Genus

ORIGINAL ARTICLE Open Access

Student mobility in Southern Italy: An empirical analysis of preferential patterns

Vincenzo Giuseppe Genova[1*], Gabriele Ruiu[2], Massimo Attanasio[1], Matteo Ermacora[2] and Marco Breschi[2]

*Correspondence: vincenzogiuseppe.genova@unipa.it

[1] Department of Economics, Business, and Statistics, University of Palermo, Palermo, Italy
[2] Department of Economics and Business, University of Sassari, Sassari, Italy

Abstract

This paper presents an empirical analysis aimed at identifying chain migration patterns among university students from Sicily, Sardinia, and Apulia. Utilizing data from the national archive of university students enrolled in an Italian university from 2008 to 2017, we construct origin–destination matrixes and calculate residuals from a log-linear model to detect key student routes. Positive residuals between specific origin–destination pairs that persist over time can suggest significant attraction and potential chain migration. The results support a "large to large" and "small to small" city mobility model, particularly for Sicily and Sardinia. Students from Palermo, Catania, Messina, and Cagliari show strong connections with large cities such as Rome and Milan. Furthermore, results show historical links between Sardinia and Tuscany date back to the 1950s, and the enduring connections between Apulian students and universities in Emilia-Romagna, Veneto, and Lombardy have been evident since the 1970s. In Sicily, mobility patterns towards the Polytechnic of Turin began in the 1970s and persist today. These patterns underscore the influence of prior migrants on the mobility decisions of university students in Southern Italy, highlighting a potential chain migration effect.

Keywords: Student mobility, Chain migration, Cluster analysis, University students

Introduction

Universities play a fundamental role for territories contributing to the creation of human capital, being a major employer, livening up the cultural life of local communities. Moreover, research has demonstrated that these institutions can act as drivers of innovation and economic growth (Wolfe, 2005).[1]

Although Italian universities are part of the economic system of their respective regions, securing the necessary resources to carry out their activities remains a pressing concern. To this end, universities must strive to increase student enrolment, as a significant portion of their funding (the Fondo Finanziamento Ordinario or FFO) is tied to

[1] See also the European Union recommendations for connecting university to regional development: https://ec.europa.eu/regional_policy/sources/docgener/presenta/universities2011/universities2011_en.pdf

Springer Open

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

student numbers. Additionally, larger student populations generate greater tuition revenues, and foreign students can provide supplementary funding.[2]

In Italy, in the last twenty years, there have been significant changes in terms of student flows and mobility. Student enrolment has decreased from 2008 to 2015 with consistent recovery in the last five to six years. It has always increased since 2008 with a slight reduction in 2017. Attanasio and Enea (2019) highlighted that university mobility flows in Italy are unidirectional from the South and Islands to the Center-North: 10.4% (15.9%) of students from the South (Islands) were enrolled in a Northern university and 11% (8.8%) of them moved towards the Center in 2014, while movements in the opposite direction are non-existent. Similar patterns have been observed during the transition from bachelor's to master's programs (Genova et al., 2021) as well as in Ph.D. programs (Ruiu et al., 2019; Tocchioni & Petrucci, 2021).

This paper aims to explore preferential students' mobility patterns in Italy aiming at detecting the eventual presence of students' chain migration in some mobility routes. This phenomenon is similar to the usual chain migration—where movers share information with primary social contacts about labour market opportunities and/or quality of life at destination.

Our main research question is:

Is there evidence of preferential mobility patterns over time among Southern students, which can be interpreted as a chain mobility process?

To answer our research question, we utilized students' career data referred to three Southern Italian regions focusing on the student routes (outside their own region) from 2008 to 2017.

In Italy, students' mobility, based on public information, has been widely dealt by statisticians (Attanasio & Priulla, 2020), by economists (Aina et al., 2022), and by sociologists (Ballarino et al., 2022). Migration processes, including student mobility, are influenced by both public information (such as socio-economic conditions, job market prospects, and university rankings) and personal factors (such as family and friendship ties, individual motivations, and so on).

The novelty of this work lies in introducing a quantitative approach that attempts to analyse student migration using origin–destination tables, enabling the identification of preferential mobility patterns. Specifically, we construct origin–destination matrixes and calculate residuals from a log-linear model to detect key student routes. Positive residuals between specific origin–destination pairs that persist over time can suggest significant attraction among the areas of origin and the universities of destination, suggesting a potential chain migration. This quantitative approach provides a first step for further qualitative research to unveil and underscore social capital's importance in shaping student mobility decisions. Indeed, one of the limitations lies in the number of years considered in our analysis—a greater number of years could have facilitated the identification of mobility patterns among students that could be considered migratory chains.

The work is organized as follows: in "Definition and historical origin of students' mobility routes in Italy" section, there is a brief review of chain migration literature and an attempt to offer a historical reconstruction of student mobility pathways starting from the post World War II period; in "Data and method" section, we illustrate the data

[2] See Fadda et al. (2022) for details about the system of public funding of Italian Universities.

and the statistical approaches to analyze the chain migration; in "Results" section, we present the results and some comments. Finally, we offer some conclusions.

Definition and historical origin of students' mobility routes in Italy

Chain migration and student mobility

In this work, we are trying to detect some chain migration routes to analyze students' mobility in Italy. Chain migration is based on the migration network theory that explains how migrants create and maintain social ties facilitating further migration over time. Mabogunje (1970) argued the importance of a feedback mechanism between migrants at destination and fellow villagers at the point of origin. Indeed, positive information transmitted to the place of origin encourages further migration and establishes structured mobility patterns between the country of origin and destination (Mabogunje, 1970).

Furthermore, these migrant networks reduce the economic, social, and psychological costs related to migration, and they increase the probability of moving for the next generation, of whatever social origin (Massey, 1990).

Those mobility processes have been highlighted by several studies conducted in different destination countries and considering various ethnic groups. MacDonald and MacDonald (1964, 1970) analyze Italian migration to the U.S. and in Australia, Palloni et al. (2001) and Liu et al. (1991) consider the emigration of, respectively, Mexicans and Filipinos in the U.S. Böcker, 1994 studies the migratory chain of the Turkish community in the Netherlands, while Eurenius (2020) emphasises the role of chains in Swedish emigration to the United States in the late nineteenth century. Reyneri (1998) investigates the case of immigrants employed in the cheap labor market in Italy, pointing out the fundamental role of the chain migration.

Again, the prerequisite for a chain migration is the existence of social capital within a community, that is, the existence of a set of relational resources—within a group—aimed at providing concrete help to newcomers (Haug, 2008).

In this work, we pay attention to a particular type of migration chain: the migration chain process of university students. It certainly has characteristics quite different from those of traditional migratory chains for various reasons. The first one is the limited temporal perspective of the students' mobility flows, which can eventually produce a turnover of the community members. Indeed, in the case of students, the term mobility is often preferred in literature. Person and Rosenbaum (2006) use the term "chain enrolment" to identify the cumulative process in student mobility. Therefore, keeping in mind the differences between chain migration and student mobility, to better remark the role of communities in mobility choices, we will use the term student chain migration in this paper too.

There are very few studies on this topic in the international literature, and these are essentially based on qualitative methods (Brooks & Waters, 2010; Pérez & McDonough, 2008; Person & Rosenbuam, 2006). The analysis carried out by Pérez and McDonough (2008) relies, for instance, on semi-structured interviews collected in a non-representative sample of students. They show that the presence of members of the Latinos' community in the U.S. plays a role in the Latinos college choices. Similarly, Brooks and Waters (2010) show, through the same technique, the importance of social/family ties in the choice of studying in the United Kingdom for

a small group of 83 international students. Person and Rosenbaum's work combines semi-structured interviews with a survey conducted on students from 14 colleges. They conclude the presence of friendship/kinship ties at the destination university increases the likelihood of enrolment in that specific university, especially for students from South American countries.

To the best of our knowledge, in literature there are no quantitative techniques devoted to the students' chain migration. The limitation of previous studies on student chain migration lies in the small size of the samples and in the temporal span. As aforementioned, they are frequently limited to a single community within a single university in a specific year. This contribution, spanning over 10 years, aims to propose a statistical approach that can 'discover' students' mobility patterns that could be a first step to point out students' chain migration. However, it is not our scope to determine what are the causes of the mobility, because our database is administrative, comprehending only university careers.

