustainability: the case of the ILVA steel plant in Taranto (Italy), *The Journal of Management and Governance*, Vol. 23, No. 1, pp. 67-109.

LAI A., PANFILO S., STACCHEZZINI R. 2019b. Accounting, Soci(et)al Risks, and Public Reason: Governmental Risk Discourses About the ILVA Steel Plant in

Taranto (Italy). In *Multiple perspectives in risk and risk management: ERRN 8th European risk conference 2018*, Katowice, Poland: Springer, pp.55-86. Springer proceedings in business and economics ringer.

MANN H. B., WHITNEY D. R. 1947. On a test whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, Vol. 18, pp. 50-60.

MOLLONA E., PARESCHI L. 2020. A Gramscian perspective on field dynamics. The case of the privatization of Italian steel industry, *Scandinavian Journal of Management*, Vol. 36, No. 4, pp. 101-128.

OLMEZ G.B, DILEK F.B., KARANFIL T., YETIS, U. 2016. The environmental impacts of iron and steel industry: a life cycle assessment study. Journal of Cleaner Production, Vol. 130, pp. 195-201.

RANIERI R. 2019. Steel and the State in Italy and the UK. The public sector of the steel industry in comparative perspective (1945-1996). In FELDENKIRCHEN W., GOURVISH T. Eds. *The European Yearbook of Business History*. London: Routledge, pp. 357-373.

SÁEZ-GARCÍA, M. A. 2016. Business and State in the development of the steel industry in Spain and Italy (c.1880–1929). *Business History*, Vol. 59, No. 2, pp. 159–178.

WEINEL M., ANTONAZZO L., STROUD D., BEHREND C., COLLA V., GOTI A., SCHRÖDER A.J. 2024. Introduction: The Historic Importance and Continued Relevance of Steel-Making in Europe. In STROUD, D., ET AL. Eds. *Industry 4.0 and the Road to Sustainable Steelmaking in Europe. Topics in Mining, Metallurgy and Materials Engineering*. Cham: Springer, pp. 1-14.

WILCOXON, F. 1945. Individual comparisons by ranking methods. *Biometrics*, Vol. 1, pp. 80-83.

WORLD HEALTH ORGANIZATION. 2023. *Health impact assessment of steel plant activities in Taranto, Italy*. Regional Office for Europe: World Health Organization.

Annamaria FIORE, ARTI – Agenzia Regionale per la Tecnologia e l'Innovazione, fiore@arti.puglia.it

Lucia MONGELLI, Istat, mongelli@istat.it

RETHINKING ITALIAN INNER AREAS: LESSONS FROM THE POTENTIAL DEMOGRAPHY[1]

Francesca Bitonti, Angelo Mazza

Abstract. The National Strategy for Inner Areas (NSIA) categorizes Italian municipalities based on their distance from "poles", i.e. centres providing essential services. Inner areas (IAs) are municipalities located within a certain distance from the poles. Although the NSIA aims to counteract depopulation and ageing processes affecting IAs, traditional demographic indicators, as the ones officially released by Istat annually, have not been considered to identify them.
For this reason, we aim to provide a demographically nuanced representation of IAs. In particular, we exploit the concepts of potential demography to delineate a demographic profile of the Italian IAs, both in actual and predictive terms. The idea of Potential Years of Life is used to assess the future economic potential of the IAs' population.
The preliminary application shows a complex situation that the NSIA classification failed to capture. A discrepancy between the IAs experiencing the strongest demographic distress and the territories receiving national funding has emerged.

1. Introduction: behind the demographically distressed territories

A large part of the Italian territory is characterized by a spatial organization based on "minor centres", often small in size, which in many cases are not able to guarantee residents accessibility to essential services. These areas have been defined as "inner areas" (IA) by the National Strategy for Inner Areas (NSIA) in 2014. NSIA is a government policy for development and territorial cohesion to counteract marginalisation and demographic decline within IAs throughout the country, which was defined by the "*Accordo di partenariato 2014-2020*" approved by decision of the European Commission on October 29, 2014 (Barca *et al.*, 2014). Since the 1950s, a significant part of the Italian rural and inland territories has undergone a process of marginalization which, first of all, has exhibited an intense phenomenon of de-anthropization, consisting in a) reduction in the population and demographic ageing; b) reduction in employment and in the degree of exploitation of territorial capital (Barca *et al.*, 2014). Secondly, this process has brought about a progressive

[1] Authors' contribution: conceptualization, final draft and editing: FB, AM; supervision, methodology: AM; data wrangling, statistical analysis, original draft, visualization: FB.

quantitative and qualitative reduction of the local offer of public, private, and collective services – that is, the services which "define the quality of citizenship in contemporary European society" (Vendemmia *et al.*, 2021).

NSIA categorizes all Italian municipalities on the basis of their distance from the centres providing services, defined as "poles", the identification of which constitutes the first step of the classification process. Poles are defined as municipalities able to offer the presence of simultaneously:

1. at least one hospital with an emergency department;
2. good offer of secondary schools (i.e., at least two different types of high schools);
3. one railway station (Italian silver level).

The other municipalities are then classified according to the travel time residents must spend to reach the nearest service provider centre:

1. Belt areas, travel time under 20 minutes;
2. Intermediate areas, between 20 and 40 minutes;
3. Remote areas, between 40 and 75 minutes;
4. Ultra-remote areas, with a travel time above 75 minutes.

Inner areas are defined as all the municipalities whose travel time is above 20 minutes, i.e. intermediate, remote, and ultra-remote areas. The distances considered are the road distances measured in terms of time in ideal travel conditions. According to this classification, inner areas cover 60% of the national territory, 52% of all municipalities, and 22% of the Italian population (De Matteis, 2013).

A new classification was published at the beginning of 2022 (Nota tecnica Nuvap, 2022). The methodology used has not changed; however, among the updates, notable is the introduction of more advanced and precise distance calculation techniques (Istat, 2022).

NSIA has, as the ultimate objective, the reversal of the demographic trend, both in terms of number of residents and in terms of composition by age and birth rate (Barca *et al.*, 2014; Barca & Carrosio, 2020). Therefore, the demographic situation has to be central to formulating an economic development plan for IAs. Despite this, no demographic indicator is considered to identify IAs affected by chronic population drain. The model proposed by NSIA has received appreciation in the scientific debate, both for the proposed classificatory work and the definition of "inner area". At the same time, despite the extensive premises accompanying the NSIA processing, it appears to be strongly focused on the provision of the three services indicated. The same classification model proposed, based solely on distances from the municipalities considered "poles", determines the categorization of the whole national territory, neglecting the differences in terms of demographic, economic conditions, and infrastructural accessibility. Moreover, there remain some inconsistencies such as that, while recognizing that inland areas are affected by the

exodus of young people, NSIA still adopts indicators such as health services, which are of little significance in terms of improving the supply of services that create opportunities to prevent this.

Here, we intended to provide a more articulated reading of the phenomenon in question, arriving at a classification of the territory capable of going beyond the availability of the three services of general interest (schools, hospitals, stations) considered by the NSIA as essential to ensure territorial cohesion and the development of disadvantaged areas. This work aims to provide a demographic profile of the Sicilian municipalities to stress the importance of the population structure and dynamics in identifying the areas suffering from young and active forces drain.

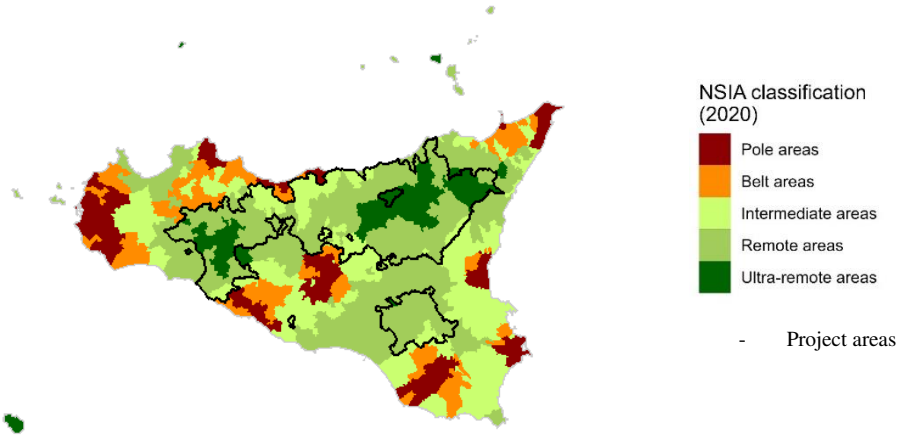
The remainder of the work is structured as follows: section 2 presents the geographical context of the application and elaborates on the demographic profile of the Sicilian municipalities; section 3 reports the discussions and conclusions of the work.

2. The demographic profile of the Sicilian municipalities

The southern Italian island of Sicily will be a case application of a new classification method based on the concept of potential demography that will be introduced in the next section. According to the 2020 NSIA classification (Dipartimento per le politiche di coesione - Mappa Aree Interne 2020 (governo.it)), the Sicilian IAs include 310 of the 390 municipalities of the region (areas coloured in shades of green in Figure 1).

Among them, 34 municipalities have been classified as ultra-remote areas and are mainly located in the northeastern and western inland areas (without considering the minor isles). Between 2014 and 2022, NSIA selected eight project areas (PAs) among the Sicilian IAs: Calatino, Madonie, Nebrodi, Terre Sicanie, Valle del Simeto – Etna, Troina, Bronte and Corleone (Regione Sicilia, 2022). The PAs are groups of municipalities intended as pilot areas where targeted actions to counter territorial marginalization are to be implemented. For each PA, a framework program agreement (ApQ) is established, outlining all the interventions to be carried out, the allocated financial resources, the scheduling of activities, and the expected results associated with each intervention (Tantillo and Lucatelli, 2018).

Figure 1 – *NSIA classification of Sicilian municipalities updated in 2020. Source: authors' elaboration on NSIA data.*



IAs account for 75.64% of the whole area of the region while hosting almost half (47.84%) of the total residing population in 2020 (Table 1). Similar figures highlight the importance of IAs both geographically and demographically.

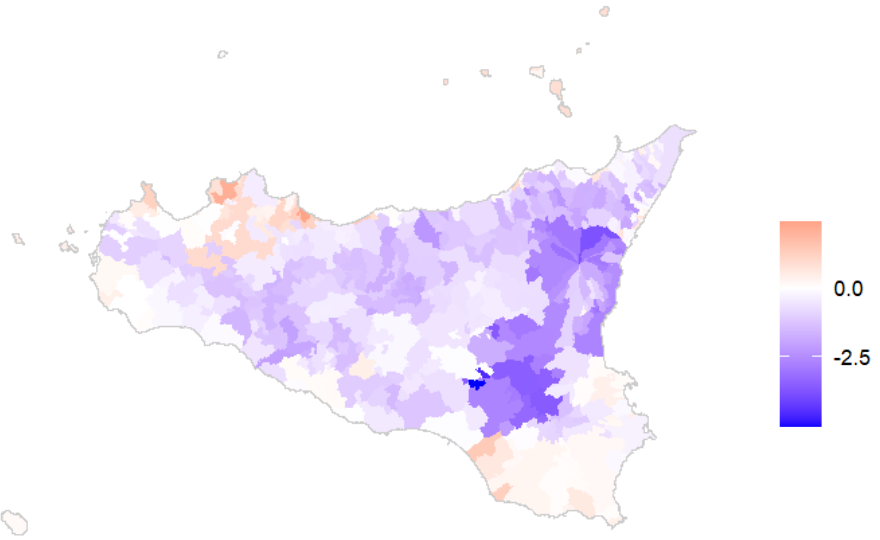
Table 1 – *Territorial area and population counts in Sicilian NSIA areas. Percentage values in brackets. Source: authors' elaboration on NSIA data.*

NSIA classification	Geographic area (km2 – 2019)	Total population (2020)
Pole areas	2882.30 (11.16)	1758272 (36.38)
Belt areas	3409.22 (13.20)	763426 (15.79)
Intermediate areas	6584.51 (25.50)	1151185 (23.82)
Remote areas	10749.76 (41.61)	1059083 (21.91)
Ultra-remote areas	2206.71 (8.53)	101739 (2.11)
Sicily	25832.54 (100.00)	4833705 (100.00)

In Sicily during the last decades, economic globalization and the expansion of urban areas have aggravated the contrast between the coastal metropolitan cities which offer public and private services (Palermo, Catania, and Messina) and the depopulation of the inland areas, historically characterized by the presence of agricultural activities (Scrofani and Novembre, 2015), as illustrated in Figure 2.

These areas are often inaccessible, sometimes unknown to foreigners, and suffer from economic and cultural marginalities. The census data from the last three decades highlighted that the local population has shown a progressive tendency to concentrate along the Northern and Eastern coasts, especially in the vicinity of the large metropolitan centres (Bitonti *et al.*, 2023). This trend has accentuated the differences between urbanized and inland areas, which are even more isolated, scarcely populated and almost completely lacking in basic services.

Figure 2 – *Population mean growth rate during the period 2002-2022 in Sicily. Source: authors' elaboration on Istat data.*



According to our elaborations based on Istat data, just 20.5% of the Sicilian municipalities registered a positive population mean growth rate in 2002-2022 (i.e. 80 municipalities out of 390). Moreover, areas recording an average growth in population between 2002-2022 have, on average, a higher share of young and adult people in 2022 compared to the areas registering population loss (Table 2). This means that the municipalities affected by depopulation have been losing the active part of the population. In other terms, less and less working-age individuals are left to support the increasing share of the elderly. As noted by Barca *et al.*, 2014, heterogeneous dynamics emerged, with areas losing and gaining population over time not always coinciding with IAs and poles. In our opinion, such "demographic distress" should be explicitly considered in the definition of IAs.

Table 2 – *Population structure of Sicilian municipalities in 2022 based on the population mean growth rate for 2002-2022. Source: authors' elaborations on Istat data..*

Population mean growth rate (2002-2022)	Young people [0-19]	Adult people [20-65)	Older people [65+)
Positive	18.9%	59.8%	21.2%
Negative	16.6%	57.7%	25.6%

Another relevant dimension to consider is, in our thoughts, the temporal dynamics of the demographic change involving the different municipalities. Areas experiencing systematic demographic drain over time should be separated from areas affected by a contingent population loss. Figure 3 shows the mean growth rates between 2002 and 2022 of different dependency indices (subfigs.3.a-c-e) and their absolute values recorded in 2022 (subfigs.3.b-d-f).
Values reported in the right-hand side maps refer to 2022 and highlight a higher demographic burden for people in the working-age class in IAs and especially in PAs (contoured by black solid lines). The lower proportions of young people (Y) and the higher proportions of older people (E) over the working-age individuals (A) emerging in IAs (and PAs in particular) may seem proof of the efficacy of the current NSIA classification. Yet, looking at the left-hand side maps, the situation becomes more complex, and the temporal dynamics of dependency ratios configure a multifaceted situation where the increasing burden of elderly over working people and the youngsters drain appear heterogeneous. Indeed, the areas characterized by an average increase in the demographic burden do not always correspond to the PAs selected among the IAs identified according to the current methodology.

Table 3 provides a comprehensive comparison of the mean rates and growth rates (2002–2022) for selected dependency ratios across three categories of municipalities: those not defined as IAs, IAs not included among PAs and PAs. In of mean dependency ratios, the PAs exhibit the highest values for all the indices considering the elderly proportions over the other parts of the population, indicating a greater demographic burden. Conversely, the Y/P and Y/A indices, which measure the proportion of youth relative to total and active populations, are marginally lower in PAs compared to other groups, signaling a reduced share of younger residents.

When examining mean growth rates of dependency ratios, a more nuanced picture emerges. While the E/P, E/A, and E/Y indices have increased across all groups, the rate of growth is significantly slower in PAs compared to the other municipalities. This slower growth in dependency ratios for PAs suggests a relative stabilization of the elderly burden, potentially reflecting the effects of NSIA interventions.

Figure 3 – *Dependency ratios (as mean growth rates between 2002-2022 in subfigs.4.a-c-e and as annual values in 2022 in subfigs.4.b-d-f) in Sicily. Note: Y = population aged [0,19]; A = population aged [20,65); E = population aged [65+); Dependency ratio: D_A = E/A + Y/A. PAs are contoured by solid black or white lines. Source: authors' elaboration on Istat and NSIA data.*

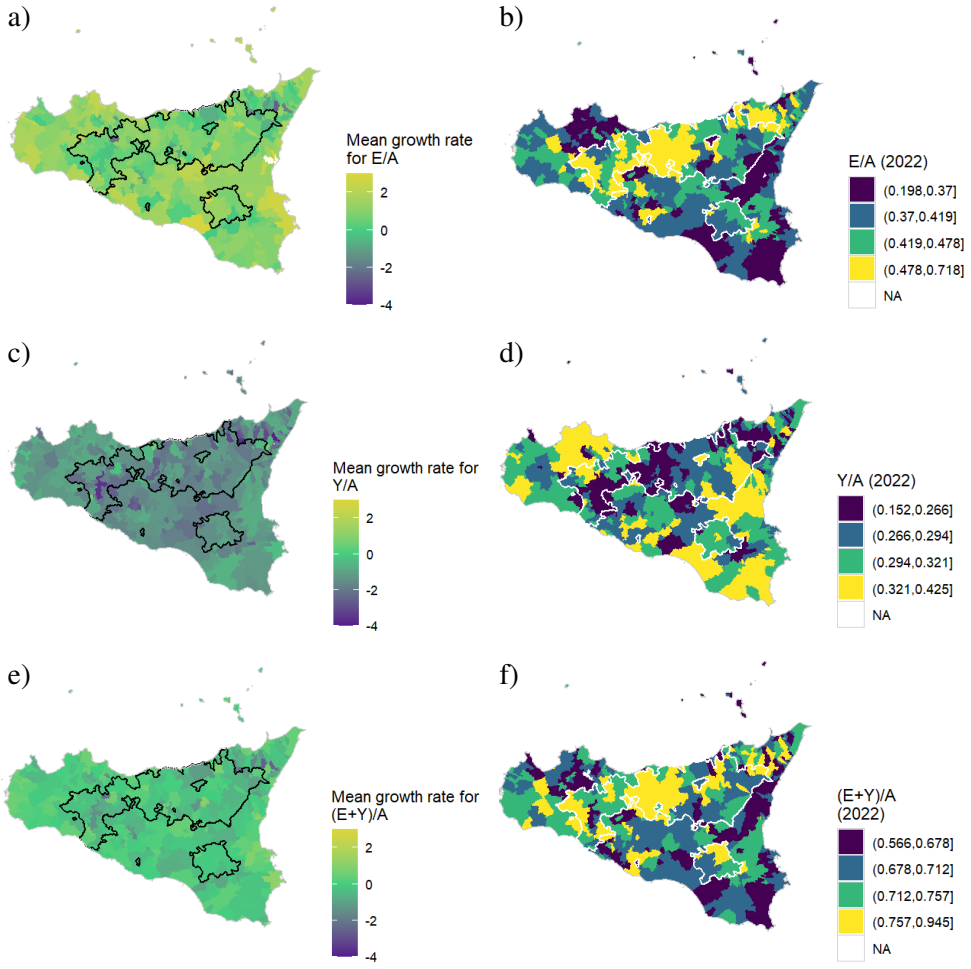


Table 3 – *Comparison between the mean rates and the mean growth rates (2002-2022) of selected dependency ratios between municipalities not defined as IAs, IAs not included among the PAs and PAs. Note: P = total population; Y = population aged [0,19]; A = population aged [20,65); E = population aged [65+). Source: authors' elaboration on Istat and NSIA data.*

	Index	Not IAs	Inner areas (no PAs)	Project areas
Mean rates (2002-2022)	E/P	0.200	0.226	0.250
	E/A	0.334	0.394	0.443
	E/Y	1.032	1.255	1.438
	Y/P	0.199	0.192	0.183
	Y/A	0.331	0.330	0.322
Mean growth rates (2002-2022)	E/P	1.731	1.015	0.841
	E/A	1.908	0.913	0.752
	E/Y	2.895	2.293	2.163
	Y/P	-1.202	-1.365	-1.408
	Y/A	-1.020	-1.468	-1.495

Conversely, the municipalities not classified as IAs exhibit the lowest average dependency ratios for elderly but show the highest mean growth rates for these ratios. This indicates that, although “Not IAs” have a lower initial burden from the elderly population, the pace of aging in these municipalities is accelerating more rapidly than in other regions. On the other end, the mean growth rates for Y/P and Y/A show consistent negative growth across all groups, with the decline being slightly more pronounced in PAs, reflecting ongoing challenges in maintaining or increasing the proportion of younger residents even in areas receiving funding. The slightly higher retention of youth in Not IAs further highlights the advantage of more accessible, urbanized regions in terms of demographic vitality and potential for future socio-economic growth. Table 3 reveals a clear overarching trend affecting all Sicilian municipalities, irrespective of their classification. Across the three categories, a common pattern emerges: the aging population is a prominent feature, with a significant demographic burden placed on the working-age population, particularly in more peripheral areas.

Overall, the depopulation processes occurring in Sicilian IAs have been causing two demographic trends: the first is the aggravation of the demographic burden of the elderly over working age individuals; the second is the shrinking of the young age classes, those which can contribute the most to the future socioeconomic vitality of IAs. The ongoing demographic dynamics induce future social, economic and cultural changes, which are worthy of attention. A crucial aspect that should be

addressed pertains to examining the labour force potential, which each area will need to confront the emerging challenges in the forthcoming decades (Blangiardo, 2013).

According to the potential demography paradigm, a population's future can be considered an economic asset, and a population possessing more "future years" can be considered wealthier (Blangiardo and Rimoldi, 2012). The assessment of a population's future is determined by calculating its "potential years of life" (PYL), a measure first introduced by Hersch (1944). PYL is the sum of the life expectancies of all its members. During a year, PYL increases through births and net migrations, while it decreases by the simple "consumption" of the remaining years of life (due to time flow) and by deaths (Blangiardo, 2012). Hersch's fundamental idea was straightforward. For an individual of a specific age x, their PYL is represented by the life expectancy, e_x (Panush and Peritz, 1996). If the age distribution of a population is given by P_x, then:

$$PYL = \frac{1}{2}\sum_{x=0}^{100+} P_x(e_x + e_{x+1}) \quad (1)$$

The PYL can be broken down in several meaningful ways. Thus, following the indication of Panush & Peritz (1996), the PYL spent in the working life ages is:

$$PYL_W = \frac{1}{2}\sum_{x=20}^{64}[P_x\,(e_{x:\overline{65-x}} + e_{x+1:\overline{65-x-1}})] + e_{20:\overline{65-20}} \cdot l_{20}\sum_{x=0}^{19} P_x L_x^{-1} \quad (2)$$

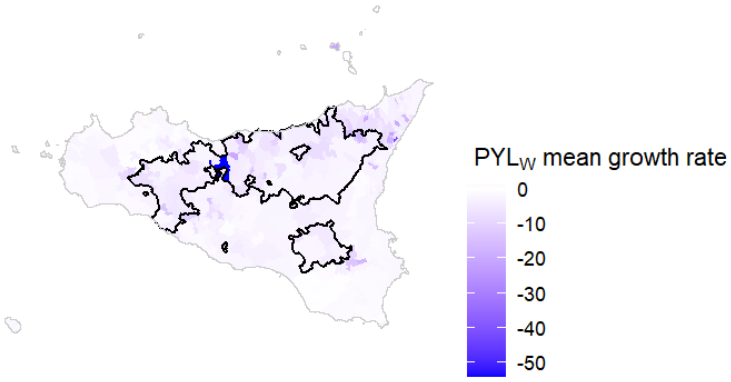
where $e_{x:\overline{65-x}}$ is the expected number of years lived before age 65 by a person now aged *x*. The PYL can be seen as the years of life that the present population will likely experience. The mean growth rate for PYL_W at the municipality level in 2002-2022 is illustrated in Figure 4. Quite all the Sicilian municipalities have experienced a decrease in PYL_w on average during the time period considered. Among them, the areas registering the largest loss are mainly located in the northeastern mountainous regions of Mt. Etna and Nebrodi, and in some inland regions in North-West and South-East. Not all of them are included among the PAs.

This preliminary, and still partial, picture of the demographic distress of Sicilian IAs is intended as a means of discussion and contributes to the debate on NSIA classification with the final scope of effectively allocating resources to the more deprived and marginalized areas of Italy.

3. Discussion and conclusions

The demographic profile of Sicilian municipalities has been highlighted as a crucial aspect that needs to be considered in the identification of areas suffering from population drain. We have demonstrated the complex nature of demographic distress in different areas by analyzing the dynamics of population growth and age distribution over time, and the PYL in working ages. Considering such dynamics can

Figure 4 – *Mean growth rate for PYL_W at the municipality level during 2002-2022. Note: the population age classes considered for the PYL_W computation are [0,19] and [20,65]. Source: authors' elaboration on Istat and NSIA data.*



lead to more targeted and effective intervention plans. The results have been represented cartographically, making it easy to visualize the areas facing greater difficulties and those that are not as peripheral. This visual representation can aid policymakers and stakeholders in identifying priority areas for investment and development initiatives.

The demographic impact of the NSIA project on Italian municipalities during the period 2002–2022 is best understood by considering the initiative's objectives, timeline, and contextual dynamics. While the NSIA was formally implemented from 2014/2015 onward, the historical trends of depopulation and demographic aging in IAs date back several decades, reflecting broader structural issues (Barca et al., 2014). Therefore, the project's direct influence on reversing these trends within the analyzed period is expected to be modest, particularly in the initial years (2002–2014). Evidence from integrated place-based policies suggests that, while measurable impacts on population size or age composition often remain limited in the short term, these initiatives can mitigate the outmigration of young adults and partially rejuvenate local communities when sustained over time (De Rossi, 2021). Consequently, the demographic dynamics observed in NSIA-targeted areas from 2015 to 2022 are likely influenced by a combination of project interventions and prevailing external pressures, including urbanization and global migration patterns (Scrofani and Novembre, 2015). Further longitudinal analyses are necessary to distinguish the initiative's specific contributions from these broader trends.Overall, this study contributes to the ongoing debate on how to support and revitalize inner areas in Italy effectively, and it emphasizes the importance of taking into account various factors beyond the provision of essential services. By adopting a more

comprehensive approach, we can develop better-informed policies that address different municipalities' unique needs and challenges, ultimately fostering sustainable development and minimizing wasteful efforts. As a final note, it is essential to recognize that regional contexts may vary significantly, and the proposed reclassification should be adapted and validated accordingly for other regions in Italy and beyond. Continued research and refinement of such classification models will be essential to ensure that development policies are well-tailored and effective in supporting the growth and well-being of all communities across the country.

Acknowledgements

This study has been partially supported by the PRIN-PNRR research project "Foreign population and territory: integration processes, demographic imbalances, challenges and oppor-tunities for the social and economic sustainability of the different local contexts (For.Pop.Ter)" [P2022WNLM7], funded by European Union—Next Generation EU, compo-nent M4C2, Investment 1.1. It has also been partially supported by the Italian Ministerial grant PRIN 2022 "Depopulation Risk: Experiences, Mobility and Subjective Well-being (DREAMS)" [2022RNKSEL] CUP E53D23010380006. It has also been partially supported by the PNRR - Missione 4 "Istruzione e Ricerca" - Componente 2 "Dalla ricerca all'impresa" - Investimento 1.3 - Partenariato Esteso "GRINS – Growing Resilient, INclusive and Sus-tainable", cod. PE0000018, CUP E63C22002120006, Spoke 8 - "Social sustainabili-ty", finanziato dall'Unione Europea – NextGenerationEU. Finally, it was also partially funded by the European Union - *NextGenerationEU*, in the framework of the *GRINS -Growing Resilient, INclusive and Sustainable* project (GRINS PE00000018 – CUP E63C22002120006). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

References

BARCA F., CASAVOLA P., LUCATELLI S. 2014. Strategia Nazionale Per le Aree Interne. Definizioni, Obiettivi e Strumenti di Governance. Materiali UVAL 2014.

BARCA, F., CARROSIO, G. 2020. Un modello di policy place-based: la Strategia Nazionale per le Aree Interne. In OSTI. G. AND JACHIA E. (Eds.) *Un disegno di rinascita delle aree interne*, Il Mulino, Bologna.

BITONTI F., MAZZA A., MUCCIARDI M., SCROFANI L. 2023. Urban transformations and the spatial distribution of foreign immigrants in Messina. In E. BRENTARI E., CHIODI M., WIT E. (Eds). *Models for Data Analysis: SIS 2018*, Palermo, Italy, June 20–22. Cham: Springer International Publishing, pp. 53-67.

BLANGIARDO G.C. 2013. Italians of today and tomorrow: awareness of next scenarios and search of new equilibrium. *Rivista Italiana di Economia, Demografia e Statistica,* Vol.17, No.2.

BLANGIARDO G.C. 2012. Discovering the "Demographic GDP". *Rivista Internazionale di Scienze Sociali,* Vol.1, pp.45-58.

BLANGIARDO G.C., RIMOLDI S.M.L. 2012. The potential demography: a tool for evaluating differences among countries in the European Union. *Genus,* Vol.68, No.3, pp.63-81.

DE MATTEIS G. 2013. Montagne e aree interne nelle politiche di coesione territoriale italiane ed europee. *Territorio*, Vol.66, pp.7-15.

HERSCH L. 1944. De la démographie actuelle à la démographie potentielle, *Publications de la Faculté des Sciences Économiques et Sociales, 8.* Université de Genève.

ISTAT. 2022. La geografia delle aree interne nel 2020: vasti territori tra potenzialità e debolezze. https://www.istat.it/it/files//2022/07/FOCUS-AREE-INTERNE-2021.pdf. Last Access: 21.11.2023.

NOTA TECNICA NUVAP. 2022. Aggiornamento 2020 della mappa delle aree interne. 20220214-mappa-ai-2020-nota-tecnica-nuvap_rev.pdf.

PANUSH N., PERITZ E. 1996. Potential demography: a second look. *European Journal of Population / Revue Européenne de Démographie,* Vol.12, No.1, pp.27-39.

R CORE TEAM. 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

REGIONE SICILIA. 2022. Strategie Territoriali Regione Siciliana 2021-2027.

SCROFANI L., NOVEMBRE C. 2015. The inland areas of Sicily. From rural development to territorial reorganization. Semestrale di Studi e Ricerche di Geografia Roma - XXVII, 1, gennaio-giugno.

TANTILLO F., LUCATELLI S. 2018. La Strategia nazionale per le aree interne. *Riabitare l'Italia. Le Aree Interne Tra ab-Bandoni e Riconquiste*. In DE ROSSI A. (Ed). Rome: Donzelli Editore, pp. 403–416.

VENDEMMIA B., PUCCI P., BERIA P. 2021. An institutional periphery in discussion. Rethinking the inner areas in Italy. *Applied Geography*, Vol.135, No.102537.

Francesca BITONTI, University of Catania, francesca.bitonti@unict.it
Angelo MAZZA, University of Catania, a.mazza@unict.it

ASSESSING THE DEGREE OF TERRITORIAL EMBEDDEDNESS AMONG POPULATION SUBGROUPS: A FIRST PROPOSAL[1]

Stefania Rimoldi, Federico Benassi

Abstract. In studies examining the settlement patterns of foreign communities within a region, the primary aim often involves assessing the level of spatial segregation compared to the native population. However, such analyses typically overlook the assessment of the historical presence of these communities in the region. Particularly with small-sized communities, their long-standing presence in the region becomes a significant factor in characterizing both the community and the region itself. This permanence can be seen as a form of embeddedness. It's important to note that the duration of a subgroup's presence in the population may not align with the individual members' length of stay. In some cases, while individual members may have short stays, the community as a whole may maintain a consistent presence over time due to ongoing turnover. To address this aspect, we propose an index that measures the permanence of a population subgroup within a region over a specified time interval, relying solely on its total balance, observed during that period. This index could serve as a straightforward measure of population stability or rootedness, capturing only the overall gains or losses (in terms of time) within the subgroup rather than focusing on the separate contribution of births, deaths, immigration, and emigration. This index is complemented by the evaluation of its uncertainty, by assuming the goodness of the indicator is based on the stationarity of the historical series of values computed on the sub-periods. To support the utility of this index, we apply it to population subgroups categorized by citizenship in the Metropolitan Area of Milan.

1. Introduction

In studies examining the settlement patterns of foreign populations within a region, the primary aim often involves assessing the level of spatial segregation compared to the native population. Research in this field is abundant, covering both the development of measures, (among the first, Hoover, 1941; Duncan and Duncan, 1955; Isard, 1960) and their applications (among the latest referring to the Italian context, Bitonti *et al.*, 2023; Conti *et al.*, 2023; Pratschke and Benassi, 2024; Rimoldi *et al.*, 2024). However, such analyses typically overlook the assessment of the historical presence of these populations in the region. A population group's enduring

[1] Sections 1 and 3 are written by Federico Benassi, sections 2 and 4 are written by Stefania Rimoldi.

presence indicates its embeddedness in the area, influencing individual behaviours and affecting the dynamics of segregation and socio-economic transformation within those spaces (Logan *et al.*, 2002; Massey and Denton, 2019; Sensenbrenner and Portes, 2018). For migrant groups, territorial embeddedness can occur as they establish roots in a new location, creating a 'home away from home.' In cities with a high influx of migrants, certain neighbourhoods might become known for their association with specific ethnic or cultural groups. This process allows migrants to preserve their cultural identity and adapt to their new setting. However, it can also pose challenges if territorial embeddedness leads to social isolation or limited integration with the broader community. One important note is that the duration of a group's presence in a territory may not align with the length of stay of individual members. In some cases, while individual members may have short stays, the group as a whole may maintain a consistent presence over time due to ongoing turnover. This is particularly evident in the case of foreign populations, which are often characterised by high variability, with significant fluctuations in inflows and outflows, especially in small areas. This article aims to propose a method for measuring the dynamics of population group permanence within a given territory over a given time interval. The indicator introduced in this study can serve as a complementary measure in population change research, that addresses the role and impact of the different demographic components on demographic shifts (Billari, 2022; Plane and Rogerson, 1994; Rees *et al.*, 2017; Rogerson and Bagchi-Sen, 2023). The focus is to measure the degree of embeddedness of a specific population subgroup within a particular territory, observed over a defined time interval, irrespective of the dynamics of individual components. Being based on the total balance, and considering the cumulative years lived in the area by the subgroup's annual net flows over the period, this indicator helps to shed light on the dynamic of a population group in a given area. Although the application provided here is demonstrated with population groups by country of citizenship, it is also applicable to subgroups defined according to other classification characteristics.

The paper is structured as follows: section 2 is devoted to the methodological construction of the indicator, then section 3 presents its application, and finally, section 4 discusses the limitations and concludes the paper.

2. The indicator of territorial embeddedness

Let X the group of the population, i the area ($i = 1, \dots, r$), t the order number of years $t = 1, \dots, n$, with $n>1$, and $P(X)$ the population of group X, on 1st January of the years t, the total balance of population in each year t is:

$$B_i^t(X) = P_i^{1.1.t+1}(X) - P_i^{1.1.t}(X) \quad (1)$$

Hereafter, we use B^t for $B_i^t(X)$, to simplify the notation.

Under the hypothesis of uniform distribution of flows (births, deaths, immigrations, emigrations and Italian citizenship acquisitions and losses) during each year, we assume that the contribution of each yearly demographic balance in the period $[1; n]$ is:

$$s^t = n - t + 0.5 \quad (2)$$

Then, the total amount of years lived (hereafter YL) by the group X during the period $[1; n]$ is:

$$YL = \sum_{t=1}^{n} B^t \cdot s^t \quad (3)$$

Finally, the average time spent (average years lived, AYL) by the total balance of the group X during the period $[1; n]$ is:

$$AYL = \frac{YL}{\sum_{t=1}^{n} |B^t|} \quad (4)$$

Notice that the total amount of years (3) can be either positive or negative. AYL is indeterminate when $B^t = 0$ for any $t = 1, \dots, n$. In this special case we assume that the permanence of the group in the area is zero, which corresponds to a neutral situation with respect to increasing or decreasing embeddedness. Moreover, $B^t = 0$ for any $t = 1, \dots, n$ corresponds both to null inflows and outflows and compensative flows. This does not influence the average dynamics of the embeddedness of the group in the territory.

The AYL (4) can be made independent from the length of the period examined, by computing the per-year average dynamics of embeddedness:

$$\overline{AYL} = \frac{AYL}{n - 0.5} \quad (5)$$

The $\overline{AYL}$ varies in the interval $[-1; +1]$. It assumes the maximum value (+1) when $\sum_{t=1}^{n} B^t = \sum_{t=1}^{n} |B^t|$ and $B^t = 0 \; \forall \, \mathrm{t} = 2, \dots, \mathrm{n}; \; B^1 > 0$, while conversely it assumes the minimum value (-1) when $\sum_{t=1}^{n} B^t = -\sum_{t=1}^{n} \lfloor B^t \rfloor$ and $B^t = 0 \; \forall \, \mathrm{t} = 2, \dots, \mathrm{n}; \; B^1 < 0$ (see the toy example in Table 1). The maximum value (+1) indicates

full embeddedness, that is a stable presence overtime, while conversely, the minimum value (-1) indicates full disembeddedness, meaning that the population *X* left from the area permanently during the interval $t = 1, \dots, n$.

Table 1 – *Toy example: n=10,* $\sum_{t=1}^{n}|B^t| = \mathbf{100}$*, and the two extreme situations corresponding to: (a) positive balance; (b) negative balance.*

		(a) Positive balance		(b) Negative balance	
t	s^t	B^t	YL	B^t	YL
1	9.5	100	950	-100	-950
2	8.5	0	0	0	0
3	7.5	0	0	0	0
4	6.5	0	0	0	0
5	5.5	0	0	0	0
6	4.5	0	0	0	0
7	3.5	0	0	0	0
8	2.5	0	0	0	0
9	1.5	0	0	0	0
10	0.5	0	0	0	0
		$\sum_{t=1}^{n}\|B^t\|$=100	$\overline{AYL} = 1$	$\sum_{t=1}^{n}\|B^t\|$=100	$\overline{AYL} = -1$

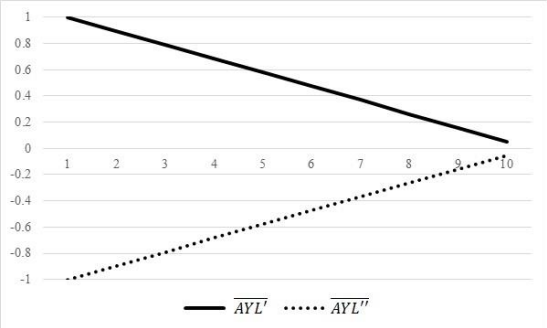
The maximum and minimum circumstances help to shed light on the behaviour of $\overline{AYL}$ overtime. Indeed, when the total amount of balances $\sum_{t=1}^{n}|B^t|$ occurs in just one year *t* of the period, $\overline{AYL}$ is expressed by the following linear equations (6), and illustrated in Figure 1 with reference to the toy example of Table 1.

$$\begin{cases} \overline{AYL'} = -\dfrac{1}{n-0.5}\cdot t + \dfrac{n+0.5}{n-0.5} \quad (\mathrm{t} = 1, \dots, \mathrm{n}) \\ \text{if } \displaystyle\sum_{t=1}^{n} B^t = \sum_{t=1}^{n}|B^t| \text{ and } B^t = 0 \;\forall\, \mathrm{t} \text{ except one} > 0 \\ \overline{AYL''} = \dfrac{1}{n-0.5}\cdot t - \dfrac{n+0.5}{n-0.5} \quad (\mathrm{t} = 1, \dots, \mathrm{n}) \\ \text{if } \displaystyle\sum_{t=1}^{n} B^t = -\sum_{t=1}^{n}|B^t| \text{ and } B^t = 0 \;\forall\, \mathrm{t} \text{ except one} < 0 \end{cases} \tag{6}$$

From Figure 1 it emerges that the indicator proposed is coherent with its meaning: it decreases approaching zero when the total balance occurs in the most recent years, compared to the reference year of the indicator, signalling low embeddedness of *X* if positive, and low disembeddedness if negative, while conversely it increases up to

1 (or -1) when the total balance occurs in the most distant years of the interval, showing a high stable embeddedness (if positive) or disembeddedness (if negative).

Figure 1 – *Toy example:* $\boldsymbol{n = 10, \sum_{t=1}^{n} |B^t| = \pm 100}$ *,* $\overline{\boldsymbol{AYL'}}$ *and* $\overline{\boldsymbol{AYL''}}$ *computed as in (6).*



2.1. Evaluation of the reliability of the indicator

The implicit question when calculating $\overline{AYL}$ is: "Is the value obtained reliable?"

A first consideration refers to the length of the interval, *n*. We theoretically can suppose that the greater is *n*, the more reliable is $\overline{AYL}$. This consideration can serve as a starting point for the empirical strategy we chose. Thus, we imagine period *n* being broken up into sub-periods, and $\overline{AYL}$ is calculated for all the possible n-2 sub-periods (with at least 2 balances), starting from the most recent year: $k = [n; n-2], [n; n-3], \ldots, [n, 2]$.[2] We assume the goodness of indicator be based on the assumption of stationarity of the historical series of $\overline{AYL}_k$. We test the stationarity with the Dickey–Fuller Test (1979), to test whether the variable follows a random walk. The null hypothesis is that the variable follows a random walk. The Test provides results for the three critical values 0.01, 0.05, and 0.1.

2.2. Alternative strategy: the location quotient of embeddedness

It is essential to note that stationarity is not always a valid assumption, particularly in dynamic or evolving systems where trends, cycles, or structural breaks may occur. An alternative strategy consists in comparing the index computed with reference to the specific population group *X* to the total population *P*, or to another group *Y*. Considering the groups by country of citizenship, for example, we may be

[2] Notice that $\overline{AYL}_k$ with $k = [n, 1]$, gives the $\overline{AYL}$

interested to compare the intensity of embeddedness of group *X* in a certain territory (relative to its average embeddedness in the entire region), with the same indicator computed for total population *P*. To do this, it is first necessary to apply the min/max procedure to the index so that it varies between 0 and 1:

$$\overline{\overline{AYL}} = \frac{\overline{AYL} - (-1)}{1 - (-1)} = \frac{\overline{AYL} + 1}{2} \tag{7}$$

with $\overline{\overline{AYL}} \in (0.5;1]$ indicating embeddedness, $\overline{\overline{AYL}} \in [0;0.5)$ indicating disembeddedness, and $\overline{\overline{AYL}}=0.5$ indicating the case when $B^t = 0 \; \forall \; t = 1, \ldots, n$, (the AYL is indeterminate), where we assume that the embeddedness of the group in the area is zero, which corresponds to a neutral situation with respect to increasing or decreasing embeddedness.

Then, in analogy to the calculus of the Location Quotients (Isard, 1960), given the group *X* of the population *P*, in each sub-area *i* of the region *R*, we can compare the ratio between the embeddedness of *X* in *i*, $\overline{\overline{AYL}}_i(X)$, and the embeddedness of total population *P* in *i*, $\overline{\overline{AYL}}_i(P)$, with the ratio between the average embeddedness of *X* in *R*, $\overline{\overline{AYL}}_R(X)$, and the average embeddedness of *P* in *R*, $\overline{\overline{AYL}}_R(P)$, deriving the following relative measure of location quotient of embeddedness:

$$ELQ_i^{X/P} = \frac{\overline{\overline{AYL}}_i(X) / \overline{\overline{AYL}}_i(P)}{\overline{\overline{AYL}}_R(X) / \overline{\overline{AYL}}_R(P)} \tag{8}$$

The interpretation of the ELQ is analogous to the interpretation of the *LQ*; it provides a measure of comparison between the embeddedness of the group *X* relative to *P* in the sub-region *i*, compared to the embeddedness of group *X* relative to *P* in the whole region *R*.

3. Empirical applications

Solely for the purpose of illustrating the performance of the indicator, we propose the following application, where we use the total population on 1st January of the years 2004–2022 by country of citizenship, Italians and foreigners. The geographical setting analysed is the Metropolitan Area of Milan, i.e. the province of Milan, with the addition of the municipalities of the province of Monza-Brianza (formerly to 2009 included in the province of Milan). The city of Milan is subdivided in the 88 administrative neighbourhood called NIL (Nuclei di Identità Locale). The indicator

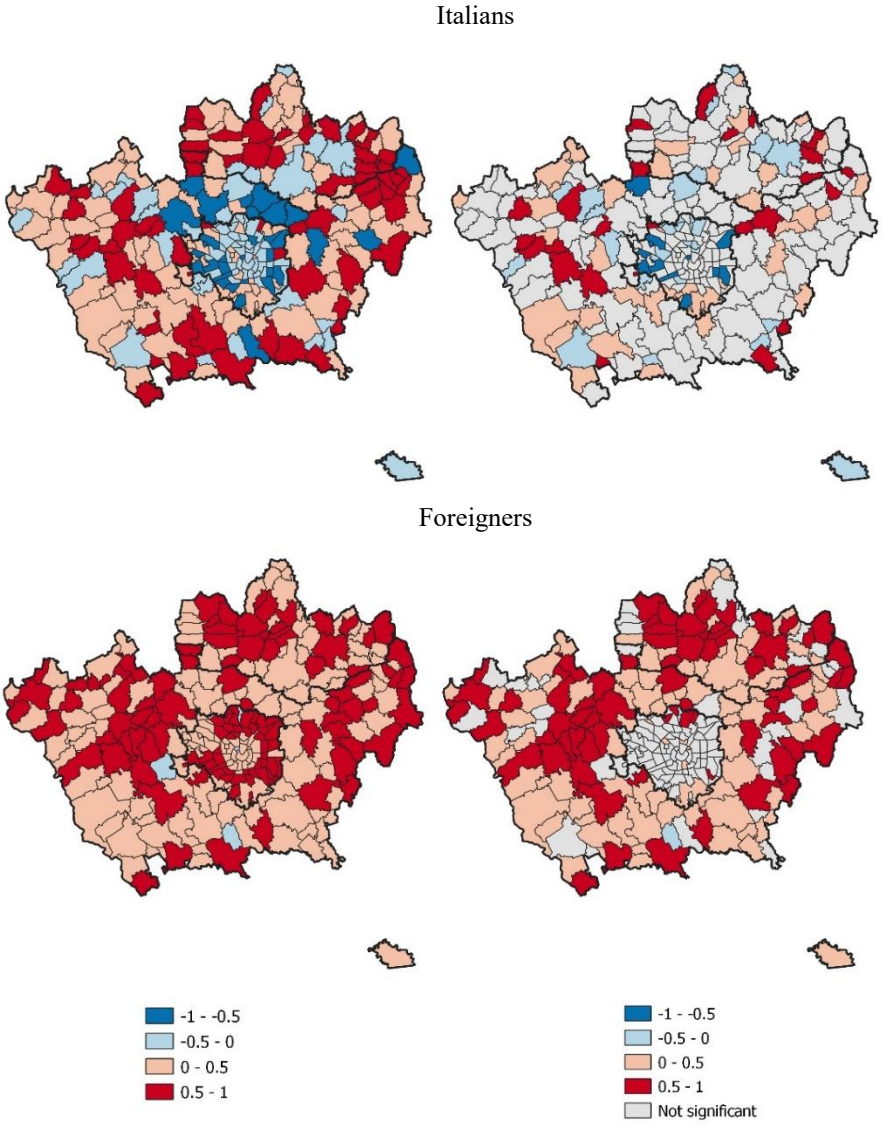
was also tested using the Dickey–Fuller Test, with a critical value set at 0.1. Regarding the critical value of the test, we opted for the most inclusive value concerning the time series stability. Table 2 provides the results of the test for the three critical values.

Table 2 – *Dickey-Fuller Test for* $\overline{\boldsymbol{AYL}}$*: number of significant values by critical values, for Italians and foreigners.*

	Critical values			
	0.01	0.05	0.1	Total
Italians	79	17	12	108
Foreigners	133	20	11	164

Figure 2 (left side) highlights how, during the period from 2004 to 2022, the historical centre of Milan was characterised by disembeddedness for both populations. Italians (upper panel) show the highest level of embeddedness in the most peripheral municipalities of the metropolitan area and the highest levels of disembeddedness in the city and in the contiguous northern municipalities. However, these results show little reliability (upper panel, right side), except for the following cases. We find a high level of embeddedness (0.5 – 1] in: the NlL of *Muggiano*, showing the highest level of embeddedness (0.72), the cluster of municipalities in the north-western direction (composed of *Cisliano, Corbetta, Marcallo con Casone, Mesero*, and *Santo Stefano Ticino*) −this area, known as Parco del Gelso, is mainly characterised by agricultural activity−, the small cluster of *Cernusco S.N.* and *Vimodrone*, north-eastern municipalities contiguous to Milan, and some municipalities in the province of Monza-Brianza, like *Corezzana*, *Sulbiate*, and *Sovico*. A moderate (0 – 0.5] embeddedness is observed in the cluster of territorial units composed of peripheral contiguous southern NILs in the city, such as *Quintosole, Stadera, Cantalupa, Parco delle Abbazie*, and *Parco dei Navigli*, and municipalities (*Buccinasco, Opera*, and *Trezzano S.N.*). Conversely, a relatively high level of disembeddedness (-1 – -0.5) is observed in some peripheral NILs all around the city, such as *Ortomercato* (-0.58)*, Baggio, Parco Lambro and Gallaratese* (-0.57)*, Mecenate and Trenno* (-0.54)*, Gratosoglio, Giambellino and Forze Armate* (-0.53)*,* and *Quinto Romano (-0.51).*

Figure 2 – $\overline{\boldsymbol{AYL}}$ *(left side) and* $\overline{\boldsymbol{AYL}}$ *filtered with the Dickey–Fuller Test (right side), for Italians and foreign citizens, in the Milan MA, 2004-2022.*



Foreigners (lower panel) exhibit positive embeddedness almost everywhere, particularly high (0.5–1) in the city outside the historical centre (left side), though not significantly stable (right side). The main areas of higher stable embeddedness

are: 18 contiguous municipalities in the western part of the metropolitan area −a large industrial area stretching from *Gaggiano* in the south to *Cerro Maggiore* in the north−, 16 contiguous municipalities, all in the province of Monza-Brianza (except for *Paderno D.*) −including *Desio, Cesano M., Carate B.*, up to *Lentate S.S.*, known as the "Distretto del mobile della Brianza" −, a small cluster of four municipalities in the province of Monza-Brianza around *Vimercate*, and 5 contiguous municipalities at the border of the province of Milan with the province of Bergamo, including *Trezzo sull'Adda* and *Cornate d'Adda*.

Figure 3 – *ELQ for foreigners (total population as reference), in the Milan MA, 2004-2022.*

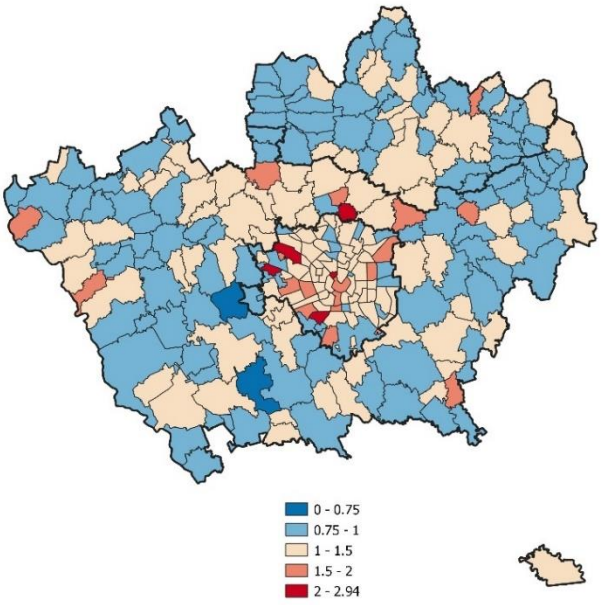


Figure 3 shows the results of $ELQ_i^{X/P}$ for foreigners, with reference to the total population P. Overall, the local relative embeddedness of foreigners is markedly different from the relative embeddedness of total population in very few areas. It is sensibly low (< 0.75) in only two municipalities, *Cusago* and *Noviglio*, in the western part of the Milan MA, while it is very high (≥ 2) in only the 4 NILs of *Parco Sempione* (2.94)[3], *Barona* (2.49), *Gallaratese* (2.14) and *Quinto Romano* (2.11), and in the municipality of *Bresso* (2.06). High (1-5 – 2.0) intensity of ELQ is found in

[3] Notice that in *Parco Sempione* population is reported as 1 to 3 units, during the period.

few NILs (*Duomo, Ticinese* and *Tibaldi* in the historic centre, *Gratosoglio* and *Ronchetto S.N.* in the south, *Bande Nere* and *Forze Armate* in the west, and finally, *Parco Lambro* and *Città Studi* in the east), and in sparse municipalities, such as *Cologno Monzese* and *Cusano Milanino* near the city border, *Garbagnate M.* in the north, *Turbigo* and *Boffalora T.* in the west, and finally *Bussero* and *Carnate* (the only one in the province of Monza- Brianza) in the north-east. Finally, a large stream of contiguous municipalities in the north is characterised by moderate (1 – 1.5) intensity of ELQ.