The origin of mobility routes from the 1950s to the 1990s

To better understand the recent students' mobility, we need to look back to the past mobility trends in the analyzed areas (Sardinia, Sicily, and Apulia).[3] Indeed, the topic of university student mobility in Italy during the first Republican era (1946–1994) has not been widely covered in the literature. Among the reasons of this under coverage are the lack of official micro-level data after the liberalization of university access (1969) and the different ways of collecting data (1953–1968 by provinces, then by regions; Cammelli, 2000).[4] However, some information could be obtained from the statistics produced by Italian National Institute of Statistics—Istat—(1953–1984) and by the Italian Ministry of Education and University (MIUR). Due to the historical imbalance in the territorial distribution of Italian universities since the unification of Italy (1861), the main migratory flow was represented by students who left the South and Islands and moved toward the universities of the Centre-North regions. From 1950 to 1990, between 50 and 65% of the internal student mobility followed the south-north direction. This flow increased during the so-called Italian "economic miracle" (1958–1963) and the following decades. Just in 1973 oil shock abruptly stopped the process of regional economic convergence (Viesti et al., 2011). During the Fifties and the Sixties, the students involved in migrations were mainly males and belonged to middle and upper class; male students attended medicine, law, pharmacy, and engineering, while female students attended humanities and to minor extent natural sciences and biology. While wealthy students could freely choose courses and city sites of university, low-income students could deal with long-range migrations when: i) they had scholarship; ii) relatives and kins could offer support in the destination cities; iii) they were hosted in colleges for free, many of them related to the Catholic Church (Milan, Pavia, and

[3] We will explain the choice of these 3 regions in Sect. "Data"

[4] The 1969 university reform in Italy notably liberalized access to higher education, dismantling previous restrictions and opening doors for a wider spectrum of students to pursue university studies. Before 1969, only "elite" students from "licei" (lyceum) could enroll in university, while students coming from technical (and vocational) high schools were excluded.

Table 1 First year enrolled students. % Stayers /movers 1955–1967

	1955–56		1964–65		1967–68	
	Stayers	**Movers**	**Stayers**	**Movers**	**Stayers**	**Movers**
Siciliy	92.3	7.7	92.2	7.8	91.8	8.2
Sardinia	91.0	8.0	85.5	14.5	84.7	15.3
Apulia	75.8	24.2	85.0	15.0	82.2	17.8

Padua). For many low-income students, option (ii) was predominant. It must be remarked that the migrations during the Fifties and Sixties had an unprecedented role as they paved the way for students' routes and contributed to create regional students' communities in the Centre and in Northern regions. The official statistical surveys on students enrolled in first year (Istat 1955–56, 1964–65, 1967–68) of Sicily, Apulia, and Sardinia show: a) the general increase of the enrollments; b) the increase of the movers (Table 1); c) the formation of migratory couples ("city-to city" links).

The different availability of universities and degree courses (such as STEM – Science, Technology, Engineering, and Mathematics – courses) in Sardinia and Apulia also played a role in mobility decisions. Official data by provinces in 1955–1967 show "regional" migratory routes and patterns: A) Sicily: *Agrigento:* Roma, Milano, Turin, Florence; *Caltanissetta:* Rome, Milan, Florence, Turin; *Messina:* Rome, Milan, Turin; *Ragusa and Syracuse:* Rome, Milan, Turin, Florence, Padua; B) Sardinia: *Cagliari:* Rome, Florence, Pisa; *Nuoro:* Rome, Turin, Pisa, Florence, Milan; *Sassari:* Rome, Milan, Pisa, Florence. C) Apulia: *Bari:* Naples, Rome, Milan; *Brindisi and Foggia:* Naples, Bologna, Rome, Milan, Florence, Pisa, Padua; *Lecce and Taranto:* Naples, Bologna, Rome, Milan, Turin, Florence, Pisa, Padua. Due to the easier access to the national transportation system, Apulia shows a wider range of destinations. In this phase we can see some processes of route consolidations. For instance, between 1960 and 1965 the Sardinian students in the University of Pisa increased four times. Another case was Rome; in 1967–68, about half (42%) of the whole movers from islands and southern regions attended the University of Rome "La Sapienza". Tables 4 and 5 (in the appendix) report the percentage of first year students moving from the Apulian, Sardinian and Sicilian provinces to each regional destinations in the academic years 1955/56 and 1967/68.[5]

The general increase of enrollments contributed to strengthen routes and destinations during Seventies and Eighties; qualitative sources (memoirs, testimonies, diaries) suggest that, after the 1968 students movement, 25% of the mobility choices were conditioned by friendship (Formes 1970, p. 48), sometimes on political militancy and the decision often was made within the peer group or due to a chain migration based on the presence of others students or siblings already moved in Rome, Bologna, Turin, Pisa or Milan (Giocondo, 2022, p. 126; Atti parlamentari 1979, p. 229). In this perspective, the local community in the other cities supported the newcomers. In this framework, students from Sicily and Apulia preferred Turin and Bologna (STEM courses; law and

[5] Note that in Table 4 and 5 it was not possible to reconstruct the AreaOri given that data from historical sources were available only at the provincial level.

medicine, respectively); in the period 1971–1977 Apulia students in Bologna passed from being 918 to 2.444 (+166.2%), representing the 4.1% of the whole students at the university. In the 1980s, they travelled along the Adriatic route, and Bologna was preferred over Rome and Naples. The "Sardinian community" in Pisa formed during the previous decades was strengthened by the institution of computer science degree in mid Eighties; in 1983–1989 Sardinian students passed from 312 to 548. Overall, the official data available for the beginning of the Nineties (MURST, 1991) confirmed aforementioned routes; in order of importance: (A) Sicily: Tuscany, Latium, Lombardy, Piedmont; (B) Sardinia: Tuscany, Latium, Lombardy; (C) Apulia: Emilia, Latium, Lombardy, Tuscany, Piedmont. Beyond the case of Roma and the so-called "Industrial Triangle" (Milan, Turin, Genua), data show that movers from islands and southern regions strengthen their presence in Tuscany, Emilia, Veneto and contributed to the growth of Pisa, Firenze, Bologna and Padua universities.

Beyond the search of prestigious qualifications and the aspiration to live in big and modern cities, our historical analysis suggests that the consolidation of a community of friends and relatives in the destination areas played a role in establishing mobility routes.

Data and method

Data

The data used in this study were obtained from the Ministry of Education, University and Research (MIUR) database, which contains longitudinal information on the careers of all Italian university students from 2008 to 2019. In our analysis, we excluded students enrolled in online degree programs because they have different characteristics compared to other students. Additionally, we excluded students in healthcare area because their choice of university is strongly influenced by the availability of spots and are based on a national competition, leading to mostly forced mobility.

The geographical characteristics of Italy are a pivotal aspect of our analysis, as highlighted in Fig. 1. Sicily and Sardinia, as island regions, share a distinct insularity that significantly constrains mobility beyond their confines, apart from the Strait of Messina area. In contrast, students originating from Apulia can commute to Campania and Lazio on a weekly basis. On the other hand, preferential routes for Apulian students towards Emilia-Romagna and Veneto are facilitated by historical transportation infrastructure, particularly the Adriatic railway.

Moreover, these three regions have quite similar GDP per capita with some differences in the transition rates to university (47.7% for Sicily, 44.5% for Apulia, and 45% for Sardinia). The three regions exhibit very different average mobility rates, with high intra-regional variability (Sicily and Apulia) and low variability (Sardinia), as well as differences in the territorial allocation of the universities. Sicily is distinguished by its three "old" universities—Palermo, Messina, and Catania. In contrast, Apulia houses the historical University of Bari and Politecnico of Bari, both in Bari, supplemented by a smaller university in the southern part of the region. Sardinia, instead, features two historical universities, one in the north and another in the southern area of the island.

The other regions of Southern Italy were not considered due to their unique characteristics. Campania, for instance, is home to a large university, Federico II, and serves as a historical attraction center for other southern regions, with several areas

Fig. 1 Italy's Regions. Yellow represents the selected origins, while blue represents the destinations

near Rome. On the other hand, Calabria and Basilicata have a tradition of university mobility mostly towards Latium and Campania, given that their universities were established within the last 30 years and do not offer comprehensive degree courses, particularly at the master's level.