4. Discussion and conclusions

In this paper, we propose a new index for evaluating the dynamics of the permanence of a population group within a given territory over a specified time period. The index $\overline{AYL}$ is intended to convey the meaning of embeddedness in the territory, whereby the stable presence of a community contributes to defining local identity by reflecting the interaction between social and physical factors. These identities significantly influence how specific neighbourhoods are perceived as desirable places to reside and settle (Robertson *et al.*, 2010). Based exclusively on yearly total balances, the index includes all types of dynamics—natural, migratory, and administrative. The latter is particularly relevant for migrants who may acquire the citizenship of the host country during the period of observation, thus representing a flow. Overall, for the purposes of the index, the distinction between natural and migratory balances is not relevant, as the concept of embeddedness of a population group can derive from both. Conversely, flows resulting from changes in citizenship status should be carefully managed, when relevant (i.e. when dealing with national contexts), due to their impact on the concept of the group's embeddedness. Also at a local level, anyway, the effect of changes in citizenship status can be relevant and constitute a limitation of the index, in this form. Intuitively, a possible solution could involve applying a corrective factor based on the difference between residents by citizenship and by country of birth, data recently made available by the National Institute of Statistics (Istat) at the municipal level.

Moreover, as mentioned in the introduction, the duration of a group's presence in a territory may not align with the length of stay of individual members. Individual members may have short stays, yet the group as a whole can maintain a consistent presence over time due to ongoing turnover. In summary, since the index measures the average time spent by the yearly net balances of the population group during the period, the length of the period n, becomes a primary concern. It is desirable for n to be sufficiently long to give meaning to the concept of embeddedness. However, a period that is too long may encompass contextual events (such as years of economic

recession) that can temporarily but significantly alter the dynamics of the balances. This eventuality also impacts the stationarity of the historical series of values of the indicator, for which a test is proposed. Beyond the knowledge of the general context, then, a preliminary analysis of the trend of balances is unavoidable, and a series of indices computed for sub-periods can be useful to evaluate the overall dynamics.

Further limitations can derive from the size of the group and of the territory. Although the index is independent of the population size, as it is based solely on annual balances, a primary consideration is the appropriateness of computing such an index for very small populations and very small territories. Secondly, the gravitational effects of larger populations on the territory and their growth rate should be considered. In our application, focusing on Italians, we explored these relationships through simple correlations. Both the correlations between the index and the average population, and between the index and the average annual growth rate, resulted in negligible negative values.

These and other aspects of the index will be treated and experimented on other populations and contexts in next research.

Acknowledgements

This research contribution is conceived and realized as part of two projects:
- METROMOSAIC project, Cariplo Foundation 2021-1690.
- PRIN 2022-PNRR research project "Foreign population and territory: integration processes, demographic imbalances, challenges and opportunities for the social and economic sustainability of the different local contexts (For.Pop.Ter)" [P2022WNLM7]. Funded by European Union – Next Generation EU, component M4C2, Investment 1.1.

References

BILLARI, F. C. 2022. Demography: Fast and slow. *Population and development review*, Vol. 48, No. 1, pp. 9-30.

BITONTI, F., BENASSI, F., MAZZA, A., STROZZA, S. 2023. From South Asia to Southern Europe: A comparative analysis of Sri Lankans' residential segregation in the main Italian cities using high-resolution data on regular lattice geographies, *Genus*, Vol. 79, No. 1, pp. 1-27.

CONTI, C., MUCCIARDI, M., SIMONE, M. 2023. Exploring the Settlement Models of the Main Foreign Communities Residing in Italy (2003–2021). *Social Sciences*, Vol. 12, No. 9, pp. 1-16.

DICKEY, D. A., FULLER, W. A. 1979. Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, Vol. 74, pp. 427–431.

DUNCAN, O. D., DUNCAN, B. 1955. A methodological analysis of segregation indexes, *American Sociological Review,* Vol. 20, No. 2, pp. 210-217.

HOOVER, E. M. 1941. Interstate Redistribution of Population, 1850-1940, *The Journal of Economic History*, Vol. 1, No. 2, pp. 199-205.

ISARD, W. 1960. *Methods of regional analysis: an introduction to regional science*. Cambridge: The MIT Press.

LOGAN, J. R., ALBA, R. D., ZHANG, W. 2002. Immigrant enclaves and ethnic communities in New York and Los Angeles. *American sociological review*, Vol. 67, No. 2, pp. 299-322.

MASSEY, D. S., DENTON, N. A. 2019. American apartheid: Segregation and the making of the underclass. In: GRUSKY D. (Ed.) *Social Stratification, Class, Race, and Gender in Sociological Perspective, Second Edition*, pp. 660-670. Routledge.

PLANE, D.A., ROGERSON, P.A. 1994. *The geographical analysis of population with applications to business and planning.* New York: Wiley.

PRATSCHKE, J., BENASSI, F. 2024. Population Change and Residential Segregation in Italian Small Areas, 2011–2021: An Analysis With New Spatial Units. *Spatial Demography*, Vol. 12, No. 2, pp. 1-30.

REES, P., BELL, M., KUPISZEWSKI, M., KUPISZEWSKA, D., UEFFING, P., BERNARD, A., CHARLES-EDWARDS, E., STILLWELL, J. 2017. The impact of internal migration on population redistribution: An international comparison. *Population, Space and Place*, Vol. 23, No. 6, e2036.

RIMOLDI, S.M.L., CRISCI, M., BENASSI, F., RAYMER, J. 2024. Intra-urban residential mobility and segregation of foreigners in Rome, *Population Space & Place*, e2777

ROBERTSON, D., MCINTOSH, I., SMYTH, J. 2010. Neighbourhood identity: The path dependency of class and place. *Housing, Theory and Society*, Vol. 27, No. 3, pp. 258-273.

ROGERSON, P. A., BAGCHI-SEN, S. 2023. The average effectiveness of local migration exchanges. *Population, Space and Place,* Vol. 29, No. 7, e2690.

SENSENBRENNER, J., PORTES, A. 2018. Embeddedness and immigration: Notes on the social determinants of economic action. In: SENSENBRENNER J, PORTES A. (Eds.) *The sociology of economic life*, Routledge, pp. 93-115.

Stefania M.L. RIMOLDI, Università di Milano-Bicocca, stefania.rimoldi@unimib.it
Federico BENASSI, Università di Napoli Federico II, federico.benassi@unina.it

ROBUSTNESS TECHNIQUES FOR THE EVALUATION OF MUNICIPAL DATA: AN EXAMPLE ON THE "A MISURA DI COMUNE" INFORMATION SYSTEM[1]

Agata Maria Madia Carucci, Valeria Marzocca, Roberto Antonello Palumbo

Abstract. The "A misura di Comune" information system is the result of the "Misure di benessere e programmazione a livello comunale" project started in 2016 and it was created with the explicit objective of supporting public decision-makers in intercepting the needs of a territory. Through a complex set of indicators useful for the planning, programming and management tasks of Local Authorities, it is also a support tool for the preparation of the Single Programming Document (SDP).

In this work, after having presented the main methods of collecting statistical information, the robustness techniques of the proposed indicators will be analysed.

From 2022, the indicators are widespread in a time series and therefore alignment with the new geography and robustness analysis of the indicator are necessary annually not only in cross-sectional terms but also longitudinally. In particular, the robustness tests of the indicators depend on the type of indicator produced and three examples will be presented in this work.

1. Introduction

Territorial development and growth policies must necessarily start with analytical and continuously updated knowledge of the social, economic, demographic and environmental conditions of the area on which action must be taken. At the basis of good governance there are data and the ability to make decisions based on exact knowledge of the area. The availability of territorial statistical data is an indispensable requirement for defining, implementing and evaluating policies. To obtain measurable policies, it is necessary to have quality, timely and comparable data across time and space (Istat, 2018).

Among the objectives of official statistics there is certainly that of expanding the offer of territorial statistics on the main environmental and socio-economic phenomena of the country, making timely information available that reads the transformations of Italian society (Crescenzi F., Lipizzi F., 2020).

[1]The paper was carried out by the joint work of the authors. More in detail, the single paragraphs are attributed as follows: paragraphs 3 and 3.1 to Agata Maria Madia Carucci; paragraphs 1, 2, 4 to Valeria Marzocca; paragraphs 3.2 and 3.3 to Roberto Antonello Palumbo.

"A misura di Comune" represents an information tool designed by Istat which collects numerous municipal indicators on the main socio-economic and environmental issues. The tool, easy to consult, is also of central interest in the preparation of the Single Planning Document[2].

In this work, after having illustrated the structure and potential of "A misura di Comune" and the related use of indicators by municipal authorities, the techniques used to control and validate the constructed indicators will be illustrated through examples.

The growing demand for data at municipal level is also justified by the presence, in Italy, of a strong territorial lack of homogeneity which is linked to different access to services. Small and very small municipalities (less than 1,000 inhabitants) represent 25% of the municipalities in Italy and over one million inhabitants live in them. In the latest annual report, Istat dedicates an entire chapter to the territories, with the view that "The territory, with its economic, demographic, social and cultural specificities, represents a moment of synthesis of the complex transformations taking place at a national level and global, highlighting specific potential and constraints compared to what is illustrated at a national level. The theme of accessibility, understood as the possibility of access for citizens and businesses to various services, is closely linked to that of the peripherality of territories and to policy strategies oriented towards territorial planning" (Istat, 2024).

2. The information system "A misura di Comune"

The thematic structuring of "A misura di Comune", reported in Table 1, presents a strong coherence with the classification of indicators for measuring well-being and sustainable development and it tries to respond to the main information needs described by the Single Programming Documents. The information system, with the latest update in March 2024, collects 72 indicators for 13 thematic areas.

"A misura di Comune" is a multi-source system, in which sources of an experimental nature are valorized alongside other more consolidated ones. Among the experimental sources, an important place goes to the databases created within the ARCH.I.M.E.DE Project, which deals with the construction and updating of databases for territorial analysis within the Integrated Microdata System of the Istat and whose data are disseminated for municipalities with more than 5,000 inhabitants. A significant contribution also concerns the use of Open Data made available by other Sistan entities, such as the Ministry of the Interior, the Ministry of Economy and Finance, the Italian Institute for Environmental Protection and Research (ISPRA), the Italian Automobile Club (ACI) and the National institute for the evaluation of the education and training system (INVALSI). Other sources are

[2] https://www.istat.it/statistica-sperimentale/aggiornamento-degli-indicatori-del-sistema-informativo-a-misura-di-comune/

included in the current statistical production of Istat, such as for example demographic statistics, statistical registers of active businesses and non-profit institutions, the "Dati ambientali nelle città" survey (Table 2).

Table 1 – *"A misura di comune" Indicators by thematic area, release 2022 and 2024*.

THEMATIC AREA	N. Indicators 2022	N. Indicators 2024	% Indicators by thematic area 2024
Popolation	10	10	13,9
Household	0	6	8,3
Education	3	4	5,6
Labour	0	4	5,6
Economic well-being	1	4	5,6
Politics and institutions	5	5	6,9
Culture	2	3	4,2
Gender Gap	0	10	13,9
Social services	0	2	2,8
Territory and environment	5	9	12,5
Settled economy	5	6	8,3
Research and innovation	1	1	1,4
Infrastructure and mobility	7	8	11,0
TOTAL	39	72	100

Source: Authors elaboration based on ISTAT data.

Table 2 – *"A misura di comune" Indicators by source*.

Source	N. Indicators	% Indicators by source
ISTAT	**59**	**81,9**
ARCH.I.M.E.DE (Integrated administrative sources)	*6*	*8,3*
ASIA Business Register	*11*	*15,3*
PERMANENT CENSUS OF POPULATION AND HOUSING	*7*	*9,7*
Current statistics	*35*	*48,6*
OPEN DATA	**13**	**18,1**
TOTAL	72	100

Source: Authors elaboration based on ISTAT data.

In the latest dissemination, the indicators were presented in a time series starting from 2014 and 80% of them are available at the individual municipality level.

3. Indicator Selection: construction and control methods

The choice of indicators, in addition to being derived from the information needs of the institutions, is linked to the availability of the data and its quality (Istat, 2012). In general, the indicators were chosen if they meet the general criteria of:

- Timeliness
- Coverage
- Completeness and
- Consistency.

According to the type of indicator and the source that feeds it, macro and micro controls have been envisaged on the input data and on the output produced.

In general, for each input dataset, it is verified that the underlying definitions have not changed to ensure the comparability of the information with each other and subsequently a univariate and multivariate analysis is carried out of the time series data.

With univariate analysis, the single time series is verified in its temporal dynamics. The analysis starts from the basic data, on which the indicator is built, and is supported by the graphic representation of the values for each indicator, starting from the regional domain and gradually descending towards territorially finer domains.

From the graphic representation and through synthetic measures, calculated on each individual basic data, qualitative considerations can be drawn on the series and "causal factors" are "recognised" within the temporal dynamics of the data (in the period considered for example, many trends affected by the pandemic event), aimed at capturing the characteristics of the historical series.

With multivariate analysis, the connections and relationships between the observed phenomenon and others considered correlated (population trend, value added of enterprises trend, etc.) are subsequently verified.

Once the consistency of the data in the historical series has been verified, the coverage is analysed at a territorial level, reconstructing the territorial perimeter of the series at the last reference year and favouring the data for which there is exhaustive municipal coverage.

The analysis of the basic data is followed by the construction of the indicator and again the verification of the indicator produced through simple and complex statistical methods to identify irregularities and inhomogeneities in the indicators produced.

The analysis, conducted on three structurally very different indicators, will be shown below:

- Taxable income per taxpayer,
- Entrepreneurship rate
- Libraries registered in the National Library Registry per 100,000 inhabitants.

3.1 Taxable income per taxpayer

The average income per taxpayer is an indicator from an open data source, the Ministry of Economy and Finance, which annually disseminates statistical data on tax returns by taxpayer characteristics, type of income and municipality.

At a territorial level the data is exhaustive, all municipalities are surveyed annually. Furthermore, the data are disseminated every year, in the month of April, on the MEF's open data platform; they are comparable over time and space as it was not subject to any change in definition or data collection methodology over the period considered.

After analysing the input data with descriptive statistical analysis techniques, the trend of the historical series was compared with the trend of the main "well-being" indicators published by official statistics to identify anomalies or particularities in the trend of the data and to highlight the territorial characteristics of information.

Below are some simple examples of statistical analysis on the input data and the indicator that, expressed in euros, is calculated as the ratio between Taxable income and Taxable income taxpayers.

On average, an Italian taxpayer has an income of 21,000 euros, stable over the years, confirming the well-known Italian North-South dualism. In particular, the income per taxpayer in the North-West of Italy is 24,600 euros while in the South is around 17,500 euros (Table 3).

Table 3 – *Taxable income per taxpayer by territory. Years 2014-2021.*

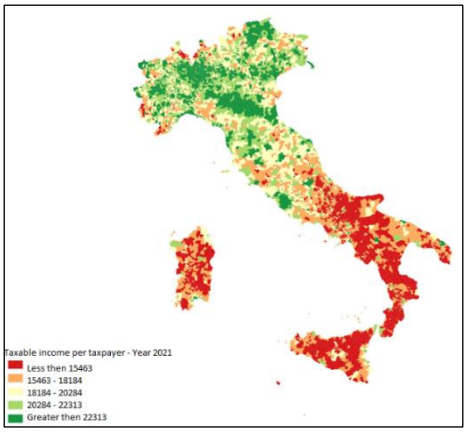
TERRITORY	2014	2015	2016	2017	2018	2019	2020	2021
Nord-West	22,335	22,741	22,910	23,074	23,608	23,789	23,518	24,664
Nord-East	20,734	21,119	21,413	21,565	22,043	22,285	22,099	23,120
Center	20,566	20,782	20,996	21,078	21,519	21,713	21,566	22,428
South	15,872	16,105	16,244	16,326	16,658	16,831	16,864	17,566
Islands	16,131	16,232	16,383	16,513	16,845	17,004	17,016	17,672
ITALY	19,720	20,017	20,214	20,355	20,800	20,993	20,852	21,777

Source: Authors elaboration based on MEF data.

To identify any outlier values of the income indicators at municipal level, the analysis was paid particular attention to the marked territorial configuration of the income phenomenon in Italy. Figure 1 represents the municipal values of the Income per taxpayer indicator.

The distribution of the municipalities in the five income classes is immediately consistent with the well-known regional differences in per capita income, also confirmed by the MEF data. The analysis of the data of the municipalities that are in the two tails of the distribution of per capita income also allows us to highlight the different cases within the general pattern of pronounced lack of homogeneity of the North-South area.

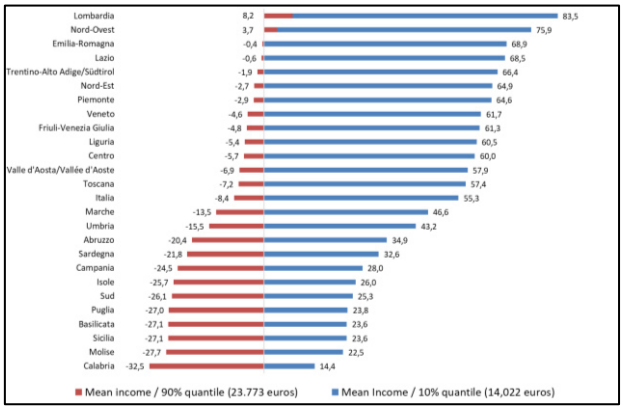
Figure 1 – *Taxable income per taxpayer- Year 2021.*



Source: Authors elaboration based on MEF data.

Considering, for the year 2021, the total of Italian municipalities for which information is available, the ninth decile threshold is equivalent to 23,773 euros, which is 9.2% higher than the national average value, but even lower than the average value for Lombardy. Conversely, this threshold is more than 20% higher than the income per taxpayer of all the regions of Southern Italy. Examining the first decile threshold (equal to 14,022 euros), it is 35.7 percentage points lower than the national average value, but the average income in Lombardy is 83.5% higher. On the contrary, the average income in the South is 25% higher than this threshold, in Calabria the average income is just 15% higher than the first decile (Figure 2).

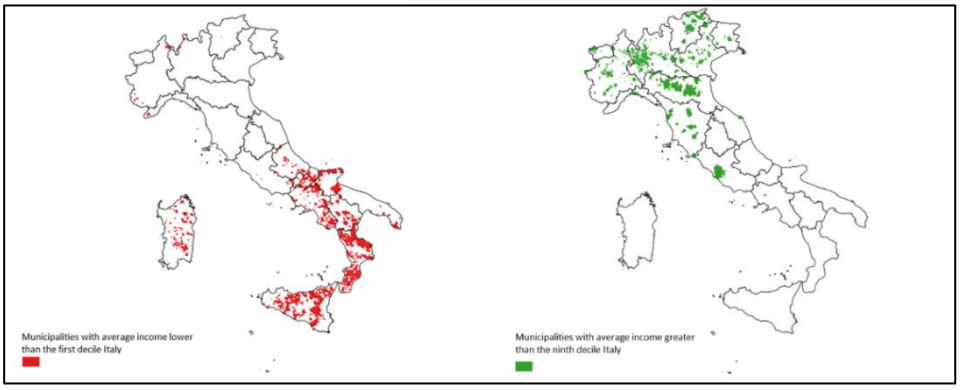
Figure 2 – *Taxable income per taxpayer: percentage distance between Mean income of the territory and first decile or ninth decile of Italy. Year 2021.*



Source: Authors elaboration based on MEF data.

At the level of the individual municipality, the municipalities in the first decile are almost exclusively in the regions of the South, apart from a few cross-border municipalities, while the municipalities whose indicator falls in the ninth decile of the distribution are in the regions of the North and Lazio, particularly in the Roman belt (Figure 3).

Figure 3 – *Taxable Income Per Taxpayer- Municipality in the first and ninth decile. Year 2021.*



Source: Authors elaboration based on ISTAT data.

In conclusion, the results of this analysis support the belief in the "spreadability" of the indicator, which is stable over the years, complete in terms of coverage, timely due to the availability of the source and coherent at a territorial level.

3.2 Entrepreneurship rate

The entrepreneurship rate is a classic indicator used for the analysis of the production system at a territorial level and the data appear to be particularly "solid" as they derive from the Business Register ASIA (Eurostat, 2010). The indicator therefore does not present coverage and timeliness problems, and the data is available for the entire historical series. The trend of the indicator was compared with the trends of the main indicators on the structure of enterprises released by official statistics and did not highlight any "irregularities" in the information collected. In addition, the output analysis cannot ignore the territorial analysis, as although the basic data are very robust by construction, the indicators at municipal level could highlight outliers value. On average, there are just over 70 enterprises per thousand inhabitants and this value is substantially stable over the years (Table 4). In the North of Italy, there are 81.2 enterprises per thousand inhabitants while in the South there are 64.1 enterprises.

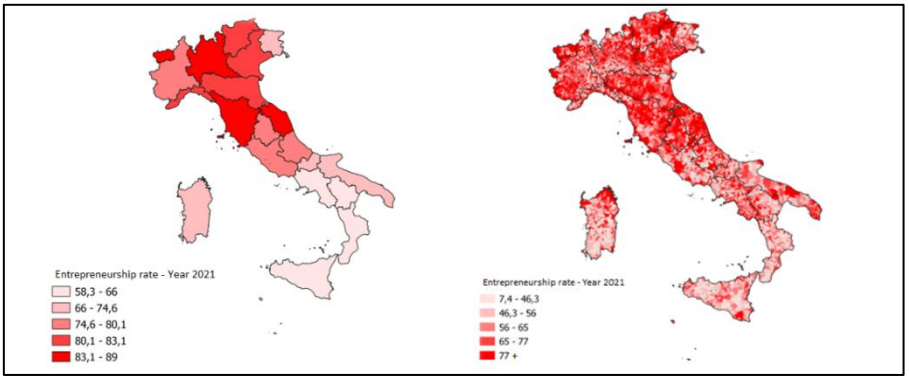
Table 4 – *Entrepreneurship Rate- enterprises per 1,000 inhabitants. Years 2014-2021.*

YEARS	2014	2015	2016	2017	2018	2019	2020	2021
Mean of the indicator	72.3	72.0	73.0	73.3	73.6	72.2	73.5	75.7
50% quantile	58.8	58.4	58.9	58.8	58.9	57.3	58.8	60.6
90% quantile	85.2	84.9	85.5	85.4	85.5	83.3	84.9	87.0
Max value	267.2	273.4	276.9	269.4	272	271.7	260.2	255.6

Source: Authors elaboration based on ISTAT data.

From the regional map, we can see the classic North-South dualism with two exceptions: Friuli-Venezia Giulia and Piedmont. Among the regions of Southern Italy, Abruzzo records the highest indicator value. Dualism is confirmed by the municipal representation even if some regional specificity is highlighted (Figure 4).

Figure 4 – *Entrepreneurship rate, Regional and municipal map, Year 2021.*



Source: Authors elaboration based on ISTAT data.

There is a low variability of the indicator between years (coefficient of variation less than 5% in 61% of the municipalities) and this variability is even lower as the municipal size increases, the 88.6% of the municipalities with a coefficient of variation greater than 30% has fewer than 1000 inhabitants (Table 5).

Table 5 – *Entrepreneurship rate: CV of the municipalities by demographic size.*

COEFFICIENT OF VARIATION CLASSES	N. OF MUNICIPALITIES (class of inhabitants)				
	Totals	< 1,000	> 1,000	> 5,000	> 10,000
Less then 5%	4,822	313	4,509	2,216	1,143
5%-10%	2,221	938	1,283	148	59
10%-20%	717	624	93	4	2
20%-30%	100	94	6	2	0
Greater then 30%	44	39	5	1	0

Source: Authors elaboration based on ISTAT data.

3.3 Libraries registered in the National Library Registry per 100,000 inhabitants.

The last indicator analyzed is the number of libraries per 100,000 inhabitants. It is a particularly different indicator compared to the previous ones.

From the point of view of the availability and timeliness of the data, they are provided periodically for all Italian municipalities, upon specific request, directly from the National Library Registry. As regards the quality analysis, it mainly involves detecting the presence or absence of the library (zero or one library, 27.3% and 58.5% respectively). On average, there are 20 libraries per 100,000 inhabitants and the number is stable over the years (Table 6).

Table 6 – *Libraries Per 100,000 Inhabitants.*

YEARS	2016	2017	2018	2019	2020	2021
Freq Mcp with value=0	2,023	2,139	2,320	2,535	2,452	2,158
Mean	24.1	23.2	22.9	20.5	20.8	21.9
50% quantile	20.9	19.8	18.3	16.1	17.0	19.7
90% quantile	108.0	105.5	103.8	101.4	104.2	111.4
MAX	2,816.9	1,470.6	1,626.0	3,076.9	2,739.7	2,439.0

Source: Authors elaboration based on ISTAT data.

From the cartograms we do not find a territorial dichotomy but rather a strong relationship between the number of libraries and the municipal size. In very small municipalities the presence of even just one library leads to a significant increase in the indicator. The maximum value, for example, is recorded in the municipality of Moncenisio (TO) where there is a library and 42 inhabitants.

The areas with the highest value of the indicator are in fact generally those in which there is a greater presence of small and very small municipalities (Table 7).

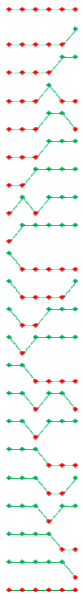
Table 7 – *Number of libraries per 100,000 inhabitants (I) by demographic size of municipalities (Mcp)- Year 2021.*

	< 1,000 inhabitants		1,000-5,000 inhabitants		5,000-10,000 inhabitants		10,000 + inhabitants		Total	
	# Mcp	*I*	# Mcp	*I*	# Mcp	*I*	# Mcp	*I*	# Mcp	*I*
North-West	401	174.5	1,027	41.6	362	16.8	278	17.9	2,068	24.4
North-East	101	162.6	565	43.7	298	18.1	269	24.1	1,233	26.4
Center	58	155.9	292	41.7	137	21.4	211	22.2	698	24.0
South	158	164.7	456	46.4	174	20.3	276	16.4	1,064	21.0
Islands	143	174.0	321	43.4	92	17.9	127	14.3	683	20.7
ITALY	861	169.7	2,661	43.0	1,063	18.4	1,161	19.4	5,746	23.6

Source: Authors elaboration based on ISTAT data.

The analysis of the presence or absence of the library can also be conducted through simple graphic representations, the so-called sparklines which allow cases of inconsistency to be intuitively highlighted (Table 8).

Table 8 – *Sparklines of the municipal libraries*.

2016	2017	2018	2019	2020	2021	Freq	% Freq	
0	**0**	**0**	**0**	**0**	**0**	**1,648**	**20.85**	
0	**0**	**0**	**0**	**0**	**1**	**168**	**2.13**	
0	**0**	**0**	**0**	**1**	**1**	31	0.39	
0	**0**	**0**	**1**	**0**	**0**	1	0.01	
0	**0**	**0**	**1**	**1**	**0**	6	0.08	
0	**0**	**0**	**1**	**1**	**1**	**130**	**1.64**	
0	**0**	**1**	**1**	**1**	**1**	27	0.34	
0	**1**	**0**	**1**	**1**	**1**	1	0.01	
0	**1**	**1**	**1**	**1**	**1**	10	0.13	
1	**0**	**0**	**0**	**0**	**0**	3	0.04	
1	**0**	**0**	**0**	**0**	**1**	5	0.06	
1	**0**	**0**	**1**	**1**	**1**	104	*1.32*	
1	**0**	**1**	**1**	**1**	**1**	15	0.19	
1	**1**	**0**	**0**	**0**	**0**	4	0.05	
1	**1**	**0**	**1**	**1**	**0**	1	0.01	
1	**1**	**0**	**1**	**1**	**1**	218	*2.76*	
1	**1**	**1**	**0**	**0**	**0**	469	5.93	
1	**1**	**1**	**0**	**0**	**1**	151	1.91	
1	**1**	**1**	**0**	**1**	**1**	55	*0.7*	
1	**1**	**1**	**1**	**0**	**0**	2	0.03	
1	**1**	**1**	**1**	**1**	**0**	23	0.29	
1	**1**	**1**	**1**	**1**	**1**	**4,831**	**61.13**	

Source: Authors elaboration based on ISTAT data.

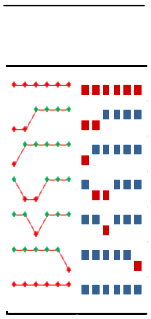
In the entire period, 20% of Italian municipalities do not have a library, just as 61% of municipalities always have at least one library and this behavior of the indicator can be said to be reasonable. Equally reasonable is that the event is present from a certain point onwards: in 2% of cases, the library is taken over starting from 2021, as well as in 1.6% of cases from 2019. Especially due to pandemic events, even the closure of a library from a certain year onwards cannot be considered an anomalous event.

The anomalous values, on the other hand, on which it is appropriate to reflect, are those for which the event is present in some years, then is no longer present and then appears present again, a condition graphically represented with a "V" in the graph. For the years under review, approximately 5% of the municipalities exhibit behavior of this type. In this case, it is necessary to understand whether, due to an exogenous event, the library was temporarily closed or whether this apparent inconsistency hides an error in the detection of the basic data.

To understand whether this is an acceptable phenomenon, given that it affects a relatively low percentage of municipalities, the same sparkline analysis was conducted at a regional level. From the regional analysis, there are no inconsistencies in all regions except in Sardinia, where 40% of municipalities lacked the signal in 2018 (Table 9). This certainly "dirty" data is affected by incorrect recording of the information. Data of this type, which cannot be justified by the presence of exogenous events, is probably due to an inconsistency in data input.

Table 9 – *Sparklines of Sardinia libraries.*

2016	2017	2018	2019	2020	2021	Freq	%Freq Italy	%Freq Area	
0	**0**	**0**	**0**	**0**	**0**	6	0,08	1,6	
0	**0**	**1**	**1**	**1**	**1**	1	0,01	0,3	
0	**1**	**1**	**1**	**1**	**1**	1	0,01	0,3	
1	**0**	**0**	**1**	**1**	**1**	1	0,01	0,3	
1	**1**	**0**	**1**	**1**	**1**	**153**	**1,94**	**40,6**	
1	**1**	**1**	**1**	**1**	**0**	3	0,04	0,8	
1	**1**	**1**	**1**	**1**	**1**	212	2,68	56,2	



Source: Authors elaboration based on ISTAT data.

This inconsistency did not make it possible to disseminate the indicator for the entire reference period. While waiting for a timely check on the data, in the March release, it was decided to release the indicator for all municipalities but only for the period 2019-2021. The same analysis technique, simple and easy to interpret, was used for other qualitative indicators: for example, the number of anti-violence centres in a municipality. Also in this case, the method was particularly effective in identifying anomalous situations of signal absence in particular years.

4. Conclusion

For a decade, official statistics have been committed, both at a national and international level, to producing data that is as exhaustive as possible from a territorial point of view and timely. Register data and open data have certainly contributed to increasing the richness of information. Nonetheless, the dissemination of data for specific territorial domains presupposes important analytical work as it can bring out anomalous values that would not be detected at a macro-area level. A fundamental prerequisite for the dissemination of municipal or sub-municipal indicators is the robustness analysis of the indicators. Depending on the type of indicator chosen, it is advisable to provide an analysis, whether transversal, longitudinal or which aims to analyse the presence or absence of a signal.

In the next years, the even more massive use of the integration between statistical and administrative data will lead on the one hand to the dissemination of more data even at sub-municipal level and on the other to the possibility of testing the available data more effectively. This will be combined with the use of new sources deriving from the web and usable through scraping techniques and the current use of big data.

The public decision-maker now needs data that allows him to comprehensively describe his territorial reality and needs indicators for monitoring public actions. This is not only a need for larger municipalities but also for small and very small municipalities, distant from essential services which therefore require targeted cohesion policies. The new frontier of official statistics is undoubtedly the integrated use of statistical, administrative and big data which constitute a very important information basin especially for municipal and sub-municipal analyses. Official statistics is investing heavily in this direction, through numerous research projects, aimed at identifying the big data sources useful for official statistics, at transforming big data into statistical data by verifying their quality through robustness analysis techniques possibly also based on machine learning. Official statistics are following this path, and research must set up increasingly effective robustness techniques.

References

BIANCHINO A., CARUCCI A.M.M., VANNELLA G. 2018. Il nuovo sistema statistico "a misura di comune": un'applicazione sul territorio lucano, *Metodi e analisi statistiche*, Università degli studi di Bari Aldo Moro, pp. 83-108.

CRESCENZI F., LIPIZZI F. 2020. The integration of geographic and territorial data sources into the base register of territorial and geographical entities, *Statistical Journal of the IAOS*, Vol. 36, No.1, pp. 143-149.

EUROSTAT. 2010. Business Register. Recommendations Manual (2010 Edition). Methodologies and Working papers. Luxembourg.

ISTAT. 2012. *Quality Guidelines for Statistical Processes*, Versione 1.1, Dicembre 2012 – Istat –Roma Quality guidelines_English version_rev (istat.it)

ISTAT. 2018. *Linee Guida per la Qualità delle Statistiche del Sistema Statistico Nazionale*, Versione 1.0, Marzo 2018 – Istat –Roma

ISTAT. 2024. *Rapporto Annuale. La situazione del Paese*. Maggio 2024 –Roma

---

Agata Maria Madia CARUCCI, ISTAT, carucci@istat.it
Valeria MARZOCCA, ISTAT, valeria.marzocca@istat.it
Roberto Antonello PALUMBO, ISTAT, palumbo@istat.it

DOES MEASUREMENT ERROR BIAS EMPLOYMENT PATHWAYS? THE CASE OF ITALY[1]

Dimitris Pavlopoulos, Roberta Varriale, Silvia Loriga

Abstract. The exploration of employment trajectories over time may be significantly biased due to measurement errors in the data used for the analysis. This paper addresses this issue by employing a mixture hidden Markov model (MHMM) that detects and corrects for measurement errors. Specifically, we use an MHMM that includes two indicators for employment status, derived from linked data from the Italian Labour Force Survey and Administrative Data for the period 2017-2021.

1. Introduction

In recent years, the processual approach has been gaining momentum in labour market research (Abbott, 1983). Instead of 'merely' studying single transitions, a process-based method allows us to examine entire sequences and has the advantage of being able to provide a complete and detailed picture of entire trajectories, e.g. in the field of working careers (Aisenbrey and Fasang, 2017; Mattijssen and Pavlopoulos, 2019; Studer and Ritschard, 2016). A potential threat to this research is measurement error that exists in socio-economic data. In data coming from surveys, measurement error may be caused by cognitive processes, social desirability, design and implementation of the survey (Groves, 2004; Sudman *et al.*, 1996; Tourangeau *et al.*, 2000). In administrative data, measurement error can result from the misalignment between the definition of the administrative variable and the target one, administrative delays, wrong registration, or erroneous administrative procedures (Bakker and Daas, 2012; Oberski *et al.*, 2017). Research has shown that the measurement error can have severe effects on our estimates of mobility over time. Even a small magnitude of measurement error can bias considerably transition rates from unemployment to employment (Filipponi *et al.*, 2021) or from temporary employment to permanent employment (Pavlopoulos and Vermunt, 2015). The use of hidden Markov models (HMMs) is a promising approach for detecting and correcting measurement error in categorical longitudinal data (Biemer, 2011). The advantages of HMMs

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

are that they are very flexible and do not assume the existence of a primary (error-free) source of information among those available (Filipponi et al., 2021). HMMs treat the 'true' employment status as a latent variable that is imperfectly measured by one or more observed indicators. Using linked data from different sources allows modelling realistic error specifications with an HMM. In more detail, when multiple data sources are available, HMMs can model both random and systematic measurement error. An extension of HMM is the mixture hidden Markov model (MHMM) that estimates the true employment status trajectories (formally: mixtures) (Vermunt, 2010). In this paper, we investigate the effect of measurement error on employment trajectories in Italy for the period 2017-2021. Furthermore, we aim to produce an error-corrected estimate of these trajectories. For this purpose, we apply a MHMM with two independent indicators for employment status. These indicators come respectively from the Italian Labour Force Survey (LFS) and Administrative Data (AD) managed by the Italian National Statistical Institute (Istat).

The remaining of this paper is organized as follows. In section 2, we describe the data that are used in the analysis. In section 3, we build the MHMM that is employed, and in section 4, we discuss the conclusions of our findings as well as the steps for further research.

2. The data

The Italian Labour Force Survey (LFS) follows the standards set by EU Regulation 2019/1700 of the European Parliament and the Council. LFS's primary goal is to provide employment statistics. LFS is a continuous household survey, conducted throughout the year and covers approximately 1.2% of the entire Italian population. The sample is selected according to a two-stage stratified sampling design: municipalities represent the primary statistical units and are stratified based on their population size; the final statistical units are the households, which are selected through a simple random sampling. The Italian LFS adopts a quarterly rotation scheme. Selected households are interviewed four times within a 15-months period. Each household is interviewed for two consecutive quarters, followed by a two-quarter break and then two consecutive survey quarters. Interviews are spread across all weeks of the quarter (reference weeks), maintaining the sample representativeness on monthly basis. The information on labour market participation refers to the reference week. For further details on the LFS contents, methodologies and organization see (Istat, 2006). Italian Administrative Data (AD) relevant to labour statistics are collected by the social security and tax authorities and are gathered and appropriately processed by Istat. These data go through different, source-specific editing and harmonization procedures (Baldi *et al.*, 2018).

Our sample in this study refers to the LFS cohorts that entered the survey from January 2017 to December 2021. From the LFS, information from all available survey waves in which these individuals participated is retained. The actual number of LFS observations in the data may be four or less than four in case of attrition or whether the LFS rotation scheme started before 2017 or ended after 2021. For the same set of individuals, quarterly information from the AD is retained, covering all quarters from January 2017 to December 2021. For each individual, there is a maximum of four observations from the LFS, whereas AD information are potentially available for all the twenty observations. The linkage between LFS and AD was conducted at the individual level through a pseudonymized code. We include, in our sample, individuals aged 25 to 55 who participated at least once in the LFS within this period. As our statistical model is computationally demanding, a 10% sample of units was randomly selected, stratifying the sample by the month of the first LFS interview. This procedure resulted in a sample of 39,847 individuals.

Our target variable on employment status is an aggregation of the International Classification of Status in Employment (ICSE-18), established by the International Labour Office and takes 4 values: (1) employees with a permanent contract (PE), (2) employees with a temporary contract (TE), (3) self-employed (SE), which encompasses employers, independent workers without employees, contributing family workers and dependent contractors and (4) not employed (NE). The creation of such target variable according to this simplified classification is relatively straightforward for LFS data, given that ICSE criteria inspired the definitions of LFS variables. However, in the AD, such a derivation presented several challenges as it uses different sources. In this respect, the administrative classification may not fully align with the statistical classification implemented in the LFS, as it is based on administrative concepts; on the other hand, LFS variables, based on self-reported information, may be affected by errors due to the subjective interpretation of the question by the respondent (especially in case of proxy answer). A potential source of discrepancies between LFS and AD is the different coverage: LFS aims at covering all employed population, including informal work, while AD contains information only for formal employment. Additional conceptual discrepancies between LFS and AD can be attributed to shortcomings in the data collection process. These include e.g. temporal misalignment of sources, particularly for occasional employment and discrepancies in the definitions of work signals across available sources. Tables 1 and 2 display the transition rates between the different employment categories in adjacent quarters in the LFS and AD. Despite the abovementioned sources of discrepancies, the disparities between the datasets are relatively minor.

Table 1 – *Observed transitions in the Labour Force Survey. Years 2017-2021.*

	Employment status *t*			
Employment status *t-1*	PE	TE	SE	NE
Permanent contract	0.962	0.012	0.006	0.020
Temporary contract	0.074	0.739	0.013	0.174
Self-employed	0.015	0.010	0.944	0.031
Not employed	0.019	0.063	0.013	0.906

Source: Istat, Labour Force Survey

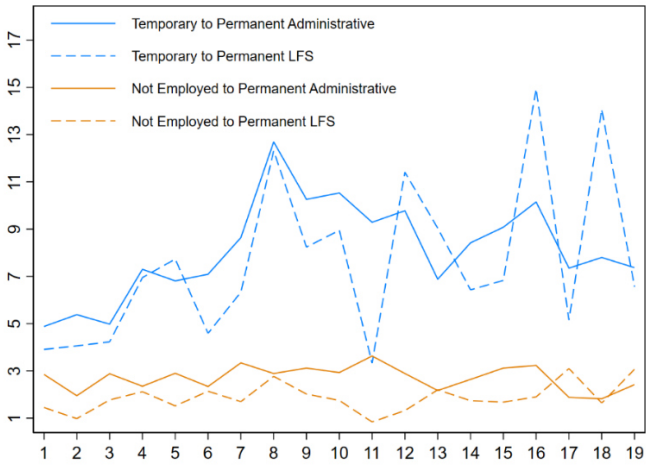
Table 2 – *Observed transitions in the Administrative Data. Years 2017-2021.*

	Employment status *t*			
Employment status *t-1*	PE	TE	SE	NE
Permanent contract	0.966	0.009	0.003	0.022
Temporary contract	0.078	0.717	0.017	0.187
Self-employed	0.011	0.012	0.956	0.022
Not employed	0.030	0.061	0.012	0.897

Source: Istat, Administrative Data

Figure 1 shows the observed transition rates between different employment categories across adjacent quarters from 2017 to 2021. The graph confirms the findings of Tables 1 and 2 that there are only minor differences in flow patterns between the LFS and AD. The largest transition rates occur from temporary to permanent employment, showing an increasing trend. This pattern is evident in both LFS and AD, suggesting that there is time dependence in transition probabilities.

Figure 1 – *Transition flows from type of contracts by quarter, years 2017-2021.*



Sources: Istat, Labour Force Survey and Administrative Data

Table 3 presents the cross-classification of employment status from LFS and AD data. The diagonal cells concern cases where the two data sources agree on the classification. In contrast, off-diagonal values represent discrepancies in classification and indicate potential classification errors in at least one of the data sources. As Table 3 illustrates, the two data sources do not align for approximately 14.6% of the total number of cases. Beyond random classification errors, these discrepancies arise from distinct reasons, as suggested by Varriale and Alfó (2023) in their analysis of employment status. Errors in AD are typically attributable to miss-specifications of statistical concepts. For example, AD lack information on irregular work, or it may encounter difficulties in correctly identifying the reference period of the information. On the other hand, errors in the LFS survey may arise from miss-classification due to respondents providing incorrect answers or having an erroneous understanding of employment categories.

Table 3 – *Cross-classification of employment status, Administrative Data and Labour Force Survey, frequencies and percentages in brackets. Years 2017-2021.*

	Employment status, LFS				
Empl. status AD	PE	TE	SE	NE	Total
Perm. contract	41326 (41.1)	1993 (2.0)	1203 (1.2)	1290 (1.3)	45812 (45.6)
Temp. contr.	1217 (1.2)	5442 (5.4)	298 (0.3)	1268 (1.3)	8225 (8.2)
Self-employed	748 (0.7)	349 (0.3)	11033 (11.0)	1095 (1.1)	13225 (13.2)
Not employed	1942 (1.9)	1394 (1.4)	1876 (1.9)	28066 (27.9)	33278 (33.1)
Total	45233 (45.0)	9178 (9.1)	14410 (14.3)	31719 (31.5)	100540 (100)

Sources: Istat, Labour Force Survey and Administrative Data

Table 4 presents the same cross classification as Table 3 but reports the percentage distribution of employment status as measured by the LFS, conditional on the AD measurement. 90.2% of cases that are recorded as permanent contracts in AD are also classified as permanent contracts according to the LFS.

Table 4 – *Distribution of employment status from Labour Force Survey conditional on Administrative Data measurement. Years 2017-2021.*

	Employment category, LFS				
Empl. status AD	PE	TE	SE	NE	Total
Perm. contract	90.2	4.4	2.6	2.8	100
Temp. contract	14.8	66.2	3.6	15.4	100
Self-empl.	5.7	2.6	83.4	8.3	100
Not empl.	5.8	4.2	5.6	84.3	100
Total	45.0	9.1	14.3	31.5	100

Sources: Istat, Labour Force Survey and Administrative Data

The relevant percentages of classification agreement for the self-employed and not employed are also quite high (83.4% and 84.3%, respectively). In these categories, the only case with a high percentage of misclassification concerns the cases that are recorded as self-employed in AD as 8.3% of them are classified as not employed in the LFS. Major classification mismatches are only observed for temporary contracts. In fact, only 66.2% of those recorded as having a temporary contract in AD are observed as having a temporary contract in the LFS, while about 14.8% are classified as having a permanent contract and 15.4% as not employed. These findings are rather stable over time.

3. The mixture hidden Markov model

Let A_{it} and S_{it} measure the observed indicators of the employment status of individual i at time t according to AD and LFS, respectively. We use quarterly data from 2017-2021, meaning that t runs from 0 to T = 19, while N equals 66,127. The true (latent) employment status is indicated by X_{it}. All A_{it}, S_{it}, and X_{it} can take on four values representing the categories of employment status (permanent contract, temporary contract, self-employment and non-employment); a particular category of the observed indicators is denoted by a_t, s_t, and the latent variable x_t.

In the MHMM, the joint probability of following a particular observed state path conditional on covariates V, t and t^2 is expressed as:

$$P(A_i = a_i, S_i = s_i | V_i, t) = \sum_{k=1}^{K} \sum_{x_0=1}^{4} \sum_{x_1=1}^{4} \dots \sum_{x_T=1}^{4} P(w_i = k) P(X_{i0} = x_0 | w_i = k)$$

$$\prod_{t=1}^{T} P\left(X_{it} = x_t | X_{i(t-1)} = x_{t-1}, w_i = k, t, t^2\right)$$

$$P(A_{i0} = c_0 | X_{i0} = x_0)$$

$$\prod_{t=1}^{T} P\left(A_{it} = a_t | X_{it} = x_t, X_{i(t-1)} = x_{t-1}, A_{i(t-1)} = a_{t-1}\right)$$

$$\prod_{t=1}^{T} P(S_{it} = s_t | X_{it} = x_t, V_{it})^{\delta_{it}} \quad (1)$$

The latent employment status X_{it} follows a first-order Markov process. This means that the realization of the true employment status at time point t, X_{it}, is

independent of all previous realizations of it conditional on its realization in the previous time point ($t - 1$).

The MHMM estimates four sets of probabilities: the initial state probabilities $P(X_{i0} = x_0)$, the latent transition probabilities $P(X_{it} = x_t|X_{i(t-1)} = x_{t-1})$ as well as the measurement error probabilities for the administrative data $P(A_{it} = a_t \vee X_{it} = x_t)$, and for the LFS $P(S_{it} = s_t|X_{it} = x_t)$. To model more realistic scenarios, the latent transition probabilities are assumed to be time-heterogeneous. In more detail, these probabilities depend on a quadratic specification of the quarter index (i.e. t and t^2), $P(X_{it} = x_t|X_{i(t-1)} = x_{t-1}, t, t^2)$). Finally, we estimate the probability of belonging to one of the K mixtures. Mixtures represent the employment trajectories as they are time-constant latent classes of individuals with similar initial latent state probabilities ($P(X_{i0} = x_0 \vee w_i = k)$) and latent transition probabilities ($P(X_{it} = x_t|X_{i(t-1)} = x_{t-1}, w_i = k, t, t^2)$).

The standard HMM maintains the Independent Classification Error assumption. This assumption means that the realizations of the observed states A_{it} (respectively, S_{it}) are independent of its previous and future realizations conditional on the value of the latent state X_{it}. However, this assumption is rather unrealistic as both survey and administrative employment data contain systematic measurement error (Pankowska *et al.*, 2021). To address this issue, we relax it by modelling error autocorrelation in both the administrative data and LFS indicator. In more detail, the measurement-error probabilities in LFS and administrative data are allowed to depend on the lagged observed and lagged latent employment status. As $X_{i(t-1)}$, $A_{i(t-1)}$ and $S_{i(t-1)}$ can take on 4 values, there are 16 ($4 * 4$) different sets of measurement-error probabilities in the administrative data and other 16 in the LFS. Each of these probabilities corresponds to each possible combination of lagged observed and latent employment status. To estimate a model with a meaningful number of probabilities as outcomes, we applied certain restrictions. For this purpose, we defined a constrained logit model that estimates one additional parameter when the same error can be made between adjacent time points. Similar restrictions have been previously applied in HMMs by others (Manzoni *et al.*, 2010; Pankowska *et al.*, 2021; Pavlopoulos and Vermunt, 2015).

The model is estimated with Maximum likelihood, while parameters are obtained using the forward-backward or Baum-Welch algorithm (Baum *et al.*, 1970), which is a variant of the Expectation-Maximization algorithm. Specifically, we use an extension of the Baum-Welch algorithm for MHMMs with covariates as described by Vermunt (2007). The estimation is done with the program Latent GOLD, version 6.0 (Vermunt and Magidson, 2016).

4. Results

Model selection involved three stages. The first stage aimed at selecting the model with measurement-error specifications that best fit the data. In the second stage, we compared model specifications with a different number of mixtures (i.e. trajectories). To accommodate the computational intensity of the MHMMs, in this, we follow the two-step approach that was introduced by Bakk and Kuha (2018). This approach involves estimating models with a different number of mixtures where the measurement-error parameters have been fixed to the values estimated in the first step of the estimation process. The third and final stage involved testing whether including predictors to the MHMM can improve model fit.

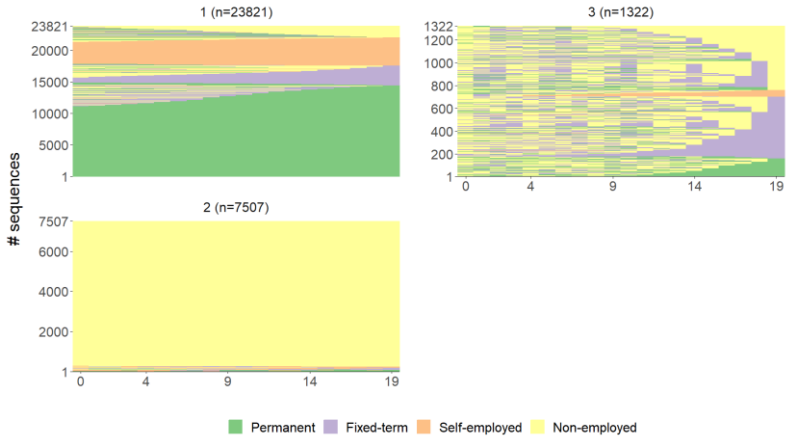
In total, in the first stage, we estimated 9 models[2]. Models 1-3 correct only for random measurement error. In more detail, Model 1 corrects for random measurement error in the indicator of the LFS, Model 2 in the indicator of the AD, and Model 3 in both indicators. Models 4-9 also add corrections for systematic measurement error. In more detail, Models 4-6 include an additional error coefficient for cases where an error was made in time point $t-1$ and can be repeated in t in the LFS indicator (Model 4), in the AD indicator (Model 5) or in both indicators (Model 6). Models 7-9 have a less restrictive specification as they always estimate an additional error coefficient for the cases where an error was made in time point $t-1$ in the LFS indicator (Model 7), the AD indicator (Model 8) or in both indicators (Model 9). Model fit measures (BIC and AIC) show that Model 5 performs best. In the next stage, we build on the selected model (Model 5) to find the optimal number of trajectories (mixtures). The two-step approach of Bakk and Kuha (2018) involves fixing the measurement model parameters to the values that we estimated by Model 5 and estimate models with an increasing number of mixtures, from 2 to 6. Model selection is, in this case, more complicated. As mixtures represent groups of individual trajectories, statistical criteria cannot provide alone the optimal solution. These statistical criteria (i.e. the AIC and the BIC) are combined with substantial considerations which refer to the size and the interpretation of trajectories. This aspect is somewhat described by the concepts of latent class separation and heterogeneity which are summarized by the measure of entropy. To sum up, besides looking at model fit measures, we select a solution with mixtures of substantial size, adequately separated with each other and having a meaningful interpretation. On the basis of these considerations, we selected an MHMM with 3 mixtures. This model has an entropy of 0.694 which, although not being particularly high, is acceptable. In the last stage, we also studied whether adding covariates (other than the indicator for the quarter) improved the model. This appeared not to be the case, so we

[2] We do not report the results of model fit indices for lack of space. These are available on request.

proceeded with the interpretation of the MHMM with 3 mixtures and only the quarter as a predictor of the latent employment state.