Moreover, Molise, despite having a university that is diffused throughout its territory, is geographically nestled between three regions—Latium, Campania, and Abruzzo. This unique position results in a significant part of the mobility flow being due to daily commuting. For example, Isernia, the region's second-largest city, is closer to Cassino (home of a university in Latium) than to Campobasso, the capital city of the region. Cassino attracts a portion of the students in Economics and Business Sciences and Engineering, courses not offered or only offered in Campobasso by the University of Molise. Lastly, the situation in Abruzzo is unique due to the memorandum established by MIUR and the University of Aquila (the region's oldest university) following

the 2009 earthquake. This memorandum allowed for the abolishment of tuition fees since the academic year 20014/15, rendering Abruzzo incomparable with other southern regions in this context.

Given these distinctive features of Italy, our analysis is focused on students who earned their high school diplomas in Sicily, Sardinia, and Apulia and subsequently enrolled in degree programs in Piedmont, Lombardy, Emilia-Romagna, Veneto, Tuscany, and Latium (Fig. 1) from 2008 to 2017.

The initial constraint placed on the construction of the areas of origin ("Determination of areas of origin" section) to make them "useful" for research purposes, particularly in sparsely populated areas of Sardinia, created a trade-off with the size of the origin areas (which should represent a more homogeneous area smaller than the provinces) and the number of students. This led to the aggregation of the cohorts into two groups, each spanning a period of five years, 2008–2012 and 2013–2017. Although a pooled analysis might seem preferable given the sample size considerations, separating the data into distinct periods helps to identify the persistence over time of potential preferential mobility patterns that may only become apparent when comparing consistent results across different time intervals ("Application of log-linear models" section). By isolating each time interval, we can track how stable some mobility routes are over time, revealing whether there are persistent connections between "specific" origin and "specific" destination areas. Those persistent connections may suggest chain migration effects. Indeed, the aggregation in five-year periods mitigates the limitations posed by sample size, while still offering enough temporal depth to observe migration behaviours.

The inclusion of a third five-year period would have been useful in better illustrating the migration routes that emerged during the decade in question. However, this was not feasible due to the impact of the COVID-19 pandemic. During this time, many Italian universities transitioned to online education, making it challenging to distinguish whether the observed mobility was genuine or artificial.

Methods

Determination of areas of origin

The first step of our analysis is to determine the origin areas (AreaOris) of the students under analysis. In the construction of these areas, we have made a "strong" assumption regarding the communication patterns among students within each AreaOris. Specifically, we have assumed that students communicate exclusively within their own AreaOris, and not with students from other AreaOris. The AreaOris are formed by multiple municipalities that gravitate around a central hub municipality, where at least one high school is located. This study is based on the premise that the university choices made by students are primarily influenced by their AreaOris and the relationships that they have with their high school and local communities. To construct the AreaOris, we utilized the data from the student cohort of 2008/09.

The approach taken for the construction of the AreaOris was consistent across all three regions analyzed, namely Sicily, Sardinia, and Apulia. The determination of the AreaOris can be summarised as follows:

1. We construct an Origin–Destination matrix M(i,j), where i represents the municipality of residence of the students and j is the municipality of the high school attended by the students (Table 2);
2. Within the list of J destinations, we select the municipalities where schools have at least 200 students who will be future freshmen at university. These chosen municipalities are the hubs used as starting point for the determination of the AreaOris;
3. If i is a hub municipality, we simply assign students in municipality i to hub j with $i=j$. (e.g., in Table 2, the students in cells (Palermo,j) ($j=1,\ldots,n$) are assigned to the Palermo hub.
4. If a municipality i is not a hub, we identify j^*, with $i \neq j^*$, as the municipality with the highest number of high school graduates living in i;

   a. If j^* is hub $\Rightarrow$ we attribute the i-th municipality to j^* (refer to Table 2: since the maximum value of the cells (Carini,j) is 50, the hub of Carini is Palermo);
   b. If j* is not a hub $\Rightarrow$ we assign students of the i-th municipality to the j-th hub municipality that is closest in terms of physical distance (e.g., refer to Table 2, for assigning students in the cells (Capaci,j) ($j=1,\ldots,n$) to the geographically closest hub, which is Palermo);

5. The aforementioned steps assigned students living in k non-hub municipalities (C_1, C_2,..., C_K) to hub municipalities. The AreaOris of a hub j is thereby defined as:

$$AreaOri(hub_j) = hub_j \cup \left\{ \bigcup_{k=1}^{K} C_k \right\}$$

Table 2 provides an exemplar of the construction of AreaOris in Sicily. One notable case is Carini, which represents the hub for students from Capaci and Isola, but the total number of students in Carini is lower than the threshold set. As a result, we assign Carini to the nearest hub, in this case, Palermo. This is due to the relatively small number of students from Carini makes it difficult to classify it as a hub municipality based on

Table 2 Examplar of construction of the AreaOris

Origin	Destination							
	Alcamo	**Carini**	**…**	**Palermo**	**…**	**Trapani**	**…**	**Sciacca**
Carini	0	30	…	50	…	2	…	10
Palermo	2	10	…	400	…	0	…	0
Capaci	3	50	…	20	…	0	…	5
Isola	2	30	…	40	…	2	…	0
Balestrate	70	3	…	10	…	50	…	3
…	…	…	…	…	…	…	…	…
Buseto Palizzolo	10	50	…	30	…	30	…	5
Total	220	100	…	1000	…	300	…	320

our pre-established criteria. Therefore, we had to reassign these students to the closest and more appropriate hub municipality for the construction of the AreaOris.

The assumption underlying the chain process is that it occurs exclusively among individuals who are part of the same AreaOris. This assumption appears to be consistent with the low levels of generalized trust documented in Southern Italian regions (Istat 2017)[4]. It is reasonable to suppose that the social connections necessary to form a chain process are established within smaller communities through daily interactions, and that individuals outside these circles of acquaintances may be considered untrustworthy. Moreover, as Pearson and Rosenbaum (2006) suggest, students who receive information from family members and close friends (strong ties) may be less inclined to rely on public information from official channels or acquaintances (weak ties) when deciding which university to attend.

Application of log-linear models

To discover preferential students' mobility pattern, we applied a log-linear model on the Origin/Destination matrixes for each region (as explained in "Data" section, we consider some central-northern regions as possible destinations) and each time period (2008–2012; 2013–2017).

$$\log\mu_{ij} = \lambda + \lambda_i^{AreeOri} + \lambda_j^{RegDest} \tag{1}$$

where:

- μ_{ij} is the expected number of students coming from the i-th AreaOris of origin and going to the j-th destination region (RegDest);
- $\lambda_i^{AreeOri}$ and $\lambda_j^{RegDest}$ are the row and column effects for the Origin/Destination matrix, respectively.

In our model, $\lambda_i^{AreeOris}$ and $\lambda_j^{RegDest}$ represent the "repulsiveness" of the i-th AreaOris and the "attractiveness" of the j-th destination region, respectively. Deviations from the expected values provided in Eq. (1) that are statistically significant provide information on special "i-j cells" that are far from independence.

Therefore, standardized residuals are used to measure how much the observed flows deviate from those estimated according to the model in Eq. (1):

$$\varepsilon_{ij} = \frac{n_{ij} - \mu_{ij}}{\sqrt{\mu_{ij}(1 - p_{i+})(1 - p_{+j})}} \tag{2}$$

where:

- n_{ij} are the flows observed from the AreaOris i towards the destination region j;
- μ_{ij} are the flows from i to j estimated from Eq. (1);
- p_{i+} and p_{+j} are the marginal proportions of the i-th row and the j-th column, respectively.

For significance level $\alpha = 0.05$, all the residuals $|\varepsilon_{ij}| > 2$ correspond to pairs (i, j) with an attraction level ($\varepsilon_{ij} > 2$) and a level of repulsion ($\varepsilon_{ij} < -2$) higher/lower than expected.

The matrix of residuals ε_{ij} is calculated for t_1 (2008–2012) and t_2 (2013–2017). If both $\varepsilon_{i'j't_1} > 2$ and $\varepsilon_{i'j't_2} > 2$, then we hypothesize a migratory chain effect, i.e., the origin i' and destination j' present flows of students leaving i' and entering j' significantly greater than those expected over time.