Figure 2 illustrates the index plot of the employment trajectories that were obtained by the Model including 3 mixtures. In this plot, every "line" represents an individual pathway across time. The status of each individual at each time point t represents actually the modal classification state according to the selected MHMM. In each of the trajectory plots, the x-axis runs from 0 to T = 19. The largest trajectory (trajectory 1 in Figure 3, which represents 75.4% of the sample) is a trajectory where employment is dominant. People in this category are most likely to be employed on a permanent contract, (less often) with a temporary contract or are self-employed. Transitions in this trajectory are rare and lead mostly to one of the three employment states. Trajectory type 2 (19.8% of the sample) is the non-employment states. Almost all pathways in this trajectory type consist exclusively of non-employment. Lastly, the smallest trajectory type (trajectory 3, 4.8% of the sample) is the trajectory with unstable and rather precarious pathways. This trajectory includes pathways with many transitions between temporary employment and non-employment.

Figure 2 – *Employment trajectories with measurement-error correction.*

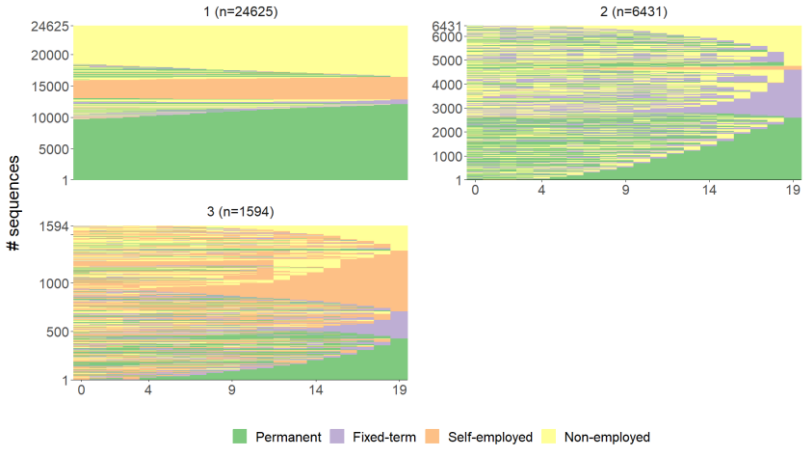


Sources: Istat, Labour Force Survey and Administrative Data

To evaluate whether measurement error biases the trajectories that we find in this analysis, we estimate the same model but this time we assume that the observed indicators are error free. In other words, we estimate a mixture Markov model with 3 mixtures. This solution has an entropy of 0.732. We present the 3 trajectories that were produced by this model in Figure 3. The size of the 3 trajectories is similar to those with measurement error correction, 73.2%, 21.2% and 5.6%. However, what these trajectories represent differs considerably from trajectories that were extracted

by the model that includes measurement error correction. Failing to correct for measurement error leads to the estimation of more heterogeneous trajectories that include many more transitions than what really exists. In more detail, the largest trajectory includes mostly stable pathways of individuals in permanent employment, self-employment and non-employment. The second largest trajectory includes unstable pathways where individuals move often between temporary employment, permanent employment and non-employment. The smallest trajectory groups also unstable pathways but contrary to the second trajectory, these pathways often include self-employment as well.

Figure 3 – *Employment trajectories without measurement-error correction.*



Sources: Istat, Labour Force Survey and Administrative Data

5. Conclusion

In the article, we estimated a mixture hidden Markov model (MHMM) to study the effect of measurement error on the typology of employment trajectories in Italy. For this purpose, we used linked data from administrative sources and the Labour Force Survey for the years 2017-2021. Our MHMM corrected for both random and systematic measurement error in both data sources.

Our results indicate that measurement error has a severe biasing effect on employment trajectories. Although the size of the estimated mixtures is not severely biased, the typologies coming out of the two analyses (i.e. with and without measurement error correction by employing an MHMM instead of a mixture Markov model) are substantially different. In this respect, research applying a processual approach to study social phenomena should not be ignorant of measurement error.

Applying error-correction is necessary to reveal existing typologies of employment or other social processes.

Acknowledgements

This paper is part of the project DYNANSE which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 864471).

References

ABBOTT A. 1983. Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, Vol. 16, No. 4, pp. 129-147.

AISENBREY S., FASANG A.E. 2017. The Interplay of Work and Family Trajectories over the Life Course: Germany and the United States in Comparison, *American Journal of Sociology*, Vol. 122, No. 5, pp. 1448-1484.

BAKK Z., KUHA J. 2018. Two-Step Estimation of Models Between Latent Classes and External Variables, *Psychometrika*, Vol. 83, No. 4, pp. 871-892.

BAKKER B.F.M., DAAS P.J.H. 2012. Methodological challenges of register-based research, *Statistica Neerlandica*, Vol. 66, No. 1, pp. 2-7.

BALDI C., CECCARELLI C., GIGANTE S., PACINI S., ROSSETTI F. 2018. The Labour Register in Italy: The New Heart of the System of Labour Statistics, *Rivista Italiana Di Economia Demografia e Statistica*, Vol. LXXII, No. 2, pp. 95-105.

BAUM L.E., PETRIE T., SOULES G., WEISS N. 1970. A Maximization Technique Occuring in the Statistical Analysis of Probabilistic Functions of Markov Chains, *The Annals of Mathematical Statistics*, Vol. 41, No. 1, pp. 164-171.

BIEMER, P. 2011. *Latent Class Analysis of Survey Error*. New Jersey: John Wiley Sons.

FILIPPONI D., GUARNERA U., VARRIALE R. 2021. Latent Mixed Markov Models for the Production of Population Census Data on Employment. In PERNA C., SALVATI N., SCHIRIPPA S. (Eds.) *Book of short papers SIS 2021*, Pearson, pp. 112-117.

GROVES R.M. 2004. Survey errors and survey costs. New Jersey: Wiley.

ISTAT 2006. La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione, *Istat, Metodi e Norme*, Vol. 32, pp. 173-196.

MANZONI A., VERMUNT J. K., LUIJKX R., MUFFELS R. 2010. Memory Bias in Retrospectively Collected Employment Careers: A Model-Based Approach to

Correct for Measurement Error, *Sociological Methodology*, Vol. 40, No. 1, pp. 39-73.

MATTIJSSEN L., PAVLOPOULOS D. 2019. A Multichannel Typology of Non-standard Employment Careers, *Social Science Research*, Vol. 77, pp. 101-114.

OBERSKI D.L., KIRCHNER A., ECKMAN S., KREUTER F. 2017. Evaluating the Quality of Survey and Administrative Data with Generalized Multitrait-Multimethod Models, *Journal of the American Statistical Association*, Vol. 520, pp. 1477-1489.

PANKOWSKA P., BAKKER B., OBERSKI D., PAVLOPOULOS, D. 2021. Dependent interviewing: A remedy or a curse for measurement error in surveys?, *Survey Research Methods*, Vol. 15, No. 2, pp. 135-146.

PAVLOPOULOS D., VERMUNT J.K. 2015. Measuring Temporary Employment. Do Survey or Register Data Tell the Truth?, *Survey Methodology*, Vol. 41. No. 1, pp. 197-214.

STUDER M., RITSCHARD G. 2016. What matters in differences between life trajectories: A comparative review of sequence dissimilarity measures, *Journal of the Royal Statistical Society. Series A: Statistics in Society*, Vol. 179, No. 2, pp. 481-511.

SUDMAN S., BRADBURN N.M., SCHWARZ N. 1995. Thinking about answers: The application of cognitive processes to survey methodology, *Quality of Life Research*, Vol. 12, pp. 719-720.

TOURANGEAU R., RIPS L.J., RASINSKI K.A. 2000. The psychology of survey response. Cambridge: University Press.

VARRIALE R., ALFO' M. 2023. Multi-source statistics on employment status in Italy, a machine learning approach, *METRON*, Vol. 81, No. 1, pp. 37-63.

VERMUNT J.K. 2007. Growth models for categorical response variables: standard, latent-class, and hybrid approaches. In VAN MONTFORT K., OUD J., SATORRA A. (Eds.) *Longitudinal models in the behavioral and related science*, Erlbaum, pp. 139-158.

VERMUNT J.K. 2010. Longitudinal Research Using Mixture Models. In MONTFORT K., OUD J.H., SATORRA A. (Eds.), *Longitudinal Research with Latent Variables*, Springer, pp. 119-152.

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

Dimitris PAVLOUPOLOS, Vrije Universiteit Amsterdam, d.pavlopoulos@vu.nl
Roberta VARRIALE, Sapienza University of Rome, roberta.varriale@uniroma1.it
Silvia LORIGA, Italian National Institute of Statistics, siloriga@istat.it

POSSIBLE IMPROVEMENTS IN THE ESTIMATION OF IT-LFS THROUGH INTEGRATION WITH THE REGISTERS SYSTEM

Antonella Iorio, Alessandro Martini

Abstract. In recent years, due to the increasing difficulties in maintaining high response rates and ensuring good representativeness of the collected samples, quality problems in surveys are getting more frequent. The impact on the final estimates is in terms of bias and lack of consistency between the different estimates. In recent decades, many NSIs have started to use administrative data, usually processed and organized in statistical registers, integrating them into statistical processes, with the aim of improving the quality of the output.
Availability of data coming from registers is increasing in the last years in Italy and it will improve even more in the next future.
The Registers should be exploited in statistical processes and in different phases: as frames for sample selections, to draw auxiliary variables needed in the estimation and to enrich the set of available information, through record linkage.
In this paper we focus on the evaluation of possible improvements on the accuracy in the IT-LFS exploiting the integration of auxiliary information coming from the ISTAT system of statistical registers. The Registers System (SIM) will be able to provide information about sex, age, citizenship, education level, employment signals and household structure for the resident population.
This more reliable information could be used through calibration in both the phases of estimation: treatment of total no response and final calibration.
For the no response treatment it would be possible to define constraints in a more flexible framework while, in the final calibration step, additional information about education level and employment signals coming from the register could be useful to improve different aspects of quality of the LFS estimates.
This experimental study is relevant to evaluate the possible improvements of the estimates and to collect evidences for introducing these changes in the next future.
The results show that with the proposed innovations the survey problems (MRT, bias) are addressed by improving the quality of the estimates.

1. Background

The Italian LFS is one of the most important source of statistical data referred to the Italian labour market. It provides monthly, quarterly and yearly figures of the main aggregates and, in a longitudinal perspective, flow estimates as well. The IT-

LFS is designed as a quarterly survey, the sample design is a two stage stratified sample with rotational pattern 2-2-2. The sample is uniformly spread across all weeks such that all geographical domains are represented in each month and in each of the four waves (rotational groups).

The estimation methodology in IT-LFS is design based, demographic administrative data are exploited through the calibration estimator.

Consistency of estimates produced by different sources is a key quality dimension for NSIs, so methodological approach and estimation procedures in IT-LFS take into account the need to provide consistent estimates for the different indicators in the statistical domain of labour market. More specifically, ISTAT is making an effort to improve coherency between LFS and different other surveys (Multipurpose Annual Survey, Time Use Survey, IT-Silc) and also with Continuous Population Census. "Coherence by design" is the idea of ensuring consistency trough the harmonization of sampling frames, designs, estimation methods and data collection protocols.

In this paper we focus on the evaluation of possible improvements on the accuracy of the estimates produced by IT-LFS exploiting the integration of auxiliary information coming from the ISTAT system of administrative and statistical registers (SIM).

The procedure is based on three step, and final weights are computed as follows:

- Step 1: Initial weights are obtained as the inverse of the inclusion probabilities of any selected household.
- Step 2: Intermediate weights are computed multiplying initial weights by correction factors for unit non-response worked out as the inverse of the response ratios; correction factors for non-response treatment are calculated using the information from the LFS theoretical samples to define 13 household typologies, considering the age, gender, citizenship of the head of the household and the number of members. When we started this kind of correction for no response, this information was not available in the population frame, but only for the municipalities selected in the sample.
- Step 3: Starting from intermediate weights, final weights are obtained solving a minimization problem under constraints through a calibration model. The estimates of some auxiliary variables have to be equal to the totals in the reference population. Final weights ensure that all members of a given household have the same weight.

Grossing weights are computed on quarterly basis, whereas annual estimates are calculated as averages of quarterly estimates.

The IT-LFS weighting procedure is more complicated than required by EU regulations, to satisfy specific national needs, in particular the weighting scheme includes, at NUTSII level:

- distribution of population by sex and seventeen 5-year age groups;

- distribution of non-national population divided as follow: male, females, other EU citizens, Non-EU citizens;
- number of households for each rotation group (1/4 of the total);
- distribution of population by sex, for each of the three months of the quarter.

The weighting scheme includes too:

- distribution of population by sex and five age groups for the thirteen large municipalities with more than 250.000 inhabitants;
- distribution of population by sex and five age groups at NUTSIII level.

According to European Regulations these control totals applied in calibration derive from demographic administrative sources. This auxiliary information is incorporated in the estimation procedure with the calibration estimator (Deville & Särndal, 1992; Särndal, 2007) to improve the quality of the estimates. As matter of fact that quarterly weights already include the benchmark to the NUTS 3 population, the annual datasets are obtained using all the quarterly interviews and annual weights are computed simply dividing the quarterly weights by four.

2. Calibration on control totals

Deville and Särndal (1992) formalized calibration in survey sampling according to the basic idea that if auxiliary variables strongly correlated with the target variables are available it "means that the weights that perform well for the auxiliary variables also should perform well for the study variables".

Consider a finite population $U = \{1, . , k, , N\}$. A sample s of fixed size n is drawn from population U according to a sampling design $(S, p(.))$, where S is the sample space and $p(\cdot)$ is a probability distribution on S. The first order inclusion probability, $\pi_k = Pr(k \in s)$, and the second order inclusion probability, $\pi_{kl} = Pr(k, l \in s)$, are assumed to be known and strictly positive. Throughout, let $\pi_{kk} = \pi_k$.

Let us assume to be interested in estimating the total of Y variable $t_y = \sum_{k \in U} y_k$ and to have collected on each element $k \in s$, besides the value of the interest variables y_k, an auxiliary vector value, $x_k = (x_{k1} \dots x_{kp} \dots x_{kP})$. Furthermore, population totals of Xs, $t_{x_p} = \sum_{k \in U} x_{k_p}$ with $p = 1, \dots, P$, are accurately known, without sampling error. Solving the optimization problem:

$$\begin{cases} min_{w_k} \{\sum_{k \in s} G(w_k, d_k) \,/\, q_k\} \\ \sum_{k \in s} w_k x_k = t_x \end{cases}$$

Deville and Särndal demonstrate that we can find a system of weights, w_k, calibrated (coherent) with the known population totals and the more the auxiliary variables are correlated with the target variables, the more efficiency of the estimates improves.

Furthermore, the calibrated weights are as close as possible to the design weights, $d_k = 1/\pi_k$, with respect to a given metric $G(\cdot)$.

When $G(\cdot)$ is the chi-squared distance, the resulting calibration estimator is equal to the generalized regression (GREG) estimator (see Cassel et al., 1979; Särndal, 1980; Isaki and Fuller, 1982; Wright, 1983; Bethlehem and Keller, 1987; Särndal et al., 1989; Fuller, 2002).

Even using other metrics with respect to the chi-squared distance (see Deville and Särndal, 1992 pp. 378-379) the calibrated estimator (CAL) is asymptotically equivalent to the GREG estimator. Therefore, also the asymptotic variance of CAL estimator can be derived referring to the GREG estimator.

3. Micro-linkage of administrative information and LFS data

Availability of data coming from administrative and statistical registers has increased significantly in recent years in Italy, and its quality and timeliness is expected to improve further over the next few years. These new sources open up the possibility of integrating such kind of data into some of the LFS production processes, usually as auxiliary information – used at the micro or macro level - for sample design, sample selection, correction for non-response, weighting, validation, imputation and modelled estimates. When administrative data are integrated into the processes at the macro level, usually there are no particular issues, apart from the consistency of definitions and the timeliness, i.e. the lag between their reference period and the moment they are available to LFS specialists for processing.

When administrative data are integrated into the processes at the micro level, a further problem arises and is related to the possibility for LFS specialists (or other dedicated specialists within statistical offices) to have an anonymized unique identifier (SIM code) that allow to link individual records (persons and households) of the administrative/statistical databases to those of the LFS databases (persons and households selected, whether interviewed or not). The process of assigning the anonymized unique identifier is called in Istat “pseudonymization”. It is a crucial task performed by a specialized unit within Istat, following stringent criteria and procedures set by The Italian Data Protection Authority.

This “pseudonymization” process has to be carried out regularly, both on the new administrative databases made available to Istat and on the actual LFS samples. This process will assign the unique identifiers to persons and households that were already

"pseudonymized" and will create a new unique identifiers to persons and households that enter for the first time into the population or sample (e.g. new born, immigrants). Clearly, the integration at the micro level of up-to-date administrative data and LFS microdata during LFS processing would help to implement much more effective methods for the treatment of total non-response and under-coverage, thus improve the quality of monthly and quarterly LFS estimates that need to be disseminated at t + 30 days (for monthly data) and t + 56 days (for quarterly estimates and for transmitting validated data to Eurostat).

Until a few months ago, the main obstacle to such integration was represented by the timing and methods of "pseudonymization" of the actual sample of the LFS survey. While the unique identifier was available for the theoretical sample already during sample selection, given the complexity of the procedure it was used to be attached to all respondents of the actual LFS sample only on an annual basis, at the end of the year, for the purpose of estimating the informal employment for National Accounts.

In order to be able to regularly integrate administrative data for the treatment of total non-response in the LFS, a new "pseudonymization" process was developed - jointly by LFS specialists and the dedicated Istat unit - specifically for such purpose. The new process is semi-automated, based on a probabilistic record linkage methodology, and runs on daily bases on the new interviews transmitted by the enumerators. For individuals who are not linked through the probabilistic approach, a deterministic linkage procedure is used. It's important to note that this semi-automated daily process excludes visual checks, but these further quality checks are performed manually only at the end of the quarter to finalize the process.

This new process successfully assigns a unique code to over 98% of cases, thereby enabling the exploitation of register information for the estimation of both monthly and quarterly LFS data.

4. Use of calibration and administrative information for Non-response treatment

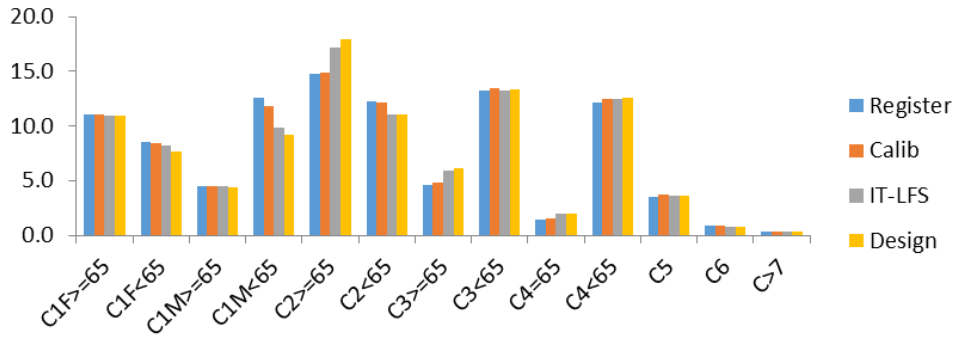
The availability to link the actual sample to the register data made it possible to review the non-response treatment. As seen above correction factors for non-response treatment are calculated with reference to the estimates obtained from LFS theoretical sample, through an Iterative proportional fitting procedure.

The availability of the population register allows to derive directly the totals for the distribution of the households by the specific typologies considered and geographical domains. Moreover, the use of calibration allows to define a more flexible framework to avoid very small adjustment cells, organizing constraints in a

more flexible framework (for instance a more detailed classification of households typologies at NUTS2 domain, a less detailed classifications at NUTS3 domain and strata level).

The results in Figure 1 show clearly how treatment of non-response trough calibration and administrative information is more effective in the context of the IT-LFS (blue and orange bars). The procedure we are currently applying is not able to entirely correct the overestimation, or underestimation, of certain categories of household, in particular the household of 2 members with the head of the household aged 65+ (C2>=65), we have an estimate of 17.9% with the base design weight (yellow bar) the intermediate weight of LFS provides an estimate of 17.2% (grey bar) while in the frame we observe 14.8% (blue bar). On the other side, households consisting of a single male member aged 18-64 (C1M<65) are 9.9% (grey bar), according to non-response adjusted weight, but they are 12.6% in the population register (blue bar). This is mainly due to the fact that the in the procedure we are currently using the correction factors, at stratum level, cannot exceed the correction factor calculated at province (Nuts3 level). This was a reasonable criterion and ensured a suitable correction when the procedure was defined but nowadays the profile of non-response by region and type of municipality has deeply changed and these bounds at province level do not allow an effective treatment of such different non-response patterns. In addition, the totals directly derived from the population register are more up-to-date, likely more reliable, and definitely more precise than the estimates we can derive from the theoretical sample.

Figure 1 – *Households by number of members and age of the head of the household, IT-LFS 2021Q4.*



Source: ISTAT, Labour Force Survey

5. Additional benchmarks derived from administrative and statistical registers

The Covid emergency had a big impact on the data collection process of IT-LFS. Since during lock-down all the interviews were conducted by CATI and the group of households that was entering the sample for the first time, was replaced selecting households that had already participated to the LFS in previous quarters and had provided a phone number for telephone interviews. So possible sources of bias arose in the data due to:

- selection of all households having phone number available;
- higher substitution rate (replacement of non-respondent households);
- higher final non-response rate.

The bias was studied comparing LFS data collected in 2020Q2 with previous quarters and with administrative sources. IT-LFS figures showed higher frequencies for elderly people, Italians, people having higher education level and employed ones. In those quarters were introduced in the calibration additional constraints regarding the distribution by education level at NUTS2 level in order to reduce the bias, maintaining the same structure of the population by education level. Updating these constraints was an issue since we had no data available for the definition of the totals.

On January 2021 new population figures were made available for the period 2011-2021, according to the results of the 2018 Population continuous Census. Consequently, LFS weights were recalculated to be coherent with IT Census population, starting from 2021Q1 and the constraints about education level have not been used anymore.

However, given the important relation noticed between the sampling distribution of the educational levels and the key labour market indicators, a further study has been conducted to evaluate the effect of the integration of auxiliary information coming from the registers on the IT-LFS estimates. This experimental study is relevant to evaluate the possible improvements of the estimates and to collect evidences for introducing these changes in the next future.

For this purpose, an additional set of constraints has been added to the final step of calibration, deriving them from the education level in the population register (RBI) and the signals of regular employment in the Labour Register (RTL/BOP). Both the registers are available with a variable timing compared to the production calendar of the LFS, a new population register (RBI) for year Y-1, referring to 31th december, is available for the estimation of quarter YQ4 while RTL/BOP has almost the same timing but with a provisional version. The same information is available for the actual sample, at microdata level, and the total population.

The first set of 24 constraints include education level in 4 groups (ISCED 0-1, 2, 3-4, 5-Higher) by sex and by 3 age groups (15-29, 30-49, 50+).

The second set of constraints include two different option to consider this auxiliary information:

- total quarterly signals of regular employment for 15-64 individuals by sex;
- annual mean of signals of regular employment for 15-64 individuals by sex.

Thus, starting from the intermediate quarterly weights, we used a calibration approach, with several different sets of constraints to get the new final weights. In this paper we present results for some of them, pointing out that all the constraints have been defined at NUTS 2 level, the level of partitioning of the calibration model.

6. Application and results

We compared the results of these different calibration models for quarter 2021Q4 LFS with the results of the census 2021. In the census, employment status at time t for unit k is modelled as a binary latent variable (t, k) (employed or not). Census survey, LFS and administrative sources are treated as imperfect measures of the target process. Coherence of employment estimates are open issue in our country, in particular for the census, that releases its figures 12 months after the reference period, with much less timeliness than then LFS.

Table 1 – *Calibration models considered.*

Calibration model	No-Response Adjustment	Final Calibration
IT-LFS	Current Iterative Proportional Fitting	Current Calibration applied in the survey
ED_FR	Current Iterative Proportional Fitting	IT-LFS + Education level in the Population register
ES_FR	Current Iterative Proportional Fitting	IT-LFS + Education level in the Population register + Regular Employment signals in Labour Register
PBCAL_ES	First step Calibration	IT-LFS + Education level in the Population register + Regular Employment signals in Labour Register

Comparing IT-LFS and the whole census framework (List-sample, Area sample, Register) LFS shows a higher level of individuals with university degrees (and a lower percentage of individuals with education up to a high school degree),

considering the education level collected on the register, in addition LFS has a higher share of respondents with no employment signal in the reference week.