Furthermore, to analyze the presence of a possible migratory chain taking into account gender and disciplinary area, divided into STEM (*Science, Technology, Engineering and Mathematics*) and no STEM, as determinants of student mobility, we replaced the model in (1), by adding the gender covariate (Eq. 3) and the field of study covariate (Eq. 4). In this way, we obtained—for each time period under analysis—the following models:

$$log\mu_{ijk} = \lambda + \lambda_i^{AreeOri} + \lambda_j^{RegDest} + \lambda_k^{Gender} \tag{3}$$

$$log\mu_{ijk} = \lambda + \lambda_i^{AreeOri} + \lambda_j^{RegDest} + \lambda_k^{STEM} \tag{4}$$

These models provide insight of the presence of preferential mobility patterns over the two 5-year periods controlling for gender or field of study.

In other words, our empirical strategy allows us to isolate a residual factor "associated" to mobility choices that is not attributable to either the push or pull factors related to the AreaOri and the destination, respectively. In our interpretation, positive residuals show "attraction" between the *i-th* AreaOri and the *j-th* destination region, while repetitive positive residuals over time show an eventual chain effect. Although other tools (field analysis through questionnaires and/or semi-structured interviews) are necessary to delve into the complex motivations for choosing a university location, our statistical approach can be seen as a simple method to understand on which mobility routes to conduct further investigations.

Clustering process and graphical representation

As a final step in our analysis, we sought to better understand the temporal patterns of mobility with respect to gender and field of study by applying a hierarchical clustering algorithm to the pairs (AreaOris (t1), RegDest (t1); AreaOris (t2), RegDest (t2)). This allowed us to isolate the (AreaOris, RegDest) pairs based on the five-year period to evaluate the persistence of attraction/repulsion with respect to gender and field of study.

We utilized the Complete Linkage clustering algorithm, which constructs a tree whose branches represent the units (pairs) grouped by similarity into one or more roots. Increasing levels of root depth correspond to larger levels of aggregation. Compared to other methods such as Single Linkage or Average Linkage, Complete Linkage provides a better separation of groups by avoiding the chaining phenomenon, typical of clustering algorithms based on nearest-neighbor distance. By utilizing the farthest-neighbor distance, Complete Linkage creates more compact and homogeneous clusters (Anderberg, 2014; Everitt et al., 2011).

Results

Description of the outflows

Fig. 2a, b, c illustrates the percentages of students who moved from Sicily, Sardinia, and Apulia to universities outside their regions during 2008–2012 and 2013–2017.

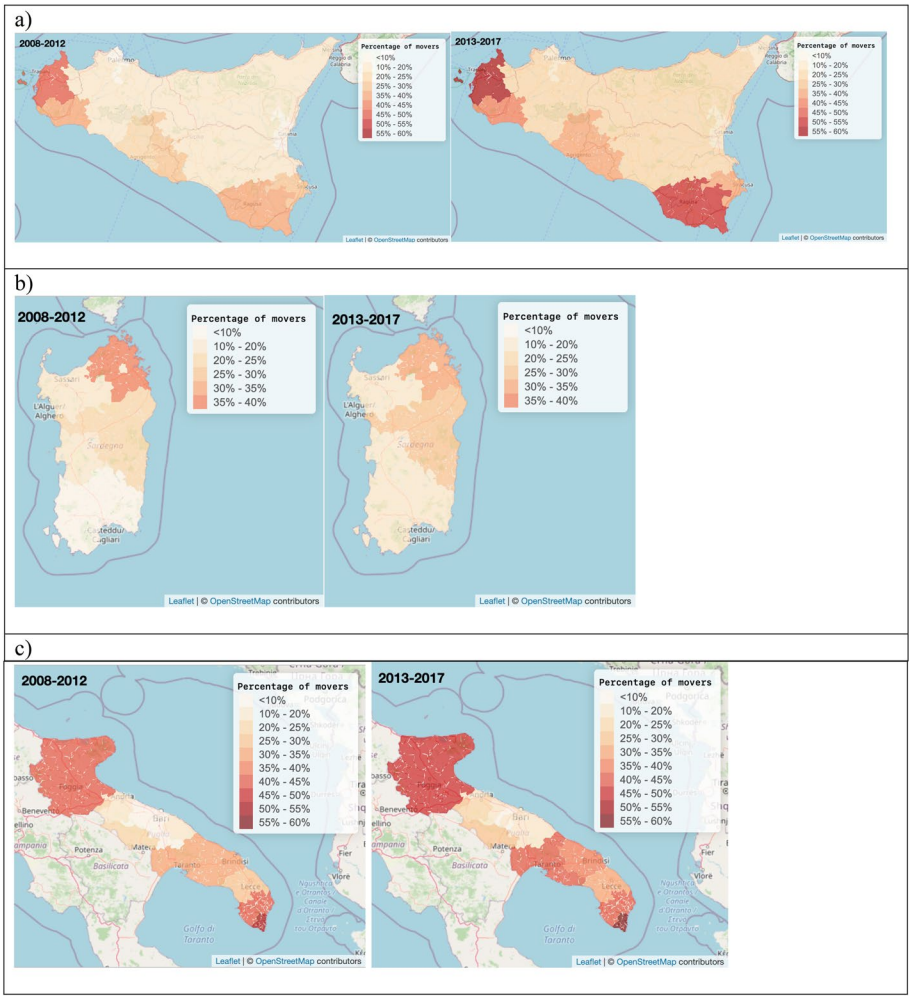

Fig. 2 Percentages of movers in Sicily (**a**), Sardinia (**b**), and Apulia (**c**) by five-year intervals

In Sicily (panel a), the areas with the biggest outflows are Trapani, Castelvetrano, Ragusa, Syracuse, and Agrigento in both periods—2008–2012, 2013–2017, with a peak in the second period, likely due to the economic recovery.[6] On the other hand, as expected, the main Sicilian cities—Palermo, Catania, and Messina, which are sites of a university—have the lowest percentage of outgoing students.

Considering Sardinia, the AreaOris with the largest outflow are Olbia-Tempio and Nuoro all over the time intervals considered in this analysis. It is worth noticing that such a phenomenon decelerates (accentuates) for Olbia-Tempio (Nuoro) in the second intervals. Furthermore, percentages of movers are significantly lower than the ones observed in Sicily, albeit remarkable. Indeed, approximately 40% of graduates from Olbia-Tempio have left the island, while the two main Sardinian cities, Sassari and Cagliari, have lower percentages of movers—since they are site of a university.

[6] According to the Eurostat Business Cycle Clock, Italy has experienced two recessions in the period under analysis, the former, which has registered the largest drop in GDP, from the first quarter of 2008 to the second quarter of 2009; the latter from the second quarter of 2011 to the second quarter of 2013.

Focusing on Apulia, it is worth to noticing that percentages of movers are higher than the Sicilian ones (see Table 3 in appendix). Considering the AreeOri of Foggia, Maglie, Casarano and Tricase, one out of two students decided to leave their hometowns to study outside in 2013–17. In Apulia, the capital city appears to have the lowest rates of mobility. However, in contrast to Sicily, relatively large cities such as Taranto, Foggia, and Lecce (which have small universities, while Taranto hosts a branch of the University of Bari) have a high percentage of students leaving the region (see Table 3 in appendix).

Analysis of origin–destination flows

Figures 3, 4, 5 6, 7 and 8 show the results of the cluster analysis on the residuals—controlling by gender and field of study, Eqs. (3) and (4)—for Sicily, Sardinia and Apulia, respectively. The red colour indicates the clusters where the standardized residuals are >2, while the blue one refers to values $\varepsilon_{ij} < -2$. Moreover, the more intense the colour, the larger the residuals.