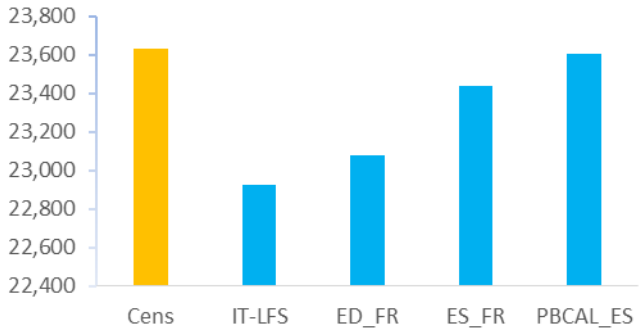
Given an estimate of employed people in the LFS that is lower of 3.0% than the census, at national level, the greatest differences are concentrated in the southern regions, in particular in Campania, Puglia, Calabria and Sicily. Overall, in these 4 regions, the Census estimates a higher number of employed people around 600 thousand (the differences exceed +10%).

Looking at Figure 2 we can summarize the results as follow. The current employment estimate from the LFS is 22.923 thousands (IT-LFS), about 3% lower than the Census figures (Cens). Introducing constraints in the calibration derived from RBI employment estimates, LFS employment estimate increases to 23 thousands including the education level (ED_FR).

Introducing the constraints on regular employment signals in the final weighting step, keeping the current adjustment for the total non-response increases the employment estimate to 23.437k (ES_FR).

Recalculating both the intermediate weight with a calibration step, as described in paragraph 4, and including in addition to the education level the constraints on regular employment signals in the final weighting step (PBCAL_ES) we get 23.603 thousand employed.

Figure 2 – *Total Employment Census 2021, LFS 2021Q4 with different weights.*



Source: ISTAT, Census, Labour Force Survey.

For employment rate 15+ at national level we get similar results, coherence with census improves, the last estimate with 2-step calibration almost closes the gap with the census figure. However, Nuts2 level analysis points out clearly a regional pattern (Figure 2). Including the auxiliary information of education level and, afterwards, of regular employment signals, the employment rate, based on the ILO definition, increases, in particular in the southern regions. In Campania employment rate raises

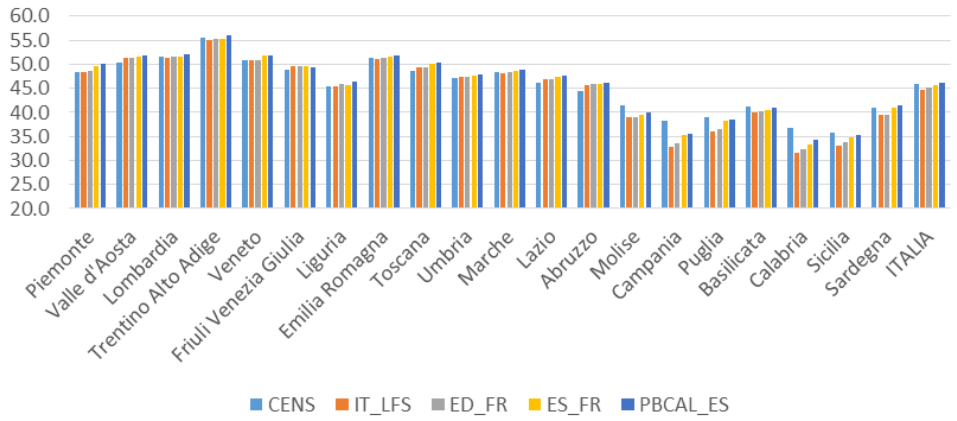
from 32.8 (the IT-LFS figure) to 33.5% including education level, 35.3% with regular employment signals and to 35.6% with 2-step calibration. Almost the same happens in Calabria, Sicilia and Puglia.

These first results represent further insights to study the differences between LFS and Census in Italy, an open issue that requires to be studied more in depth. For sure the response propensity of respondents is different for the two surveys and the under/over coverage as well.

Historically, the census is able to reach individuals who are very mobile in the country, perhaps a large part of them is residing in the south of the country but lives and works in the northern regions. Participation in the census is a strong motivation, linked to the administrative aspects of maintaining residence in a particular municipality, and the availability of CAWI interviewing mode for the Census may have been relevant for this particular subpopulation in order to improve the response rate.

On the other hand, participation of these individuals to the LFS could be more difficult, given that the survey technique is CAPI-CATI and they spend most of their time away from their residence, hence giving rise to a certain under representation in the sample. However, an advantage of the LFS, compared to census, is that data collection lasts less, so telescoping effect and measurement errors are generally reduced. The integration of the different sources and the use as auxiliary information in the LFS calibration seems to be able to reduce the bias due to the unbalanced coverage of certain sub-groups of population.

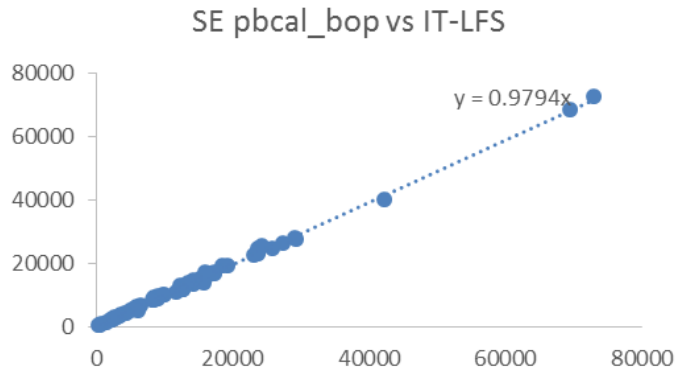
Figure 3 – *Employment rate 15+ at Nuts2 level, Census 2021, LFS 2021Q4 with different weights.*



Source: ISTAT, Census, Labour Force Survey.

Comparison of standard errors for a set of estimates obtained with 2-step calibration (PBCAL_ES) and the current calibration model (Figure 3) shows a slight increase for the precision of the estimates with the first option ($\hat{\beta}$=0.979).

Figure 4 – *Standard errors for the main aggregates at NUTS2 level, 2021Q4 LFS estimates and 2-step weights.*



Source: ISTAT, Census, Labour Force Survey.

7. Conclusions

The results show that the proposed innovations can be effective to address the survey problems (MRT, bias) and for improving the quality of the estimates. The crucial element is the integration with the information available from the registers. It is reasonable to think that introducing the same auxiliary information, education level and signals of regular employment, already included in the census latent class model, in the calibration model of LFS, leads to a better coherence of the estimates. For this purpose, calibration, beside the efficiency gain, is very effective to build a coherent system of surveys. Considering the different timeliness of the two surveys even a benchmarking of census employment to the LFS figure could be a suitable strategy, once the differences have been clearly identified.

The joint use of these kinds of auxiliary information in calibration estimators raises some questions.

Integrating administrative data into the LFS improves the accuracy of monthly and quarterly estimates? This should also be assessed by considering more detailed territorial domains, estimates of variations and particular subpopulations such as irregular workers.

Can we consider the integration of this information a good practice?

Which kind of signal for regular employment should we consider? The quarterly or the annual one?

Do they have a significant effect on the seasonal pattern of ILO Employment?

Further studies will be useful to finely tune the model and focus on specific aspects, but this work has already shown that the integration of administrative sources and its use in the calibration model is a very promising way to improve quality of IT-LFS.

References

DEVER J.A., VALLIANT R. 2010. A comparison of variance estimators for poststratification to estimated control totals, *Survey Methodology*, Statistics Canada, Catalogue No. 12-001-X, Vol. 36, No. 1, pp. 45-56.

DEVILLE J.C., SÄRNDAL C.E. 1992. Calibration Estimators in Survey Sampling, *Journal of the American Statistical Association*, Vol.87, No. 418, pp.376-382.

DI ZIO M. FILIPPONI D. 2023., Multi-source statistics in the Italian permanent census . In *Proceedings of the Workshop on Methodologies for official statistics*, Rome, pp.31-44.

FULLER W. A. 2002. Regression estimation for survey samples, *Survey Methodology*, Statistics Canada, Catalogue No. 12001XIE, Vol. 28,No. 1, pp.523.

FULLER W.A., RAO J.N.K. 2001. A Regression Composite Estimator with Application to the Labour Force Survey, *Survey Methodology*, Statistics Canada, Catalogue No. 12001, Vol. 27, No. 1, pp. 45-51.

LUNDSTROM S., SÄRNDAL C.E. 1999. Calibration as a Standard Method for Treatment of Nonresponse, *Journal of Official Statistics,* Vol. 15, No. 2, pp. 305-327.

RENSSEN R.H., NIEUWENBROEK N.J. 1997. Aligning estimates for common variables in two or more sample surveys, *Journal of the American Statistical Association*, Vol 92, Issue 437, pp. 368-374.

Antonella IORIO, ISTAT, iorio@istat.it
Alessandro MARTINI, ISTAT, alemartini@istat.it

INCREASE RESPONSE RATE IN ITALIAN SURVEYS: A REWARD MODEL PROPOSAL[1]

Michele Camisasca, Fabio Falcone, Livio Gigliuto, Giuseppina Sola

Abstract. The proliferation of new sources of information, varying in reliability, reinforces the view that society as a whole can be observed, described and quantified through an analysis of the data generated and stored in the Internet of Things (IoT). The automatic gathering of data from our actions, purchases and online browsing habits provides a direct and automated method of data collection, offering insights into our behaviour and habits.
In this context, private research institutes and national statistical offices (NSOs) must develop strategies to increase direct respondent participation and reduce the occurrence of non-response. Indeed, the use of monetary incentives is not always sufficient, and further measures must therefore be considered. The paper begins with an analysis of legal provisions concerning mandatory participation and goes on to propose a new approach that transforms the experience of participating in surveys from a "duty" to be fulfilled to a "benefit" to be enjoyed. This proposal would shift from a punitive approach involving monetary sanctions to a purely administrative approach based on verifying statistical compliance and issuing a specific attestation document. The possession of this document would then determine eligibility for rewards and/or benefits.

1. Direct surveys and other sources of statistical information

The increasing availability and reliability of new sources of data, which have emerged in competition with direct respondent participation, are reinforcing the view that the only effective approach to examining, describing and quantifying society is through the analysis of the data that each individual leaves in the internet of things. Data is collected automatically from individuals' actions, purchases and online behaviour. Much of the information that used to be provided by traditional research techniques (polls, market research and quantitative surveys), can now be captured through the use of big data. This makes respondents less likely to give information they think statistical agencies already have (Savoldelli, 2018). As a result, response rate decreases.

It would, however, be erroneous to assume that big data can entirely replace more traditional survey methods. This is due to the fact that the latter are based on

[1] Authors contributors: § 5. to M. Camisasca; § 4., 4.1 and 4.2 to F. Falcone; § 2.1 and 3.1 to L. Gigliuto; § 1. and 3. to G. Sola. Paragraphs 2.2 and 3.2 are written jointly by F. Falcone and G. Sola.

established conceptual frameworks and can be produced in a systematic and reproducible manner. Consequently, they are better suited to meet the particular cognitive requirements of researchers, public decision-makers and private users. Additionally, there is a possibility that big data may yield an inaccurate or misleading representation of reality (Pratesi, 2017).

The success of a traditional survey depends on the participation of potential respondents: setting clear cognitive objectives, preparing the perfect questionnaire and identifying accurately the target population is useless if responding units (individuals, households, businesses) decide not to respond. So, private research institutes and national statistical offices (NSOs) are obliged to experiment with new solutions to encourage such participation.

2. Analysis of the response rate

2.1 Response rate in market research

The inclination to engage in market research frequently coincides with the everyday lives of individuals and their situational and tangible readiness to dedicate the requisite amount of time and attention to a given survey.

One of the key factors influencing participation is the method used for data collection. The average response rates demonstrate a greater success for CAWI (Computer-Assisted Web Interviewing) compared to CATI (Computer-Assisted Telephone Interviewing), with the former achieving a response rate of approximately 7%, and the latter a rate of 2.6%[2]. Nevertheless, the difference between these methods is not substantial. While a notable difference is apparent, the response rate for the total number of links dispatched is comparable to the response rate for the total number of contacts utilising the CATI approach[3].

There are four principal reasons for this difference:

[2] The paper analyzed 10 surveys, each with the same sample size (just over 13,000 contacts), related to current affairs and carried out by the Piepoli Institute during the year 2019. The surveys were conducted using the CATI method. The response rate is calculated as the ratio of the completed interviews to the total number of contacts.
With regard to the CAWI method, a further five surveys, also conducted by the Piepoli Institute and carried out between 2021 and 2023, were analyzed. These related to social and topical issues and involved varying sample sizes (ranging from 10,000 to 30,000 units to whom the relevant survey participation link was sent. The response rate is calculated as the ratio of completed interviews to the links sent.

[3] On average, 13,300 people were contacted in the 10 CATI surveys; while an average of 18,000 links were sent in the 5 CAWI surveys.

a. spread of technology: the growth and increasing accessibility of the CAWI method has been enabled by the proliferation of the internet. However, this has not been universally embraced, with certain sections of the population remaining less familiar with digital technologies;
b. perceived invasiveness: telephone interviews frequently give rise to a sense of intrusion into private life, which in turn places a particular responsibility upon the role of interviewers;
c. contingency of the individuals involved is another factor to consider. People tend to respond more willingly to CAWI surveys during free time and without pressure, whereas telephone interviews require immediate engagement.
d. trust: the respondent may feel more confident answering an online survey rather than providing personal information during a phone call, especially if they do not know the interviewer.

It can be reasonably deduced that CAWI is somewhat more efficacious than CATI, as it appears to offer greater convenience to participants, thereby facilitating research participation even among younger age groups. In contrast, CATI may be a valuable tool for reaching segments of the population with limited familiarity with the Internet. It also allows for the maintenance of the valuable relationship between the interviewer and the respondent, which can help to maintain the respondent's focus.

2.2 Response rate in official statistics surveys

Official statistical surveys conducted by the Italian National Statistical Institute (Istat) generally show an uneven response rate. The trend is mainly influenced by the increase in perceived reporting burden and the expectation of sanctionability resulting from non-response. With regard of reporting burden, companies complain of: low rotation of the audience (often dictated by the methodology and criteria established by European regulations) and consequent involvement of the same company in numerous surveys; collection of information perceived as redundant, because it has already been provided to administrative authorities, for other surveys or previous editions of the same survey; the completion of the questionnaire at inappropriate times (periods of intense work activity or, conversely, periods of company closure) and/or addressed to units structurally not equipped with dedicated staff, etc. Sanctionability results in a physiological increase in response rate for

surveys whose violation results in the imposition of a fine (from 60% per year for mandatory non-sanctionable surveys up to 75% per year for sanctionable ones)[4].

In non-sanctionable mandatory surveys, the highest response rates are recorded for surveys with statistic sample limited to a few survey units[5], with which there is a tendency to establish a direct loyalty relationship, and those in which survey units perform public interest services[6]. Conversely, response rates are greatly reduced in short-term surveys[7] and if companies are involved in more than one statistical survey, due precisely to reporting burden. In particular, if survey units have to partecipate to non-sanctionable mandatory surveys and to sanctionable ones, they give priority only to sanctionable surveys.

Because of sanctionability usually tied to meeting size thresholds (number of employees and/or turnover), for sanctionable surveys the highest response rates occur for that ones whose field of observation consists exclusively of larger structural size units[8]. In contrast, for surveys whose statistic sample also consists of smaller units – for example, units operating in service sector – there is a less pronounced increase[9].

One possible explanation for these values is to be found in the more or less widespread organizational articulation of the survey units and the non-'punitive' perception of the financial penalty. In other words, some survey units consider it more cost-effective to pay the penalty rather than provide the data[10].

[4] The paper analyses the response rates of 23 surveys carried out by Istat in 2023 using the CAWI method (including 12 sanctionable surveys). The surveys were in the fields of environmental and spatial statistics, economic statistics, socio-economic statistics and national accounts.

[5] For instance, but not limited, in Rail Transport Survey, the response rate is 100 percent.

[6] For example, but not limited, the response rate for both the Survey of Museums and Similar Institutions and the Survey of Libraries is approximately 91 percent.

[7] In the Monthly Confidence Surveys, the response rate ranges from 45 percent for manufacturing firms to just under 19 percent for service firms.

[8] In the survey on employment, working hours, wages and labour costs in large enterprises, the response rate is approaching 98 percent.

[9] In Services Revenue Survey, the response rate is about 82 percent.

[10] This phenomenon includes units with very high turnover that perceive the amount of the financial penalty as insignificant, as well as units that choose to pay the penalty amount instead of dedicating an appropriate resource or turning to the accountant or an outside consultant to source the data. It also includes those units that, despite assurances of statistical confidentiality, choose not to provide the data due to privacy issues.

3. The use of incentives to increase the response rate: advantages and disadvantages

In general, respondents participate in surveys because they wish to make themselves useful, have a genuine interest in survey topic, or participation guarantees a tangible benefit, usually in the form of an incentive (Lipps, 2019, Singer, 2002).

Incentives are mainly monetary (money or goods of economic value received directly by respondent) and semi-monetary (money or goods of economic value that associations and charities obtain by virtue of the participation of third parties), and only rarely non-monetary (rewards of psychological and social value received directly by respondent) rewards or benefits[11].

Incentives increase motivation to participate by two mechanisms: they activate the reciprocity norm[12] and compensate barriers to participation, thereby increasing response rate and questionnaire completion rate, building respondent loyalty and making them likely to participate in future surveys, and reducing survey costs[13].

Conversely, offering incentives may induce to participate only people interested on the reward and not in the survey topics. They might give unmotivated or random answers just to complete the questionnaire and get the incentive. So it could introduce bias because respondents feel an obligation to give positive answers or answers in line with the proposer's expectations, instead of honest answers. Consequently, incentives must be sufficiently attractive to motivate participation, while avoiding excessive generosity.

3.1 Incentives in market research surveys

The objective of market research surveys is to facilitate improvements in products or services offered by companies. This is done by collecting feedback from consumers or potential consumers who have purchased or could purchase the products or services in question. Consequently, market research serves the dual purpose of benefiting both the company and the consumer. From the company's perspective, it enables the optimisation of the effectiveness of its offerings in the

[11] Monetary and semi-monetary incentives include: discount codes or coupons; cash or gift card rewards; free samples or trial periods; offering airtime top-ups; professional development opportunities (access to webinar or workshop); social media shoutout or shares; customized merchandise or swag; donations to associations (voluntary or scientific research or non-profit associations, etc.).
Non-monetary incentives include: appeal to self-interest; appeal to social status; appeal to public value.

[12] At a psychological level, the norm of reciprocity causes the respondent to read all questions and response options carefully before selecting the one that best represents his or her experience.

[13] As a result, less time can be spent on field data collection because the desired response rate is achieved earlier.

target market. From the consumer's standpoint, it facilitates the identification of avenues for greater satisfaction with their needs.

From the perspective of the consumer, the primary motivation for participation in market research remains the potential for enhancing the quality of the products and services they utilize, or could utilize, and, consequently, the quality of their own lives. However, despite the common perception that offering incentives to participate in market research is essential, it is crucial to carefully consider the target audience and the design of such rewards to ensure they do not introduce bias into the data collected (Marbach, 2000).

In the case of surveys conducted on enterprises, it is often found that a beneficial incentive can be in the form of a summary of the research findings provided at no cost. Demonstrating how and to what extent the results of the survey may affect the company's decision-making process is a more effective strategy for securing cooperation than the offer of a financial incentive. Conversely, if the respondents are consumers, a symbolic reward system may be employed, whereby the incentive does not unduly influence the respondents' responses. It is important to note that the distribution of prizes should be perceived as an act of goodwill. The objective is to foster a stronger relationship between the researcher and the respondents, rather than focusing on the company, or the product being tested, as a means of improving the quality of the relationship. For this reason, it is imperative that rewards are not categorised as part of the product under investigation. To illustrate, if the research pertains to shower gel, the reward cannot be the gel itself, nor can it be classified as toiletries.

3.2 Incentives in official statistics

Official statistics rapresent pubblic good[14]. They are an indispensable pillar of democracy, and because they impartially provide credible information, they enable survey units to go beyond the “perception” of their own experience to observe and quantify the surrounding reality in a national and international comparison[15]. Therefore, participating in the production of official statistical information is a duty

[14] The basic principles were set out in the ONU UN Resolution, Fundamental Principles of Official Statistics (A/RES/68/261 from 29 January 2014), https://unstats.un.org/unsd/dnss/gp/FP-Rev2013-E.pdf.

[15] This obligation at the european level was established by European Statistics Code of Practice, in particular in Principle 2 – Mandate for Data Collection and Access to Data, according to which “Administrations, enterprises and households, and the public at large may be compelled by law to allow access to or deliver data for European statistical purposes at the request of statistical authorities” and “the statistical authorities may compel response to statistical surveys”.

of survey units. In some cases, the obligation to participate can lead to penalties for failing to comply with this requirement. Generally it is a fine, applied by the national statistical offices themselves or by legal authorities on their report[16]. So, because penalties are foreseen, monetary incentives cannot be used for respondents. However, except for a single experiment by the U.S. Census Bureau in conducting two longitudinal demographic surveys (Survey of Income and Program Participation – SIPP and Survey of Program Dynamics – SPD)[17], no similar monetary incentives experiences can be found in other states, European or otherwise.

More workable, however, would seem to be the original and unusual idea of facilitating the collaboration of survey units by providing them with software or apps designed to interconnect with the management software in use and simplify data retrieval and questionnaire completion. Deserving of mention is the experience of the Australian Bureau of Statistics, which offers small and medium-sized companies, affected by the Quarterly Business Indicators Survey, a Web app (ABS Business Reporting) developed in collaboration with accountants and industry consultants. The goal is to facilitate the quarterly transmission of mandatory statistical data[18].

4. Enhancing survey units participation: a proposal for a reward model

The transformation from a top-down authoritative model to an egalitarian approach seems to be a viable alternative to ensure compliance with the principles of cooperation and good faith between survey units and PA. It is in this direction that the proposal to incentivize voluntary statistical participation moves.

4.1 Obligation to reply and sanctionability in official surveys

In Italy, the obligation to provide statistical data for surveys under the National Statistical Program (NSP) is stipulated in Article 7 of Legislative Decree No. 322 of

[16] For example, in England, under the Census Act 1920, the national statistical office cannot challenge violations, but merely supports the Crown Prosecution Service in the relevant prosecution for the purpose of applying fine.

[17] The first experiment was conducted from 1996 to 2000 for the SIPP survey. Survey units were sorted by size and divided into three incentive groups ($0, $10, $20 in paper vouchers and debit cards). A $20 incentive significantly reduced non-response rates among households, especially those with low incomes or near poverty, while the $10 incentive group did not have significantly higher response rates. At an expenditure of about $415,000 over 4 years, an increase of about 2.25 percentage points was estimated over the rate that would have been achieved without using incentives (Creighton, 2001).

[18] In-depth information can be found at the link abs.gov.au/participate-survey/business-reporting.

1989. The obligation covers all public administrations, agencies[19] and institutions. Private entities are also subject to the obligation, limited to the surveys specified in the NSP and included in the "Mandatory List". For some surveys included in the "Sanctionable List", non-response is subject to a monetary fine. Both lists are updated annually.

There are two articles of Legislative Decree No. 322 of 1989 (7 and 11) that, alone, determine the structural configuration of the administrative penalty in statistics. But it is an 'imperfect' configuration, which has failed to adapt to the many technological changes in the mode of data acquisition (e.g., questionnaire is no longer hard copy, but digital).

Article 7 is a clear expression of the authoritative power of government departments to impose their own determinations on other parties in terms of the behavior to be carried out. It has proven to have little impact compared to its original objective: in fact, despite the provision of an obligation, not all violations are subject to an administrative fine, but only those that fall under the "Sanctionability List".

This choice by the legislator renders violation inconsequential, effectively neutralizing the very scope of the legislation. Moreover, the tendency to set high sanctionability thresholds (dimensional and/or turnover) has also unintentionally led to the failure of smaller units to acquire data (as discussed in more detail in Section 2.2).

Article 11 identifies ways and subjects that enforce sanctionability. It has also manifested, over time, the rigidity and inadequacy of the dynamics of assesment, leading to a deep crisis of the sanction system.

4.2 A proposal for a reward model

In order to overcome the critical issues highlighted, a transition to an alternative model to the sanction system is proposed, within which the survey units will no longer perceive the obligation to respond as an imposition, but as a benefit to be enjoyed.

The model aims to introduce elements of simplification, moralization and transparency in order to reduce the phenomenon of non-response, involve the survey units in a multi-subjective context that incentivizes virtuous behavior, creates and consolidates the relationship of trust between the units and Istat, and recognizes them benefits and reductions of various kinds (concessions, subsidies, grants, specific economic advantages, access to certain services and/or multiple commercial and financial opportunities, etc.).

[19] Such as non-profit entities, associations, etc.

With unchanged legislation, it would be a shift from a repressive, 1:1 sanction model (determining body and offender) linked to the payment of a fine, to a reward model with a strong deterrent character and with the involvement of a varied audience of public and private stakeholders[20] that simultaneously serves as a lever and a multiplier effect to induce the detecting units to comply with the obligation. Therefore, compliance will no longer be seen as the result of adverse selection or reporting burden, but rather will paradoxically be perceived as an opportunity to be taken advantage of.

Greater is the multiplicity of rewards and stakeholders involved and greater will be the upsurge for fulfilment.

Borrowing from experience gained in other contexts different from the purely statistical sphere, the model finds its realization in Istat's issuance of a document valid for administrative purposes, called Statistical Regularity Document (SRD), containing attestation of the fulfillment of statistical obligations by the survey units.

This document assumes the function of comprehensive monitoring of the statistical participation of each survey unit and certifies its statistical regularity by assigning a rating[21].

Basically, the SRD will be subjected to a system of synthetic indicators of the fulfillment of the statistical requirement, which will take into account the different statistical quality and regularity standards (statistical rating) – including, but not limited to, number, timing, and percentage of fulfillments –, using a score between a minimum (*) and a maximum (***).

Survey unit will obtain the basic score (*), increasable by a fraction of a score (+) if it meets all requirements. Achieving three fractions (+++) will result in an additional point, until the maximum score is reached (***).