The results for Sicily (see Figs. 3 and 4) may suggest the presence of some student migration chains. In particular, if we consider the humanities (Fig. 4), the residuals are exceptionally high in both five-year periods (thus highlighting a number of students permanently higher than what might be expected on the basis of the pull and push

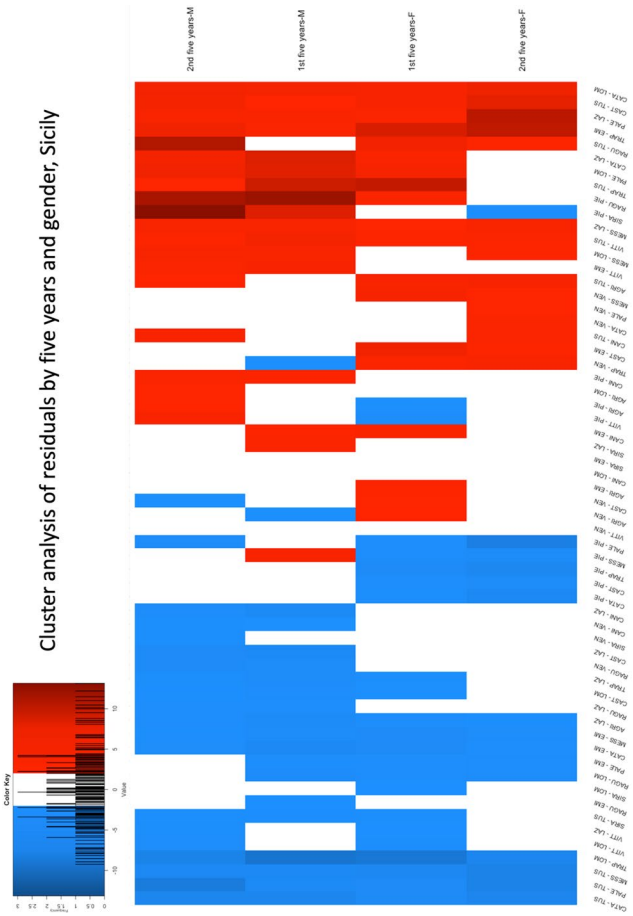

Fig. 3 Cluster analysis of residuals by 5-years for gender, Sicily

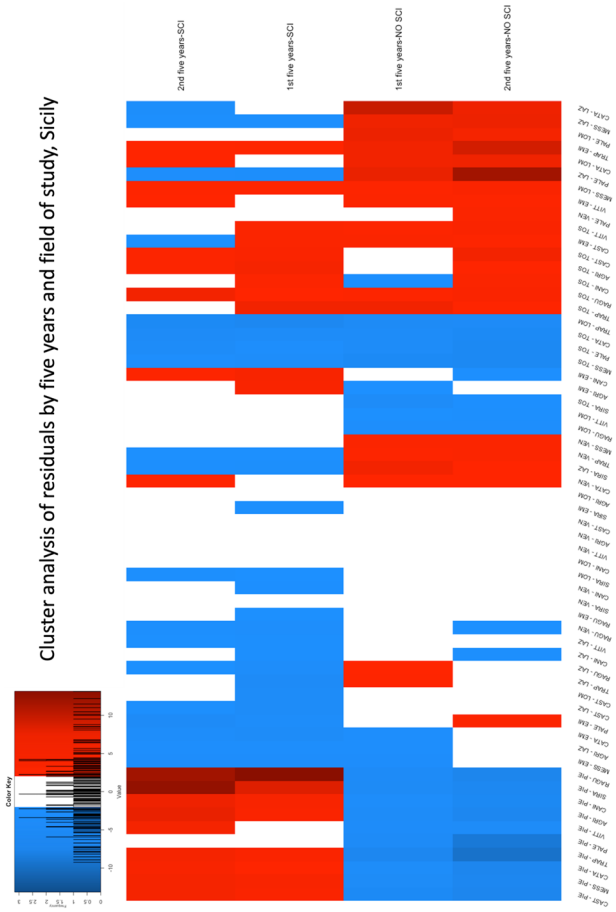

Fig. 4 Cluster analysis of residuals by 5-years for field of study, Sicily

effects) in the following pairs: Palermo-Latium, Catania-Latium, Messina-Latium, Palermo-Lombardy, Messina-Lombardy, Catania-Lombardy, Trapani-Tuscany, Trapani-Emilia-Romagna, Vittoria-Emilia-Romagna, Agrigento-Tuscany, Vittoria-Tuscany, Ragusa-Tuscany, Castelvetrano- Emilia Romagna. It is worth noticing that different patterns of mobility emerge between main Sicilian cities (Palermo, Catania, and Messina) and minor Sicilian cities. Students coming from Palermo, Catania, and Messina are more oriented toward the universities located in big cities such as Rome, Milan and Turin. On the other hand, the second group appears to be moving towards smaller destinations. One possible explanation for this observation is that moving to a big city may have higher psychological costs for a student coming from a relatively small background, especially if they do not have a social network in the destination. This might not be the case for students coming from larger cities within their regional context. After all, the choice to move to study still represents a provisional choice of duration presumably equal to the years of the university, so although the job market of a large city like Milan or Rome could be more attractive for a young person from a small city, this factor affects his/her choices especially in the context of a final (or semi-final) migratory movement for work, but it is relatively less important in the choice of the University.

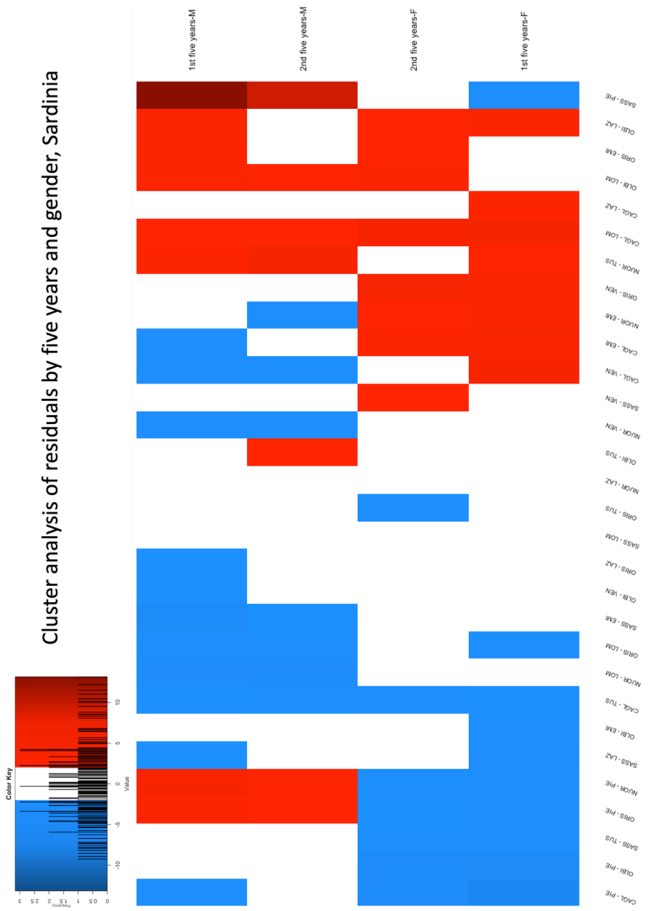

Fig. 5 Cluster analysis of residuals by 5-years for gender, Sardinia

Our data do not allow us to test this hypothesis. However, we believe it should be interesting to dig deeper into the motivations behind the observed different mobility patterns between big and small cities in future research.[7]

It is worth noticing that all the mobility patterns towards Piedmont, possibly affected by the chain migration too, refer mainly to the STEM degrees. In this region, the Polytechnic of Turin (a university mainly devoted to the Engineering degree courses) have represented and represents an important basin of attraction, especially for students from small Sicilian cities (strong positive residuals are not evident for big cities such as Palermo and Catania). Moreover, these flows have been increasing along the whole decade, so it is possible to claim a chain migration effect. It is worthwhile to remark that this link between Sicilian provinces and Polytechnic of Turin started in the 1970s and it seems to be still ongoing.

On the contrary, for the non-STEM courses, Piedmont seems to have a number of students lower than expected. Due to persistent problems of gender stereotypes (Cervia & Biancheri, 2017), STEM subjects continue to be the prerogative of men and therefore, mobility patterns in which the destination is Piedmont are mainly composed of male students.

[7] In this case a qualitative study with semi-structured interviews may be the most appropriated instrument.

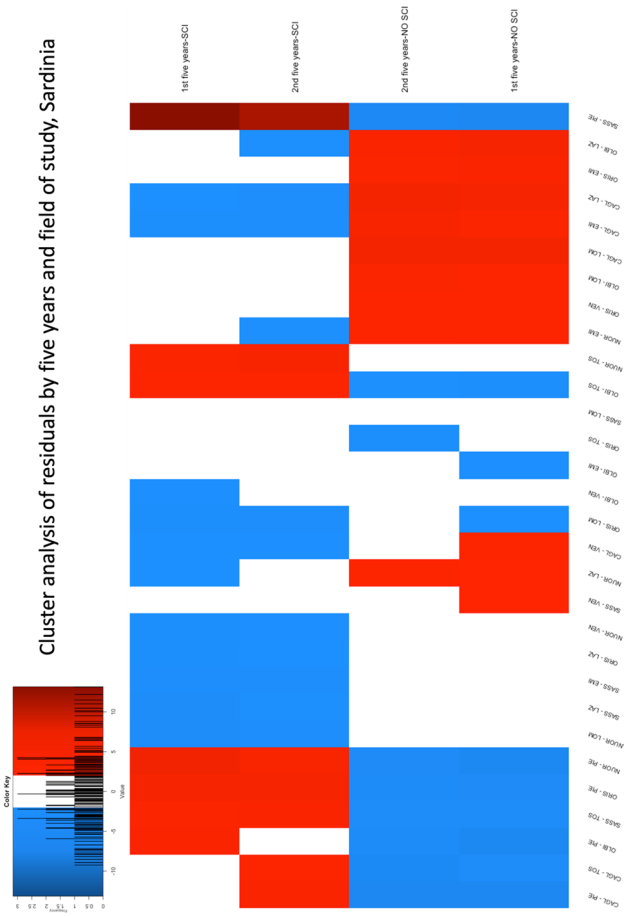

Fig. 6 Cluster analysis of residuals by 5-years for field of study, Sardinia

As for Sardinia (Figs. 5 and 6), the strong link between Sassari and Piedmont for hard science subjects is evident in both periods. The link between Piedmont with Nuoro and Oristano is milder but still significant. In the field of scientific courses, Tuscany also seems to be linked by a chain effect to the territories of Sassari, Nuoro, and Olbia-Tempio. As noted in the historical analysis of this work, the link between Sardinia and Tuscany has a long-lasting origin—since the 1950s. This offers further support to the idea that with our statistical tools, we are capturing a fundamental characteristic of the cumulative process: persistence in time and space. The network at the destination may have instead created a force of repulsion between Tuscany and Sassari in the non-STEM degrees.

The link between Veneto and Oristano for the non-STEM field of studies is not surprising given the historical relationships between these two territories. Indeed, numerous Venetian families arrived in this area of Sardinia during the fascist reclamations (Ruiu et al., 2020), and their offspring who are still living in this area may have maintained contacts with relatives in Veneto. This evidence may support the ability of the proposed approach to capture the chain effects.

Among the non-scientific disciplines, the chain effects seem to be present between Cagliari and Latium (since the 1950s), Lombardy (since the 1990s), and Emilia-Romagna.

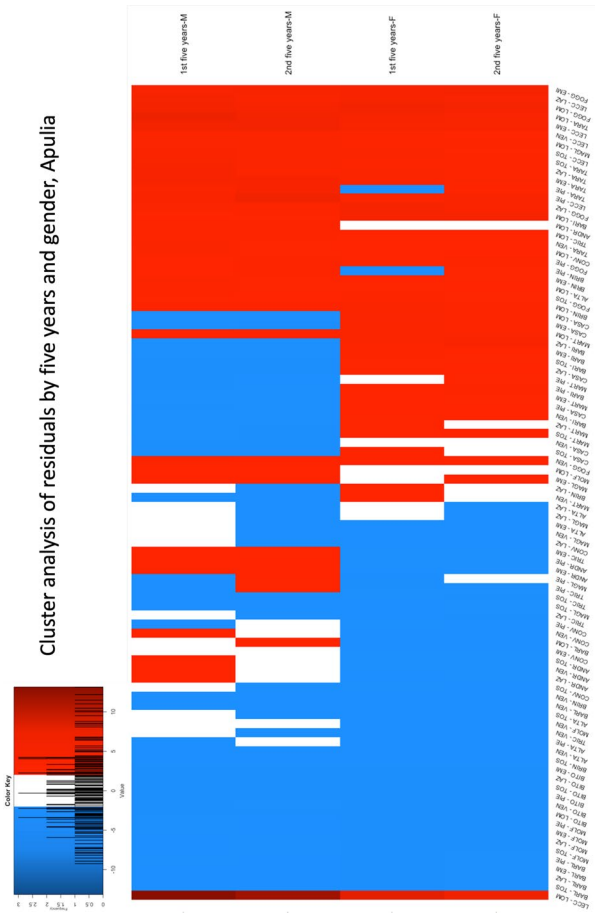

Fig. 7 Cluster analysis of residuals by 5-years for gender, Apulia

Emilia Romagna is also strongly linked to Oristano and Nuoro, while the students of Olbia-Tempio seem to have deeper ties with Lombardy. When gender is also considered for Sardinia, it is confirmed that mobility patterns across scientific subjects are mainly composed of male students. It is worth noticing that, also for Cagliari (the capital and the biggest city in Sardinia) there exists a "city to city" connection like the one revealed for Palermo, Catania, and Messina.

Finally, regarding Apulia (Figs. 7 and 8), a partially different picture emerges compared to the two islands: the mobility patterns formed for scientific subjects largely correspond to those formed for the non-STEM subjects. This difference between Apulia and the islands may be attributed to geographical factors rather than an actual difference in behaviour patterns. Sardinia and Sicily are islands; thus, connections with the mainland are only by airplane, so it is equally difficult/expensive to reach any destination in the Center-North. In contrast, for students coming from Apulia, it is easier to reach some attractive Italian universities, which could somewhat amplify the chain effect. This is also consistent with the historical pattern of mobility of Apulian students highlighted in "The origin of mobility routes from the 1950s to the 1990s" section. Indeed, the preferential mobility patterns identified with our approach have been present since the 1950s. Specifically, large cities such as Bari, Brindisi, Foggia, Lecce, and Taranto exhibit strong

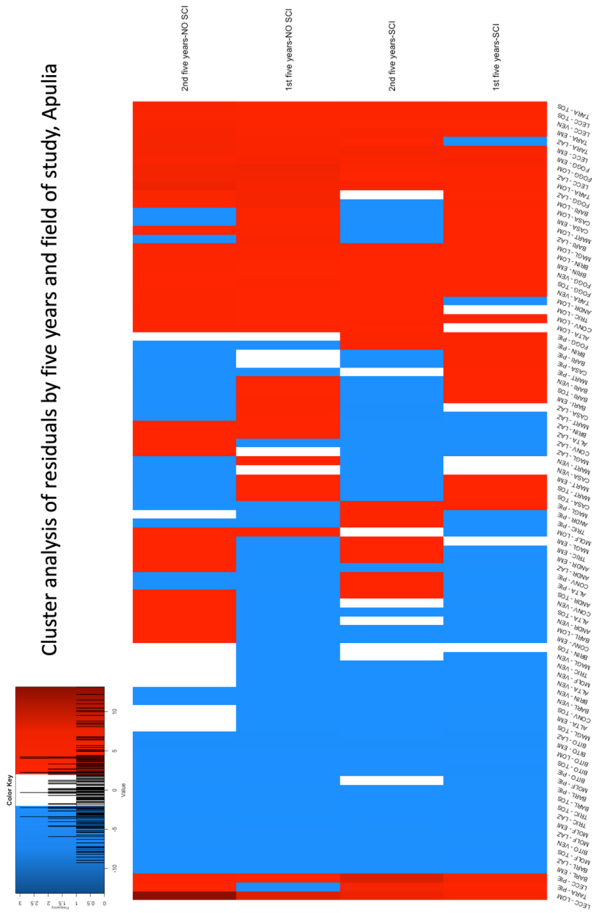

Fig. 8 Cluster analysis of residuals by 5-years for field of study, Apulia

connections with Emilia-Romagna. Notably, in the 1970s, student flows from these areas towards Bologna increased by one and a half. Other persistent mobility patterns include significant links between Bari, Brindisi, and Foggia with Lombardy and Latium, and between Lecce and Taranto with Latium, Lombardy, Veneto, Tuscany, and Piedmont. In contrast, the "medium-large city to large city" and "medium-small city to medium-small city" mobility model appears to be less pronounced with respect to Sicily and Sardinia. The relative ease of connection between Apulia and other regions may be a contributing factor. The geographical closeness likely reduces both the monetary and psychological costs of moving, making it more likely for students to study outside their region, regardless of their field of study or gender. As a result, Apulia presents the highest mobility rates for both genders. Even though a lower influence of gender stereotypes for Apulia cannot be excluded, this explanation would not seem consistent with Istat (2020). Indeed, according to Istat in 2019 the female activity rate on the labour market in Sicily and Apulia was 38.7% and 40%, respectively, while for Sardinia it rose to 55.9%. These labor market statistics suggest that a traditional male bread-winner model is still strong in Apulia. Hence, this seems to be not very coherent with the idea that a lower influence of cultural gender stereotypes is behind the higher female mobility in this southern region.