Based on the acquired score, stakeholders may grant individual units reward benefits[22]: for example, it could count as an additional score in calls for proposals, a requirement for access to reserved portions of public funding or other forms of business support (such as tax credit, tax bonus, guarantee grant, capital subsidy, interest subsidy, soft loan), or more or less extensive periods of statistical 'relief'[23] (during which some questionnaires will not be administered or the number of questions will be reduced).

---

[20] The main stakeholders will be central and local state governments, tax agencies, chambers of commerce and trade associations.

[21] The SRD has the value of a declaration of science, to be placed among the acts of certification or attestation, having a merely declaratory character of the data, held by the Public Administration.

[22] SRD does not establish the right to the bonus. It exists regardless and it is linked to the requirements for accessing the benefit. Therefore, it does not affect the origin, extinction, or forfeiture of the right, but arises solely on an administrative level.

[23] The initiative launched in Canada on January 1, 2015 to benefit small companies with a good prior history of statistical participation (Accumulated Response Burden Initiative – ARBI).

SRD can be used alone or in combination with other rating systems (e.g., with legality or sustainability ratings), making it flexible and adaptable to any assesment[24].

Due to the operational distance between the two models described (sanction system and the reward one), an immediate transition from one to the other is not conceivable. A period of coexistence will be inevitable, which should be seen as an opportunity to test the functionality of the new reward model.

In fact, by taking advantage of the provision for the proceeds of administrative penalties to flow into a special chapter in Istat's budget (Art. 7(1)), an experiment (adequately publicized) can be launched establishing a special fund aimed at guaranteeing the coverage of benefits provided exclusively to the survey units involved in certain surveys. This will allow comparison between the response rate of the experimental edition with the previous ones of the same survey, so that we can assess how much the reward system may have affected the response rate percentage.

5. Conclusions: survey units' engagement and civic partecipation

Given today's information overload, it is undeniable that official statistical information constitutes a fundamental pillar of modern participatory democracy. The essential criteria that official statistics fulfil – relevance, timeliness and punctuality, accuracy and reliability, coherence and comparability, accessibility and clarity – are the basis for facilitating citizens' individual and collective engagement with the institutions that govern them and the exercise of their right to participate in public decision-making.

Participation in statistical surveys as a respondent contributes to the generation of more optimal public policies. Such policies are more efficient, fairer and easier to implement. This can be attributed to respondents (even those considered "ordinary citizens") possessing a deeper understanding of actual needs, and thus providing information that is more accurate and beneficial for decision-making purposes.

Civic participation ensures that these services are tailored to people's real needs and preferences. Citizen participation in statistical surveys fosters a sense of ownership and trust in the data collection process. When citizens are actively involved, they are more likely to trust the results and the institutions that produce them. As a result, participation in statistical surveys as a respondent contributes to the generation of more optimal public policies.

[24] SDR could also be provided to Prefectures to determine the amount of the fine under Article 11 of Law No. 689/81.

To enhance participation and promote active and mindful involvement, it is vital to transition from a system reliant on compulsion, which has evidenced suboptimal efficacy, toward a collaborative and egalitarian approach. This entails engaging with respondents (individuals, families, businesses, and institutions) on an equal footing, in a way that is mutually beneficial. To put it another way, official statistics must abandon the practice of regarding respondents as mere "targets" to be reached and instead recognise them as genuine participants in the statistical process. Such an approach requires a focus on empowering the respondent to enhance their capacity for critical thinking and analysis. It is insufficient for respondents to have merely access to information and the capacity for independent reasoning; they must also be made aware of how the data collected in various surveys has been used by those making public policy decisions.

To facilitate greater comprehension amongst respondents, it would be advantageous to provide each individual's digital domicile with an informative booklet on an annual basis. This booklet would delineate the measures taken based on the provided data. Such an initiative might be modelled on the example of the Internal Revenue Service, which has distributed a summary table of the allocation of taxes related to tax returns. The aforementioned document delineates both how tax resources are utilized by the state and the extent to which individuals contribute to disparate domains of public expenditure. To provide tangible recognition for their participation, one may consider utilising the Statistical Regularity Document (SDR), as proposed in this study, or, with respect to individual respondents, the implementation of incentives closely aligned with the specific survey for which they provided data.

In conclusion, the implementation of an appropriate, strategic and coordinated use of a range of communication techniques and incentives represents a critical factor in advancing the quality of statistical participation. It is solely through such means that official statistics can facilitate the growth of a robust civic culture. When individuals possess greater awareness of their contributions, they tend to become more responsible and aware of the practical implications associated with each choice. This heightened awareness also encourages active engagement in decision-making forums, where they can propose innovative solutions.

References

BERĘSEWICZA M., LEHTONEN R., REIS F., DI CONSIGLIO L., KARLBERG M. 2018. An overview of methods for treating selectivity in big data sources, *Statistical Working Papers*, Eurostat.

CREIGHTON K.P., KING K.E., MARTIN E.A. 2001. *The Use of Monetary Incentives in Census Bureau Longitudinal Surveys*. Washington: U.S. Census Bureau.

FLORESCU D, KARLBERG M, REIS F, et al. 2014. Will 'big data' transform official statistics? *retrieved from http://www. q2014.at/fileadmin/user_upload/ESTAT-Q2014-BigData OS-v1a.pdf.*

LIPPS, O., HERZING, J. M. E., PEKARI, N., ERNST STÄHLI, M. POLLIEN, A., RIEDO, G., REVEILHAC, M. 2019. Incentives in surveys. *FORS Guide*, n. 08, Version 1.0. Lausanne: Swiss Centre of Expertise in the Social Sciences.

MARBACH, G. 2002. *Le ricerche di mercato negli anni della discontinuità*, Roma: UTET.

PRATESI M. 2017. I big data: il punto di vista di uno statistico. Menabò di Etica e Economia, *retrieved from https://eticaeconomia.it/i-big-data-il-punto-di-vista-di-uno-statistico/*

SAVOLDELLI S. 2018. Perché I big data non uccideranno la ricerca tradizionale, retrieved from *https://sandrosavoldelli.wordpress.com/2018/11/25/perche-i-big-data-non-uccideranno-la-ricerca-tradizionale/*.

SINGER, E. 2002. The use of incentives to reduce nonresponse in household surveys. In GROVES, R. M. , DILLMAN, A., ELTINGE J. L., & LITTLE R. J. A. (Eds.), *Survey nonresponse,* Hoboken, NJ: Wiley. pp. 163-177.

Michele CAMISASCA, Istat, michele.camisasca@istat.it
Fabio FALCONE, Istat, fabio.falcone@istat.it
Livio GIGLIUTO, Istituto Piepoli, liviogigliuto@istitutopiepoli.it
Giuseppina SOLA, Istat, giuseppina.sola@istat.it

WEIGHTING A SNOWBALL SAMPLE THROUGH A PSEUDO-CALIBRATION: THE CASE STUDY OF SAME-SEX CIVIL UNIONS IN ITALY

Marco D. Terribili

Abstract. This paper presents an approach to addressing the challenges of selection bias and non-probabilistic characteristics of snowball sampling design, particularly in the context of social research involving hidden or hard-to-reach populations. The primary aim is to refine the weighting methodology of snowball sampling by introducing the pseudo-calibration technique to adjust the direct weights, equal to one, according to some available auxiliary variables, with the ultimate goal of producing more reliable and unbiased estimates.
The data for this case study is derived from a snowball sample survey, called the "Over the Rainbow" project, carried out on Instagram users who tag their photos with popular LGBTQ+ community hashtags. The username list is collected using web-scraping tools that identify relevant users. To address the sampling design's limitations, the study employs calibration and post-stratification methods. Calibration involves adjusting the weights of the sample data to match known population totals, while post-stratification involves dividing the sample into subgroups that align with known demographic distributions.
The proposed sampling weights are benchmarked on the Istat survey on all civil unions between same-sex couples, celebrated by Italian municipalities since 2016; this known totals source represents a reliable external reference to adopt sampling weights. The expected results of this study are twofold. First, the application of calibration should yield sample weights that are more representative of the target population, by aligning the sample distribution with known population totals. Second, post-stratification is anticipated to define pseudo-weights, adjusting the direct weights equal to one, refining the sample by ensuring that the subgroups within the sample correspond proportionally to those in the broader population. The combination of these methods is expected to significantly reduce the biases associated with traditional snowball sampling and attribute sampling weights not just equal to 1, as usually happens in snowball sampling.
The paper contributes to the statistical and social research field by offering a methodologically sound approach to improving the accuracy of snowball sampling designs, with practical implications for studying hard-to-reach populations on social media platforms.

1. Introduction

Snowball sampling is a type of purposive sample, particularly useful for recruiting individuals who are difficult to identify and for whom a predefined

sampling framework is unavailable. One of the main advantages of snowball sampling is its ability to ease data collection. This method involves identifying one participant who meets the study's criteria and then asking them to recommend other individuals with similar characteristics, thereby expanding the sample. This approach can approximate a simple random sample, potentially leading to more robust and reliable results (Abdul-Quader et al., 2006).

On the other hand, snowballing is a sampling design that, due to its non-probabilistic nature, does not allow for classical and probabilistic statistical inference on the reference population: determining the appropriate sampling weights to make inferences on this kind of sample can be particularly challenging.

1.1. How to weight a snowball sample? The (few) literature suggestions.

When attempting to make inferences with a snowball sample, the choice of sampling weights depends directly on several design factors and data collection process characteristics.

If the linkage between initial and subsequent participants can be preserved, Respondent-Driven Sampling (RDS) can be used (Heckathorn, 1997). RDS allows for the adjustment of sampling weights based on the network structure of the participants, thus facilitating more reliable inferences. This sampling method entails multiple snowballing waves, repeating these waves until the condition of Markov equilibrium is verified. In this way, the probability of being included in the final wave is independent of the probability of being recruited by the first one, which gets "watered down" as more waves are added. The main advantage of this method is that it asymptotically approximates a simple random sample, but, on the other hand, it requires tracking the recruiting IDs of responding people, in contrast with the essential privacy and anonymity constraints (De Rosa et al., 2020).

However, when such linkages are not available, researchers face the challenge of weighting the sampling units appropriately. The literature about weighting issues in this data collection situation is relatively sparse, with only a few contributions offering guidance: Shafie (2010), Snijders (1992), and Vitalini (2010). These three papers contribute significantly to the understanding of snowball sampling, yet none fully bridge the knowledge gap regarding the weighting of snowball samples: Shafie's paper excels in proposing innovative weighting techniques to correct for selection bias in snowball samples, using simulations to demonstrate the effectiveness of these methods in providing more accurate estimates. Snijders' study provides a thorough analysis of the mathematical properties and challenges of snowball sampling, offering a strong theoretical foundation and highlighting the complexities involved in sample selection probabilities. Vitalini's PhD dissertation

offers a comprehensive review, underscoring the critical importance of known selection probabilities for accurate statistical inference. Despite these strengths, the issue of how to precisely weight snowball samples remains unresolved, leaving a critical gap in the methodology regarding snowball sampling.

In this context, alternative techniques such as calibration, post-stratification, and pseudo-calibration have been employed to mitigate the limitations of snowball sampling. These methods involve adjusting the sample to match known population characteristics, thereby improving the accuracy of the inferences drawn from the sample.

In this context, according to known totals availability, on which to constrain, data on same-sex people in civil unions, coming from a previous research project, has been pseudo-calibrated.

2. The case study: the collected civil union snowball sampling units.

The case study data, focusing on civil unions in Italy, are derived from the undersigned PhD research project (Terribili, 2022), titled "Over The Rainbow", OTR from now on. The data collection process involved a comprehensive survey of Instagram users who tagged their pictures with popular LGBTQ+ community hashtags. This initial list of 8290 users was compiled using a web-scraping procedure in R. Subsequently, a questionnaire was disseminated to these users through Instagram, leveraging the same social network they actively engage with. To further expand and diversify the sample, the snowball sampling method was employed, where participants were encouraged to share the survey with others in their network. This approach effectively broadened the scope and depth of the data collected for the research: 638 users, named by other respondents, actively participated in the survey, also without being initially listed by the web-scraping procedure.

Unfortunately, at least for this study, the population(s) involved in the OTR survey were extremely young[1] and non-representative of the whole population. In fact, being the OTR sample non-probabilistic, we cannot strictly talk about representativeness. Anyway, the characteristics of people caught can be considered a result too. For instance, the higher collaboration from women was an important feedback about the web-scraping phase: in fact, also if most of the collected posts by the web-scraping script (according to the hashtags contained) were male-oriented, most respondents were female.

[1] The OTR respondents' average age was 24.84 years old, a mean value which decreases observing the median (23), and even more the mode (20).

This design bias led to just 30 individuals in civil unions. This relatively exiguous sample size will be one of the most crucial issues of the weighting phase, as will be explained in the following paragraph.

2.1. Benchmarking at Istat same-sex civil unions survey

Despite the limited sample size available for our survey, we have been privileged to have a robust benchmark to refer to. Istat has conducted an extensive survey encompassing all civil unions between same-sex couples that have been registered at the civil registry offices of Italian municipalities since 2016[2]. This comprehensive data collection effort has resulted in a known total of 18962 individuals in same-sex civil unions, up to the point of our data collection. This substantial and reliable dataset serves as a crucial external reference, available at different levels of aggregation, as shown in the table below:

Table 1 – Individuals *in a civil union up to 31st January 2020 (end of OTR data collection phase).*

Level		Modalities	No. of people in Civil union
Sex		Male	6964
		Female	11998
Age class		18-39	7172
		40+	11770
Educational level		Lower High School	7848
		High School	6448
		Higher High School	4666
Sex-Geographic area of birth	Female	Abroad	610
		North-West	3546
		North-East	2093
		Centre	2575
		South & Islands	1513
	Male	Abroad	915
		North-West	2581
		North-East	1587
		Centre	1947
		South & Islands	1324
TOTAL			18692

Source: Istat civil unions survey

[2] On June 5, 2016, the law No. 76 (commonly named "Cirinnà's law", according the Senator who proposed and submitted the law proposal), introducing civil unions between same-sex couples, came into effect in Italy. During the second half of that year, 2336 civil unions were formed, a particularly significant number, reflecting couples who had long been waiting to formalize their emotional bond. After this initial boom, there was a gradual stabilization in terms of number of civil unions incurred (Istat, 2023).

By leveraging this benchmark, we can use calibration and post-stratification model to adopt sampling weights that respect these known totals making the OTR survey results more representative and statistically sound.

3. Pseudocalibrating to provide sampling weight to non-probabilistic sample units

Handling non-probabilistic samples through calibration and/or post-stratification methods entails some crucial initial choices. If the primary goal of these methods is to adjust the initial sampling weights of respondents to align with known population totals from an external survey, which initial weights to consider is not trivial, as it happens with the direct weights (given by the reciprocal of inclusion probabilities, as in the Horvitz-Thompson estimator) for probabilistic sample surveys.

The theory indeed suggests that, in the case of convenience samples, each statistical unit involved in the survey represents only itself, thus having an (initial) sample weight of 1. This does not preclude making inferences by referring to the universe, at least in cases where the population totals are known. However, such a small and constant sample weight certainly complicates the calibration procedure in terms of convergence to obtain a result.

In this context, a pseudo-calibration estimator can be adopted for integrating the non-probability sample with a probabilistic one, assuming both samples contain relevant information for estimating the population parameter. These proposed estimators employ pseudo-weights (Elliot, 2009; Baker et al., 2013), sharing a structural similarity with the adjusted projection estimators, but adopting a different inferential approach and informative setup (Golini and Righi, 2024).

The underlying idea of applied pseudo-calibration is to correct the direct weights (d_i), equal to 1 for all units belonging to the non-probabilistic sample, as suggested by the sampling literature, with an adjustment factor, to obtain a pseudo-weight (δ_i). In our case, we adopted a post-stratification to proceed with this adjustment.

The post-stratification domains represent an exhaustive and mutually exclusive partition of the population, which can be made more and more detailed and granular, for example, by geographic area, or by geographic area crossed with gender.

The easiest pseudo-weights (δ_i) set is intuitively given by the product of the direct weight d_i, as already said equal to 1, with the inverse of the sampling fraction (*n/N*) for all *n* sample units involved in the OTR survey (*s*):

$$\delta_i = d_i \cdot \frac{N}{n} = K \qquad \forall\, i \in s \tag{1}$$

Then, this sampling fraction, and its reciprocal, can also be computed for each mutually exclusive, and jointly exhaustive, subgroup - named *h* - of the population, for which the number of civil unions contracted in Italy between 2016 and January 2020 is known, providing different sets of pseudo-weights to use and to compare.

$$\delta_{i_h} = d_i \cdot \frac{N_h}{n_h} = K_h \qquad \forall\, i \in s_h \tag{2}$$

These different scenarios result in various sets of initial pseudo-weights, on which calibration was tested on the margins, given by the variables listed in Table 1.

Therefore, the final calibrated estimator is as follows:

$$\hat{t}_{Y_{CAL}} = \sum_{i \epsilon s} y_i \cdot w_i = \sum_{i \epsilon s} y_i \cdot \delta_i \cdot \gamma_i = \sum_{i \epsilon s} y_i \cdot \frac{N_h}{n_h} \cdot \gamma_i \qquad \forall\, i \in s, h \in H \tag{3}$$

$$\text{where} \begin{cases} min_{w_i}\{\sum_{k \in s} G_i(w_i - d_i)/q_i\} \\ \sum_{k \in s} x_i \cdot w_i = t_i \end{cases}$$

Basically, the final weights w_i are obtained by solving an optimization problem, in which $G_i(w_i - d_i)$ is, a non-negative, and strictly convex pseudo-distance function, continuously differentiable to w_i (Deville and Sarndal, 1992).

In this way, by progressively sharpening the pseudo weights system δ_i, a post-stratification system was first utilized, dividing the sample into exhaustive and not overlapped subgroups, according to the variables of sex and geographical area of birth, ensuring each subgroup reflects the population's known demographic distribution. Then the pseudo-weights δ_i have been calibrated to match the population's marginal distribution across variables such as sex, age classes, and geographic area, getting the final weights set w_i.

The pseudo-calibration can be, in other words, described as the process that leads to the calibrated final weight from a pseudo-weight, used to avoid the direct weight equal to 1. This technique highlights how each method strives to enhance the representativeness and reliability of survey data, when auxiliary data is available and when the minimum distance between weight sets is guaranteed and maintained.

3.1. The weights variability: something to consider and control

A calibration model was applied based on the availability of auxiliary data, with the input pseudo-weights being varied. The 1+CV² (one plus Coefficient of Variation squared) formula of Kish (1992) was then calculated for the resulting final weights.

This calculation quantifies how the variability in weights approximatively affects the estimates, allowing us to assess and manage this variability to ensure that it does not lead to increased variability in the estimates. By carefully monitoring and adjusting the weights, we aim to maintain the precision and reliability of the survey estimates, thus enhancing the robustness of our inferential procedures.

Table 2 – *Weights systems Kish's variability indicator.*

Weight system label	Direct or Pseudo-weight	$1+CV^2$	Calibration model	$1+CV^2$
a.	$d_i = 1$	1	Sex + Age classes + Educational Status	3.135
b.	$\delta_i = N/n$	1		3.277
c.	$\delta_{ih} = N_h/n_h$, where h is a Geographical Area	1.107		3.344
d.	$\delta_{ij} = N_j/n_j$, where j is a Geographical Area * Sex	1.672		3.687
e.	$d_i = 1$	1	Geographical Area * Sex	1.672

Table 1 shows that, from the perspective of sample weights and their variability, using constant direct weights equal to 1 does not offer a substantial advantage once the weights are calibrated to meet the known marginal totals available from the external source Istat. The $1+CV^2$ of the final weights increases from 3.135 to 3.687, rising by only 0.552 points, making the initial weight system increasingly granular.

Moreover, the adoption of pseudo-weights, that are not constant and equal to 1 does not result in a significant increase in the variability of the weights (from 1 to 1.672), but it also ensures adherence to the reality of the non-probabilistic design, which, being convenience-based, deviates from simple randomness and constant weights.

3.2. Comparing the estimates: which calibration model does entail the lowest bias?

The Istat survey not only provides the previously mentioned valuable known totals but also estimates of survey variables common to both the Istat survey and the OTR survey, which were deliberately included in the questionnaire of the latter. These variables include experiencing discrimination from one of the parents, following the respondents coming out, and experiencing microaggressions in the workplace (simplified here as mobbing) due to their sexual orientation.

The table below shows the OTR estimates obtained with the different weighting systems illustrated in the previous paragraph, as well as the Istat benchmark

estimates. Additionally, the relative errors, in terms of the percentage coefficient of variation (CV%) of the obtained estimates, are reported, which are understandably high due to the small sample size of the OTR.

Table 3 – *Comparison OTR and Istat estimates, varying weighting system and relative estimation errors (CV%).*

		ESTIMATES (%)						COEFFICIENT OF VARIATION (%)[3]				
Proxy Variable in common	By Sex	a.	b.	c.	d.	e.	*Istat*	a.	b.	c.	d.	e.
Bad coming out (mother)	TOT	47,7	48,0	50,0	51,4	**37,8**	*21,8*	32,5	32,1	30,9	31,8	23,2
	F	36,3	41,7	**31,3**	43,4	23,1	*28,8*	54,3	53,9	59,5	57,3	53,5
	M	54,4	**51,7**	60,8	56,0	56,0	*18,1*	41,3	41,8	37,2	39,5	21,9
Bad coming out (father)	TOT	34,5	28,1	27,0	**22,9**	31,9	*19,8*	43,2	53,1	49,9	56,9	23,9
	F	15,5	7,9	5,5	2,7	**22,0**	*18,7*	110,0	259,9	259,1	531,4	43,2
	M	45,5	39,8	39,5	**34,7**	44,0	*20,4*	46,8	51,5	47,8	52,7	27,9
Mobbing	TOT	13,2	**15,7**	14,7	15,1	28,5	*20,0*	98,1	86,9	89,9	89,5	20,4
	F	16,4	**21,2**	25,3	29,7	12,5	*21,5*	84,5	83,3	73,5	60,9	71,9
	M	11,4	**12,5**	8,6	6,7	48,4	*20,4*	154,8	137,2	175,4	228,9	14,0

Table 3 shows the propensity estimates related to the three common survey variables between the OTR survey and the Istat survey (in italics), differentiated by the biological sex of the respondent. The estimates calculated from the OTR data are obtained using the five weighting systems mentioned above, differing in how the pseudo-sampling weights have been calculated.

The differences are significant, and there is no clear method of pseudo-calibration that is unequivocally better than the others. However, what emerges is that the use of direct weights that are constant and equal to 1 never results in the estimate closest to the Istat estimates benchmark. Weighting system b., the one obtained by calibrating the constant weights to be equal to the reciprocal of the overall sampling fraction, appears to be the "best" (in bold) more frequently, but still always with substantial differences. The variable "Mobbing", related to microaggressions in the workplace, seems to be the one estimated better by OTR, perhaps also because it is the most similar, from a definition and ontology point of view, between these two questionnaires.

The second part of the table, the one on the far right, shows the relative errors in terms of CV% of the estimates just seen. The errors are extremely high, making the estimates difficult to publish. This factor inevitably limits the generalizability of the observed results and confirms how crucial sample size is, in non-probability sampling, even more so than in probability sampling.

[3] Istat estimates' coefficient of variation were not published and are not known, but are extremely lower than the OTR ones, reported in Table 3.

4. Final remarks

The previous paragraphs comprehensively analyze the challenges and strategies associated with non-probabilistic samples, particularly those collected via snowball sampling. If non-probability samples are commonly accepted and used among survey statisticians, there must be a coherent framework, such as auxiliary information to employ, and an accompanying set of measures for evaluating their quality and reliability (Kim, 2024).

One of the key takeaways is that using direct weights equal to 1, a common practice, may result in slightly lower variability of the final weights but fails to consider the specific nuances of snowball sampling, making it insensitive to the unique characteristics of the sample units.

In this context, pseudo-calibration emerges as a crucial step, as it ensures that the sample aligns with known totals at their intersections, not just their margins. This alignment seems to effectively control the increase in weight variability, which is a significant concern when dealing with non-probabilistic samples. Calibration takes this a step further by ensuring that the estimates are more closely aligned with the Istat reference values, thereby enhancing the accuracy and reliability of the models.

Importantly, the choice of the best weighting system is highly dependent on the availability of auxiliary information. This dependency underscores the necessity of having comprehensive and accurate auxiliary data to inform the weighting process.

Concluding, it is important to say that making inferences from non-probabilistic samples is indeed possible. However, it is important to emphasize the critical importance of having a sufficiently large sample size and a robust sampling strategy. These elements are essential to obtaining reliable and valid estimates, thereby addressing the inherent challenges of non-probabilistic sampling methods.

Acknowledgements

I am deeply grateful to my superiors and colleagues at Istat for granting me leave to explore this innovative topic, showcasing the institute's commitment to cutting-edge research.

Lastly, I extend my sincere thanks to all the LGBTQ individuals I interviewed: their insights were invaluable to this research.