Conclusions

In this work, a statistical approach was proposed to three Italian case studies (Apulia, Sardinia, and Sicily) to identify the student mobility patterns that can be at least partially explained by the existence of a process like the phenomenon of the chain migratory. The hypotheses used in the work are essentially three: (*i)* the areas of origin of the flows can be identified by analysing the mobility flows given by the high school students living in the same areas of origin; (*ii)* the forces of attraction and repulsion towards/from a destination/origin can be quantified using the column/row profiles of an origin–destination matrix; (*iii)* the difference between the observed flows and the expected ones (under the hypothesis of independence) over time can partially reflect the presence of student migration patterns present over time.

The first hypothesis seemed reasonable since communities that gravitate around the same area of origin are likely to share largely the same cultural traits and refer to the same local economic systems. In addition, the construction of these origin areas allows for overcoming the limitation of administrative areas, such as provinces or municipalities.

The second and third hypotheses are well established within the analysis of qualitative data of student mobility, although, to the best of our knowledge, the present work is the first to use this technique to perform a quantitative analysis based on students' mobility micro-data. This approach enables the identification of student mobility patterns that may be considered as some evidence for a migratory chain effect.

Obviously, the reasoning of the assumptions does not imply that the work is free from limitations. The first is certainly having focused the attention only on bachelor students. The reason for this choice has in fact more practical than scientific motivations. The number of master's degree students, especially in Sardinia, is low and this implies difficulties in identifying areas of origin that are not so extensive in terms of territory, as well as greater variability that could make it difficult to identify patterns. The same reason led us to aggregate the data into five-year periods. Moreover, one of the limitations lies in the number of years considered in our analysis. Indeed, a greater number of years could have facilitated the identification of mobility patterns among students that align with typical migratory chains. Unfortunately, this was not feasible due to the distinct characteristics of students enrolled during the COVID-19 pandemic. During that period, many Italian universities transitioned to online education, making it challenging to discern whether the mobility observed during this time is genuine or artificial.

Finally, the chain migration effect is intended as a residual after eliminating the push and pull factors. This argument should be supported by an analysis of students' motivations to move conducted through a survey and/or in-depth interviews at individual level. Using a metaphor, the same statistical tool that we propose in this paper can be considered as a sort of thermographic survey that indicates where to dig deeper to uncover chain mechanisms and its determinants. Obviously, we can exclude that other forces are causing the large residuals that we observer for some mobility routes, however, the temporal stability of these routes remarked also by the historical analysis give some strength to this interpretation of this phenomenon.

With these limitations in mind, some interesting results seem to emerge. First, the hypothesis that some mobility patterns seem to be affected by chain migration effects is

supported by our data. Both Sicily and Sardinia show mobility patterns associated to the field of study, while Apulia tends to replicate the same patterns of mobility for both fields of study. This difference between Apulia and the Islands is likely driven by factors related to the transportation network that connects Apulia to the Centre-North of the country, which is rather different from the Islands' network, based in the last decades essentially on airplane routes' availability.

Another noteworthy aspect is the applicability of a "medium-large city to large city" and "medium-small city to medium-small city" model to student mobility data, especially for Sardinia and Sicily. In fact, students from Palermo, Catania and Messina and those from Cagliari tend to move to the largest cities in the Centre-North of the country, namely, Milan and Rome, whereas students from small cities tend to prefer universities located in small-medium cities. Probably, in this case, psychological factors play a role. Moving from a small context to a very large one may be more frightening than directing toward one of comparable size. Furthermore, some scholars (Burt, 1992; Kuschminder, 2016; Massey, 1990; Wilson, 1998) have introduced the concept of weak and strong ties in migration. Thus, strong and weak ties play a role in migration and can be seen for students' mobility too. Strong ties provide concrete help at the beginning of a migration process, weak ties reduce the redundancy in information coming in from a dense and close community of relatives and friends. It is possible that strong ties (families and close friends) shape the transition from high school to a bachelor: a student decides to move from the area of origin to another region for their bachelor, guided by relatives and friends that can concretely help her/him at destination.

Our approach could be integrated in future research with a qualitative one to offer a better comprehension of the social mechanisms that drive students' mobility choices. For instance, it can be of interest to investigate if student mobility patterns are also related to the North–South historical internal migration routes (Panichella, 2012). Indeed, a question that remains open in our work is: are the chain effects due to the exchange of information with peers who have already moved or to the presence of relatives who moved from the South to the North in the middle of the Italian economic miracle? Although we cannot provide a definitive answer to this question at this time, we believe that our historical analysis gives some support to this interpretation. In any case, we believe that the primary contribution of this paper lies in proposing a relatively simple approach to detect preferential mobility patterns that could be related to some chain migration. The presence of long-lasting and intergenerational communities in the host cities suggests a network that can be both used for students' mobility and job placement; moreover, the flow from overcrowded universities (Milan, Rome, Naples) towards the cities of the Center-Nord, contributed to the growth of new economic areas, such Toscana, Umbria, Emilia or Veneto, the Adriatic coast (the so-called "Terza Italia"; Bagnasco, 1977). The mobility network created during the Eighties and Nineties seems to be active and available yet.

Lastly, it is known mobility is generally positive, providing new opportunities for those who have experienced it. But the Italian pathways are one-way. This could represent an additional channel for the impoverishment of the local human capital, creating another element of inequality as sustained by Hillman (2016) in his analysis of "educational deserts".

Furthermore, considering ISTAT data, the last SVIMEZ Annual Report *L'economia e la società del Mezzogiorno* (SVIMEZ, 2022) sheds light on how the economic and social disparities within the country have led to a continuous exodus of people from the South to the Center-North. Between 2002 and 2020, nearly 2.5 million individuals left the South to relocate to the Center-North, officially changing their places of residence. It is noteworthy that more than half of these emigrants are young people aged 15 to 34, and one over five holds a degree. In 2020, the percentage of graduates among those who left the South was particularly high in Basilicata (47.7%), Molise (47.2%), and Abruzzo (44.1%), while in other Southern regions, this percentage ranged from 32 to 39%. Regarding young emigrants aged 25 to 34, the percentage of graduates significantly increased, rising from 17% in 2002 to 33% in 2010, and then to 48% in 2020. In addition to these labor flows, we can almost certainly include student mobility, which is not registered by Istat, since students change their address just when they have a stable employment. Finally, another alarming consequence of these movements of students and workers to the Center-North, which cannot be addressed here, is the economic loss of the South. For the period from 2000 to 2015, the OECD estimates that approximately 30 billion euros were invested in the education and training of Southern youth. The significant migrations from the South result in a substantial loss of this investment, depriving the South of its "best" young individuals, namely graduates and students. Those who leave the South tend to have a higher socio-economic status on average than those who stay (Türk, 2019). As reported by the *"Rivista economica del Mezzogiorno, Trimestrale della Svimez"*, Southern universities receive less public funding, lower tuition fees paid by students and limited funding from local public entities, private sources, and international institutions. Consequently, Southern universities have an overall lower level of economic resources, approximately 11% below the national average. This disparity has worsened in recent years due to reduced public funding and declining enrolments at Southern universities, leading to an increasingly wide gap between universities in the South and those in the Center-North (SVIMEZ, 2017). Fadda et al. (2022) also showed that the geographical location of universities affects their ability to increase their financial resources. Notably, universities located in the Northern regions, the wealthiest part of Italy, receive higher amounts of resources. The authors emphasize that "universities located in the Northeast and Northwest receive, on average, respective recruitment budgets of €29,000 and €22,000 higher per 1000 students per year than those located in the South."