References

ABDUL-QUADER A. S., HECKATHORN D. D., SABIN K., SAIDEL T. 2006. Implementation and analysis of respondent-driven sampling: lessons learned from the field. *Journal of Urban Health*, No. 83, pp. 1-5

BAKER R., BRICK J. M., BATES N. A., BATTAGLIA M., COUPER M. P., DEVER J. A., TOURANGEAU R. 2013. Summary report of the AAPOR task force on non-probability sampling. *Journal of Survey Statistics and Methodology*, Vol. 1, No. 2, pp. 90-143

DE ROSA E., DE VITIIS C., INGLESE F., VITALINI A. 2020. Il web-Respondent Driven Sampling per lo studio della popolazione LGBT+, *RIEDS - The Italian Journal of Economic, Demographic and Statistical Studies*, Vol. 74, No. 1, pp. 73-84

DEVILLE J.C., SARNDAL C.E., 1992. Calibration Estimators in Survey Sampling, Journal of the American Statistical Association, Vol. 87, No. 418, pp. 376-382

ELLIOTT M. R. 2009. Combining Data from Probability and Non-Probability Samples Using Pseudo-Weights. *Survey Practice*, Vol. 2, No. 6

GOLINI N., RIGHI P. 2024. Integrating probability and big non-probability samples data to produce Official Statistics. *Statistical Methods & Applications*, Vol. 33, No. 2, pp. 555-580

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

KIM K.S. 2024. Methodology of Non-Probability Samples Through Data Integration. *American Journal of Biomedical Science & Research*. Vol. 21, No. 5

KISH L. 1992. Weighting for unequal Pi. *Journal of Official Statistics*, No. 8, pp. 183-200

ISTAT. 2023. Matrimoni e unioni civili in ripresa ma ancora non ai livelli pre-pandemia. *Statistiche Report*

SHAFIE T. 2010. Design-based estimators for snowball sampling. Available at SSRN 2471006

SNIJDERS T. A. 1992. Estimation on the basis of snowball samples: how to weight?. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, Vol. 36, No. 1, pp. 59-70

TERRIBILI, M. D. 2022. Surveying the LGBTQ population (s) through social media. *AG About Gender - International Journal of Gender Studies*, Vol. 11, No. 21

VITALINI, A. 2010. L'uso delle reti sociali per la costruzione di campioni probabilistici: possibilità e limiti per lo studio di popolazioni senza lista di campionamento. Milan: Università Cattolica del Sacro Cuore

---

Marco D. TERRIBILI, Istat, terribili@istat.it

A CLUSTERING APPROACH FOR DETERMINING STRATIFICATION VARIABLES IN SBS SURVEYS

Ilaria Bombelli, Giorgia Sacco, Alessio Guandalini

Abstract. Many Structural Business Statistics (SBSs) surveys, according to European Regulations, must move from considering the Legal Unit (LU) as unit of interest towards considering the Enterprise (ENT) as such. This transition is not trivial, as many NSIs still need to provide estimates at a LU level, for comparability through the time.

Consequently, to modify and enhance the standard sample design based on LU to address this shift, it could be required to investigate an alternative stratification of the sample. To address this task, we propose to use a clustering algorithm, i.e., the K-prototype, to obtain groups of ENT and assess the variables' importance in the clustering result.

The algorithm is applied to several input datasets, obtained by sub-setting the ASIA ENT 2021 register, which includes all enterprises carrying on economic activities. The input datasets include ENT working on different sections of the statistical classification of economic activities in the European Community (NACE) and ENT included in the target population of the Community Innovation Survey (CIS) carried out by ISTAT. The clustering is applied separately to each aforementioned dataset. From the clustering result, we assess the variables' importance and identify the variables that mostly influence the obtained partition.

The most influential variables are used to build the new stratification of the ENT, hence they contribute to a new definition of the strata. The proposed stratification is used to allocate a sample of the same dimension as the one extracted with the current stratification. From the sample, we estimate some of the survey's target variables and their coefficient of variation (CV). The CVs are compared with the ones resulting from the current stratification. The comparison reveals that the efficiency of the estimates is preserved. In addition, the new stratification allows for reducing the number of strata and therefore also the processing time is limited.

1. Motivation of the study

Structural Business Statistics (SBS) surveys are carried out periodically by the Italian National Institute of Statistics (ISTAT), as well as the other National Statistical Institutes (NSI) in Europe. SBS surveys are used to collect information about, among the others, the organisation's structure, the economic activity and performance in terms of turnover, production and innovation of businesses over

time[1]. In addition, as required by the EUROSTAT regulations, definitions and harmonisation policies, the statistical units in SBS surveys carried out by European NSI must have the same definition to make comparable estimates.

Therefore, nowadays, all European National Statistical Institutes (NSI) are facing a new challenge: the transition from considering the legal unit (LU) as the statistical unit of analysis towards considering the Enterprise (ENT) as such. The LUs include *legal persons whose existence is recognized by law independently of the individuals or institutions which may own them or are members of them* and *natural persons who are engaged in an economic activity in their own right.* Moreover*, the legal unit always forms, either by itself or sometimes in combination with other legal units, the legal basis for the statistical unit known as the enterprise*[2]. The ENT is defined as *the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit*[3].

It is important to mention that each ENT can be made up of one or more LU, which in turn can belong to one or more ENT simultaneously according to a given *share*. Specifically, in the ASIA ENT register related to 2021, which includes all enterprises carrying on economic activities in 2021, almost all ENT are made up of a single LU: among the 4462146 ENT, 4409019 (98.8%) are made up of only one LU. For a complete description of the register, the reader can refer to Section 2 of the current paper.

However, even though NSI must provide estimates at the ENT-level, as required by EUROSTAT, they also want to provide estimates at LU-level, for comparability through the time. This makes the aforementioned transition not trivial. For this reason, a solution which has been adopted is to use a cluster sampling strategy: in this way, the ENT are considered as the statistical units during the sampling process; then, once a sample of ENT has been selected, all the LUs belonging to the sampled ENT will be provided with the questionnaire. This sampling strategy will allow to collect data at a LU-level, but the estimates are possible at both LU and ENT levels, by properly adjusting the sampling weights in the calibration step. This strategy has

[1] Source: https://ec.europa.eu/eurostat/web/structural-business-statistics#:~:text=Structural%20business%20statistics%20(SBS)%20describe,European%20business%20statistics%20.

[2] EUROSTAT glossary accessible at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Legal_unit.

[3] Regulation (CEE) n. 696/93 accessible at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31993R0696.

been adopted for selecting the sample of the Community Innovation Survey (CIS) of 2022, as it will be discussed in Section 2.1.

What is proposed in this paper are some suggestions for leading the transition process. More in detail, we want to highlight some insights to adapt the sampling design to this change and to stratify ENT efficiently. In order to study a new stratification, we apply cluster analysis on ENT; then, from the clustering result, we assess the variables' importance, and we identify the variables that mostly influence the obtained partition. In this way, the most influential variables are considered as the new stratification variables.

The data and the methodological tools descriptions are provided in Section 2, 3, respectively. Results are presented and discussed in Section 4. Finally, Section 5 sums up the main outcome of the research.

2. Data

The Business Register ASIA includes all enterprises carrying on economic activities contributing to gross domestic product at market prices, in the fields of industry, commerce and services[4]. The register is therefore used to collect information about the ENT.

Table 1 - *Variables' description.*

Variable	Description	Type
number_ug_ent	number of UG belonging to the ENT	numeric
employees	number of employees working for the ENT	numeric
turnover	turnover	numeric
s_shares	sum of the shares of the LU belonging to the ENT and sharing with it the same NACE code	numeric
s_shares_diff	sum of the shares of the LU belonging to the ENT and with a NACE code different from the ENT's code	numeric
ls_ent	legal status	categorical
region_ent	region	categorical
ateco5_ent	Categories NACE code (5 digits)	categorical
ateco4_ent	Classes NACE code (4 digits)	categorical
ateco3_ent	Groups NACE code (3 digits)	categorical
ateco2_ent	Divisions NACE code (2 digits)	categorical
section_ent	Sections NACE code (1 digit)	categorical
cladd3_ent (cis)	class of employees: 10-49, 50-249, 250+	categorical

[4] For a complete description of the register the reader may refer to Siqual website accessible at: https://siqual.istat.it/SIQual/visualizza.do?id=7777976.

The most updated version of the register available, i.e. ASIA ENT 2021, includes the ENT active in 2021. The number of units (ENT) is 4462146, while the number of variables is 17.

We considered only some of the available variables to carry out our analysis: *number_ug_ent, employees, turnover*[5]*, ls_ent, region_ent, ateco5_ent*. In addition, by furtherly manipulating some of the variables we determined additional variables such as *s_shares, s_shares_diff, ateco4_ent, ateco3_ent, ateco2_ent, section_ent* and *cladd3_ent*. The variables *ateco5_ent, ateco4_ent, ateco3_ent, ateco2_ent, section* are related to the ENT's economic activity according to the *statistical classification of economic activities in the European Community* (NACE classification). The variables' descriptions and information are summarised and reported in Table 1.
Instead of considering the whole ASIA register, we implemented our study on several subsets of this register. The filtering variables were those related to economic activity.

- ENT working in the C section of the NACE classification ('manufacturing activities');
- ENT working in the M section of the NACE classification ('Professional, Scientific and Technical Activities');
- ENT working in the R section of the NACE classification ('arts, sports, entertainment and recreation activities);
- ENT working in the division codes (NACE 2-digits) of the NACE classification usually survey in Community Innovation Survey (CIS).

The following Section provides a full description of the aims and the current sampling design for the CIS survey.

2.1. CIS Survey

The Statistical Survey of Business Innovation (CIS) aims to collect information on the strategies, behaviours and innovative activities carried out by enterprises. Since 2023, the sample survey is based on a stratified cluster sampling design with equal inclusion probabilities for all population units. The enterprises were stratified, and from each stratum they were selected by simple random sampling. All legal units belonging to enterprises were sampled. In this way, data are collected at LU-level and therefore direct estimates at LU-level are easily obtained; at the same time, given

[5] To handle missing data in this variable, we used the Multivariate Imputation by Chained Equation (mice R package, van Buuren and Groothuis-Oudshoorn, 2011), even though we are aware that there exist many other imputation techniques which are used in Official Statistics.

the LU-ENT structure, it is also possible to obtain estimates at ENT-level by properly adjusting the weights in the calibration process.

The goal of the survey is to collect information on the number of employees, turnover and total innovation spending of enterprises. In particular, the target population consists of all enterprises with more than 10 employees and operating in the following economic activity by NACE Rev. 2 at one-digit (section) level: B, C, D, E, F, G, H, J, K, L, M (excluding division 75). The survey is sample-based for enterprises with 10 to 249 employees and census-based for those with 250 or more employees. The stratification variables are: a) the economic activity by NACE Rev. 2. Stratification by NACE was done at a two-digit (division) level, except for section F; b) the enterprise size according to the number of persons employed (the size-classes used were the following ones: between 10 and 49; between 50 and 249; 250 and more); c) the regional variable. The breakdown of national territory into regions was performed based on the NUTS level 2.

The estimates are provided, according to regulations, for the following 4 domains: a) Estimation domain 1: economic activity by NACE Rev. 2 at two-digit (division) level, except for section F; b) Estimation domain 2: economic activity by NACE Rev. 2 at two-digit (division) level, except for section F cross-tabulated with 3 classes of employees (3 classes: 10-49, 50-249, 250 and above); c) Estimation domain 3: the regional variable cross-tabulated with Macrosector (Eu-core sectors - Other sectors) and 2 classes of employees (2 classes: small and medium-sized enterprises - large enterprises); d) Estimation domain 4: only for Bolzano province, Trento province, and Friuli Venezia Giulia region: economic activity by NACE Rev. 2 at one-digit (section) level cross-tabulated with Eu-core/Others.

A multi-variable and multi-domain sample allocation is used. In particular, the adopted procedure is an application of the Bethel algorithm (Bethel, 1989) . It is an optimum allocation since it aims at minimising survey costs under the constraint that sampling errors (in terms of coefficient of variation, CV) of estimates of each variable of interest and for each domain don’t exceed the given upper bounds assigned to each of them.

3. Methodological Tools

Cluster analysis is used to obtain clusters of ENT and to identify the most influential variables in the partition which can be used as stratification variables: indeed, influential variables in clustering are the variables that discriminate and partition the ENT homogeneously, so it is reasonable to use them to build new strata of ENT. In this way, ENT will be stratified efficiently.

Since our dataset includes both numerical and categorical variables, we applied the K-prototype clustering algorithm (Huang, 1998), allowing mixed-data type as input. The algorithm is implemented in the clustMixType R package (Szepannek, 2018) and it is considered as an extension of the K-Means (McQueen, 1967) and K-Modes (Huang, 1997) algorithm.

To apply cluster analysis to the four subsets of interest, we decided to use a bootstrap-like procedure for handling huge data dimensions. In particular, from each dataset, 500 random samples of 1000 units without replacement have been selected. On each sample, we performed clustering. Practically, on each sample, by letting the number of clusters vary in [2,20], we ran the K-prototype and we selected the optimal number of cluster k* using the Silhouette index (Rousseeuw, 1987). The resulting partition with k* clusters has been evaluated and analysed. To study which variables were the most influencing in the resulting partition, the feature importance of each variable has been computed by using the R function FeatureImpCluster (Pfaffel, 2021), which measures the importance of a variable in terms of misclassification rate relative to the baseline cluster assignment due to a random permutation of feature values. Finally, variables are ranked according to the feature importance score.

This bootstrap-like procedure helped us to deal with high-dimensional datasets, as in our application. However, to validate the results, we implemented the clustering algorithm on the whole CIS survey target population and we assessed the variables' importance. The most influential variables can indeed be considered for a new stratification.

4. Results

For each dataset, Table 2 reports the variables that were shown to have the greatest influence in most of the iterations of the study. Specifically, the percentage distribution of the most influencing variable obtained in each of the 500 iterations has been computed. Then, the two variables which have the two highest percentage frequencies are reported.

Table 2 – *Two most influencing variables in the simulation study. Datasets C, M, R refer to the subsets of ENT of ASIA2021 working in section C, M, R of NACE classification, respectively.*

Dataset	Two most influencing variables	% of occurrences
C	*ls_ent, s_shares*	82.2, 6.0
M	*region_ent, ateco5_ent*	74.6, 21.8
R	*ateco5_ent, region_ent*	60.0, 39.4
univ_CIS	*section_ent, region_ent*	92.0, 8.0

Table 2 indicates that the region and the section, which pertain to economic macro-activity, are the two most influential variables in the dataset associated with the CIS survey. We found that location information is crucial for the two economic sectors that sections M and R identify. The location of the enterprises may have an impact on their economic activities. Additionally, the variable *ateco5_ent,* which refers to a very detailed economic activity description (being the finest economic activity classification), turns out to be a very discriminant variable given that those two sectors are very heterogeneous, including a large number of quite different economic sub-activities. Therefore, in this sector, the factors that have the biggest effects on differentiating amongst ENTs are the locations of the enterprises and the specific economic sectors in which each of them operates.

As mentioned in Section 3, we also used the K-prototype for the whole CIS dataset to properly validate the findings of the previously described study. The results indicate that the ENT in the CIS dataset is divided into 18 clusters, with *region_ent* and *section_ent* having the most influence in the obtained partition. These variables are consistent and support the output of the simulation.

Furthermore, given the peculiar composition of ENT, we divided the dataset into two subgroups to examine the results in more detail: the first subset (*ENT-Mono*) only takes into account ENTs composed of a single LU, while the second subset (*ENT-Pluri*) only takes into account ENTs composed of more than one LUs. Indeed, in ASIA ENT 2021, among the 4.462.146 ENT, only 1.2% is *ENT-Pluri*, and the remaining 98.8% is *ENT-Mono,* while in the CIS dataset, among the 156619 ENT, 15.67% is *ENT-Pluri*, and the remaining 84.43% is *ENT-Mono.*

We decided to implement two separate analyses for *ENT-Mono* and *ENT-Pluri,* as we thought that the structure and characteristics of those two types of ENT were extremely different. The K-prototype algorithm was applied to both subgroups.

Section_ent and *region_ent* are the two most important variables for *ENT-Mono*, as the simulation study also highlights. Rather, the variables with the highest relevance scores for the *ENT-Pluri*, a very small subset of the entire dataset, are *section_ent, ateco2_ent, cladd3_ent*, and *region_ent*. What is therefore underlined is

the potential for a new stratification that is more wide and comprehensive when the ENT is an *ENT-mono* and more precise and specific when the ENT is an *ENT-pluri*.

Therefore, in order to wrap up the experiment, we choose to test the newly proposed stratification, where the strata variable for each ENT i is specified as follows:

$$stratum_i = \begin{cases} region_i \times section_i & if\ the\ ENT\ i\ is\ a\ ENT-mono \\ region_i \times ateco2_i \times cladd3_i & if\ the\ ENT\ i\ is\ a\ ENT-pluri \end{cases} \quad (1)$$

It has to be noticed that we choose to include *ateco2*, which is a subset of the *section* and more accurate in characterising economic activity, in place of the *section* variable for the *ENT-pluri.*

Table 3 – *Median values of estimates' CVs resulting from the samples selected either using the standard allocation or the new one.*

DOM	Variable	CV_alloc_stand	CV_alloc_new
DOM 1	*employees*	0.017	0.064
DOM 2	*employees*	0.016	0.013
DOM 3	*employees*	0.015	0.009
DOM 4	*employees*	0.013	0.035
DOM 1	*turnover*	0.037	0.110
DOM 2	*turnover*	0.045	0.066
DOM 3	*turnover*	0.028	0.022
DOM 4	*turnover*	0.024	0.057

The proposed stratification has been employed to allocate a sample of the same dimension (39000 LU) as the one extracted using the existing stratification, and the optimal allocation in a multi-variable and multi-domain framework has been used[6] (Fasulo et al., 2021; Barcaroli et al., 2023). Compared to the current stratification, there are about 800 fewer new strata (1769 instead of 2541).

Given the new allocation, we derived the estimates of employees and turnover variables and their related CV. A common and extensively used metric of an estimate's mistake in various NSIs is the CV, which is defined as the ratio of the standard deviation to the mean. Publicising estimates with a CV less than 0.33 and classifying as totally dependable those with a CV less than 0.15[7] is standard procedure.

[6] The optimal allocation in a multi-domain, multivariate framework implemented in the R package R2BEAT has been used (see https://barcaroli.github.io/R2BEAT/index).

[7] See Statistics Canada. Sampling error, https://www150.statcan.gc.ca/n1/pub/62f0026m/2010003/section1-eng.htm

More in detail, we implemented a Monte Carlo simulation with 100 iterations: in each iteration, two samples have been selected: the former is extracted using the current allocation, while the latter is selected using the proposed allocation. Given the two samples, the estimates of the target variables and their related CV are computed. The estimates are obtained at the four different domains discussed and presented in Section 2.1.

Table 3 reports the median level of the CVs of the estimates for each combination of target variable and domain of interest.

The estimates derived from the samples extracted following the proposed allocation have median values of the CVs that are marginally higher than those associated with the current allocation in some combination but turn out to be below the acceptance threshold.

To conclude the proposed assessment, two are the most important advantages of the methodological proposal. On the one hand, the new stratification allows to obtain a lower number of strata; on the other hand, the efficiency of the estimates of the target variables for each domain of interest is still preserved.

The reduction in the number of strata turns out to be an important gain in stratified sampling strategy: indeed, one of the main concerns and issues which arise when adopting stratified sampling is related to the fact that after the survey, several strata turned out to be empty, i.e. no-one of the sampled unit responded to the survey; the *empty-strata issue* is a big concern as units belonging to the strata are no-more represented by any respondent. Instead, by considering a lower number of strata, the strata dimension at the population-level (number of population units, N_h) is higher, as well as therefore also the strata dimension at the sample-level (number of sample-allocated units, n_h); as a consequence, the risk of incurring in empty-strata issue is considerably reduced.

Furthermore, because the number of strata is reduced, the sample is less widespread and can be allocated more efficiently.

Finally, considering the editing and imputation phase of the whole statistical survey process, when using hot-deck imputation, if imputation classes exactly match the strata, then the number of potential donors inside each imputation class is high, given that the strata have a higher size and hence a higher number of respondents is obtained in each stratum.

It is worth mentioning that a reduced number of strata allows also to reduce the processing time of the whole allocation, sampling and estimation processes.

5. Final consideration

This work focuses on addressing the use of a new stratification to select a sample of ENT in SBS surveys. More in detail, a clustering approach for determining stratification variables has been proposed: K-prototype clustering algorithm has been applied to several input datasets, obtained through a subset of the ASIA ENT 2021 register, which is a mixed-data type dataframe. The input datasets include ENTs working in different sections of the statistical classification of economic activities in the European Community (NACE) and ENTs included in the target population of the CIS survey conducted by ISTAT. Finally, the most influencing variables in determining the clustering partition have been used as stratification variables.

The analysis of the CIS target population has been deeply discussed; the new proposed stratification has been used to allocate a sample of ENT of the same dimension as the one selected with the current stratification. Through a Monte Carlo simulation, the variables of interest related to the CIS survey (turnover and employees) have been estimated as well as their coefficients of variation. By comparing the results obtained by sampling using the proposed stratification and the ones obtained by sampling using the current stratification two important aspects are revealed: on the one hand, the efficiency of the estimates is preserved; on the other hand, the number of strata has been reduced, and several related advantages have been discussed in the previous Section. Among the others, it's worth recalling the reduction of the risk of incurring in empty-strata issue and the reduction in processing time.

This work aimed to provide some hints to help NSI go through the transition process from considering the Legal Unit as the statistical unit in SBS surveys towards considering the ENT as such. The results are promising and suggest that it may be necessary to think about a new sampling design. However, further investigations are needed.

References

BARCAROLI G., FASULO A., GUANDALINI A., TERRIBILI M.D. 2023. Two Stage Sampling Design and Sample Selection with the R Package R2BEAT. *The R Journal*, Vol. 15, No. 3, pp. 191-213.

BETHEL J. 1989. Sample allocation in multivariate surveys. *Survey Methodology*, Vol. 15, pp. 47 -57

FASULO A., BARCAROLI G., FALORSI S., GUANDALINI A., PAGLIUCA D., TERRIBILI M.D. 2021. R2BEAT: Multistage Sampling Allocation and Sample

Selection *R package version* 1.0.4. url:https://CRAN.R-project.org/package=R2BEAT

HUANG J. 1998. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data mining and knowledge discovery,* Vol. 2, No. 3, pp. 283-304.

HUANG Z. 1997. A fast clustering algorithm to cluster very large categorical data sets in data mining. *DMKD,* Vol. 3, No. 8, pp. 34-39.

MCQUEEN J. B. 1967. Some methods of classification and analysis of multivariate observations. In: *Proc. of 5th Berkeley Symposium on Math. Stat. and Prob*, Vol 2, No. 3, pp.281-297.

PFAFFEL O. 2021. FeatureImpCluster: Feature Importance for Partitional Clustering. *R package version 0.1.5.* url: https://CRAN.Rproject.org/package=FeatureImpCluster.

ROUSSEEUW P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, Vol. 20*,* pp. 53-65.

SZEPANNEK G. 2018. ClustMixType: User-Friendly Clustering of Mixed-Type Data in R. *R Journal*, Vol. 10, No. 2, p. 200.

VAN BUUREN S., ROOTHUIS-OUDSHOORM K 2011. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, Vol. 45, No. 3, pp. 1-67.

Ilaria BOMBELLI, ISTAT, ilaria.bombelli@istat.it
Giorgia SACCO, ISTAT, giorgia.sacco@istat.it
Alessio GUANDALINI, ISTAT, alessio.guandalini@istat.it

SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS
ISSN ASSEGNATO: 0035-6832

Direttore Responsabile: CHIARA GIGLIARANO

Iscrizione della Rivista al Tribunale di Roma del 5 dicembre 1950 N. 1864

Associazione all'Unione Stampa Periodica Italiana

TRIMESTRALE

La copertina è stata ideata e realizzata da Pardini, Apostoli, Maggi p.a.m.@tin.it – *Roma*

Stampato da CLEUP sc
"Coop. Libraria Editrice Università di Padova"
Via G. Belzoni, 118/3 – Padova (Tel. 049/650261)
www.cleup.it

ATTIVITÀ DELLA SOCIETÀ

RIUNIONI SCIENTIFICHE

XXXVII	La mobilità dei fattori produttivi nell'area del Mediterraneo (Palermo, 15-17 giugno 2000).
XXXVIII	Qualità dell'informazione statistica e strategie di programmazione a livello locale (Arcavacata di Rende, 10-12 maggio 2001).
XXXIX	L'Europa in trasformazione (Siena, 20-22 maggio 2002).
XL	Implicazioni demografiche, economiche e sociali dello sviluppo sostenibile (Bari, 15-17 maggio 2003).
XLI	Sviluppo economico e sociale e ulteriori ampliamenti dell'Unione Europea (Torino, 20-22 maggio 2004).
XLII	Sistemi urbani e riorganizzazione del territorio (Lucca, 19-21 maggio 2005).
XLIII	Mobilità delle risorse nel bacino del Mediterraneo e globalizzazione (Palermo, 25-27 maggio 2006).
XLIV	Impresa, lavoro e territorio nel quadro dei processi di localizzazione e trasformazione economica (Teramo 24-26 maggio 2007).
XLV	Geopolitica del Mediterraneo (Bari, 29-31 maggio 2008).
XLVI	Povertà ed esclusione sociale (Firenze 28-30 maggio 2009).
XLVII	Un mondo in movimento: approccio multidisciplinare ai fenomeni migratori (Milano 27-29 maggio 2010).
XLVIII	150 anni di Statistica per lo sviluppo del territorio: 1861-2011. (Roma 26-28 maggio 2011).
XLIX	Mobilità e sviluppo: il ruolo del turismo. (San Benedetto del Tronto, 24-26 maggio 2012).
L	Trasformazioni economiche e sociali agli inizi del terzo millennio: analisi e prospettive (Università Europea di Roma, 29-31 maggio 2013).
LI	Popolazione, sviluppo e ambiente: il caso del Mediterraneo (Università Federico II di Napoli, 29-31 maggio 2014).
LII	Le dinamiche economiche e sociali in tempo di crisi (Università Politecnica delle Marche, 28-30 maggio 2015).
LIII	Mutamento economico e tendenze socio-demografiche tra sfide e opportunità (Università degli Studi Internazionali di Roma, 26-28 maggio 2016).
LIV	Mobilità territoriale, sociale ed economica: modelli e metodi di analisi (Università degli Studi Internazionali di Catania, 25-26 maggio 2017).
LV	Coesione sociale, welfare e sviluppo equo e sostenibile (Università degli Studi dell'Insubria, Varese 24-25 maggio 2018).
LVI	Benessere e Territorio: Metodi e Strategie (Università Politecnica delle Marche, Ascoli Piceno 23-24 maggio 2019).

LVIII Tra marginalità e sviluppo. La sfida della sostenibilità in una prospettiva mediterranea (Università LUMSA, Palermo, 26-27 maggio 2022).

LIX Aspetti economici e sociali dell'invecchiamento demografico (Università degli Studi di Napoli Federico II Napoli, 25-26 maggio 2023).

LX Salute, ambiente e disuguaglianze: istituzioni, imprese e società (Università degli studi di Milano – Bicocca & LIUC - Università Cattaneo, Milano & Castellanza (VA), 22-24 maggio 2024).