As previously mentioned, the issue of Southern Italy losing its students and workers is a significant challenge for Italy, because emigrants are seeking more favourable living and working conditions. This phenomenon, which has existed for a long time, is likely endemic. To address challenges such as the depopulation of the South and the economic challenges faced by Southern universities, the SVIMEZ suggests, possibly drawing inspiration from some American experiences where universities have played a vital role in local economies, increasing funding for the southern universities.

Appendix

See Tables 3, 4, 5 and 6.

Table 3 Origin–destination arrays

A-Sicily										
Period	***AreaOris***	**Destination**								**Total**
		EMILIA-ROMAGNA %	**LATIUM %**	**LOMBARDY %**	**PIEDMONT %**	**TUSCANY %**	**VENETO %**	**OTHER %**	**SICILY %**	
2008–2012	AGR	4.5	1.6	5.3	2.0	5.5	0.8	2.4	77.9	4338
	CAN	6.5	2.9	6.4	2.9	5.8	0.8	3.6	71.2	3696
	CAS	11.1	4.1	8.3	3.7	11.4	1.9	6.0	53.4	2782
	CAT	0.6	1.6	2.2	0.5	0.6	0.3	1.2	92.9	12,517
	MES	1.2	2.1	3.2	1.1	1.1	0.7	3.1	87.6	8264
	PAL	0.9	1.2	2.3	0.4	0.8	0.3	1.0	93.0	15,544
	RAG	6.9	4.7	7.9	6.3	9.8	1.0	4.4	59.0	5731
	SIR	4.5	4.3	5.3	3.4	4.3	0.8	2.8	74.6	4743
	TRA	12.4	5.3	6.2	3.7	15.4	2.0	6.6	48.3	4313
	VIT	8.2	3.7	6.1	2.2	9.9	1.4	5.5	63.1	1831
	Other AreaOris	1.8	1.8	3.0	1.6	2.2	0.5	2.2	86.9	41,228
Total 2008–2012		2.9	2.3	3.7	1.8	3.4	0.6	2.5	82.7	104,987
2013–2017	AGR	5.5	2.3	7.4	6.6	7.6	1.3	4.6	64.7	3666
	CAN	6.3	2.6	6.8	7.6	8.3	1.2	5.3	62.0	3463
	CAS	9.1	4.0	8.9	7.7	13.4	1.7	7.9	47.3	2648
	CAT	1.4	1.6	3.2	1.5	0.8	0.6	1.7	89.1	11,577
	MES	2.2	2.6	4.3	2.7	1.3	0.9	3.6	82.4	7326
	PAL	1.8	2.0	2.8	1.2	0.9	0.6	1.2	89.5	14,590
	RAG	8.3	4.4	8.8	11.2	12.1	1.3	5.8	48.1	5144
	SIR	7.2	3.5	7.1	9.4	5.9	1.2	5.3	60.5	3956
	TRA	16.1	5.4	7.5	9.2	12.1	2.8	9.6	37.3	3918
	VIT	10.7	3.3	7.2	10.2	11.1	1.5	6.0	49.9	1657
	Other *AreeOri*	3.0	1.9	3.9	3.5	3.2	1.0	2.8	80.5	37,964
Total 2013–2017		4.1	2.4	4.7	4.3	4.2	1.0	3.4	76.0	95,909

Table 3 (continued)

B- Sardinia										
Period	***AreaOris***	**Destination**								**Total**
		EMILIA-ROMAGNA %	**LATIUM %**	**LOMBARDY %**	**PIEDMONT %**	**TUSCANY %**	**VENETO %**	**OTHER %**	**SARDINIA %**	
2008–2012	CAG	1.2	1.4	1.9	0.9	1.2	0.5	1.0	91.9	14,562
	NUO	3.8	2.9	3.0	4.1	4.6	1.0	3.0	77.7	5117
	OLB	4.0	7.0	8.0	5.3	6.7	1.7	2.8	64.5	2349
	ORI	3.0	1.9	1.8	3.0	3.0	1.3	1.8	84.1	3012
	SAS	1.8	1.7	2.9	4.1	2.7	0.9	1.1	84.7	6541
Total 2008–2012		2.1	2.1	2.7	2.6	2.7	0.8	1.5	85.3	31,581
2013–2017	CAG	1.8	1.2	2.3	1.9	1.6	0.7	1.2	89.4	13,709
	NUO	4.2	2.6	3.6	5.5	5.9	1.7	4.0	72.4	4862
	OLB	5.0	3.6	7.0	5.3	6.4	1.7	5.3	65.8	2323
	ORI	3.6	1.4	2.3	4.0	2.9	2.0	2.7	81.0	2554
	SAS	2.4	1.3	2.5	4.7	2.9	1.2	1.7	83.4	6217
Total 2013–2017		2.7	1.6	2.9	3.5	3.0	1.1	2.2	82.8	29,665

Table 3 (continued)

C- Apulia											
Period	***AreaOris***	**Destination**									**Total**
		CAMPANIA %	**EMILIA-ROMAGNA %**	**LATIUM %**	**LOMBARDY %**	**PIEDMONT %**	**TUSCANY %**	**VENETO %**	**OTHER %**	**APULIA %**	
2008–2012	ALT	0.3	2.2	3.6	4.1	1.6	1.7	0.4	9.8	76.3	4687
	AND	0.4	2.6	2.4	3.2	1.3	1.5	0.5	6.0	82.0	4383
	BARI	0.2	0.8	1.9	2.3	0.6	0.7	0.3	2.2	91.0	13,931
	BAR	0.2	1.7	1.6	2.2	1.2	1.2	0.4	3.8	87.7	2592
	BIT	0.3	0.9	1.1	1.1	0.5	0.4	0.5	1.9	93.4	1707
	BRI	0.3	7.8	4.6	6.4	4.7	2.5	0.6	6.7	66.4	7882
	CAS	0.3	10.5	5.7	7.7	4.8	3.3	0.9	7.3	59.5	3510
	CON	0.4	1.9	2.3	2.6	1.0	1.3	0.5	3.1	86.9	5800
	FOG	1.4	7.5	5.3	4.1	1.7	2.0	0.4	21.6	56.1	15,457
	LEC	0.3	5.5	5.6	6.3	3.5	2.1	0.8	5.9	70.0	10,172
	MAG	0.2	10.0	6.3	7.9	5.0	3.6	0.9	7.3	58.9	3796
	MAR	0.4	5.5	4.5	5.4	4.0	4.5	0.8	7.1	67.7	3843
	MOL	0.4	0.8	1.1	2.1	0.5	0.8	0.4	2.8	91.0	4236
	TAR	0.5	6.2	5.3	4.5	2.3	4.7	0.7	8.6	67.3	11,421
	TRI	0.2	16.2	8.8	7.5	5.5	3.5	1.3	10.6	46.4	2160
Total 2008–2012		0.5	5.1	4.1	4.4	2.3	2.2	0.6	8.3	72.5	95,577

Table 3 (continued)

C- Apulia											
Period	***AreaOris***	**Destination**									**Total**
		CAMPANIA %	**EMILIA-ROMAGNA %**	**LATIUM %**	**LOMBARDY %**	**PIEDMONT %**	**TUSCANY %**	**VENETO %**	**OTHER %**	**APULIA %**	
2013–2017	ALT	0.7	2.9	2.5	5.6	2.4	2.7	0.8	10.5	71.8	4655
	AND	1.2	3.3	2.0	4.3	3.4	1.7	1.0	6.6	76.5	4350
	BARI	0.5	1.9	1.7	3.0	1.3	0.7	0.6	3.2	87.1	12,721
	BAR	0.5	1.7	1.6	3.3	1.4	1.4	0.4	4.9	84.8	2597
	BIT	0.4	0.6	0.9	1.1	0.8	0.6	0.7	2.1	92.7	1577
	BRI	0.6	8.7	3.7	6.5	7.7	2.8	0.7	9.0	60.2	6709
	CAS	0.3	10.9	2.9	7.3	8.8	3.4	1.0	8.8	56.5	3293
	CON	0.4	2.8	1.9	2.7	2.2	1.4	0.6	4.3	83.7	5111
	FOG	2.5	8.4	4.2	5.0	2.9	2.1	0.9	21.1	52.9	14,586
	LEC	0.6	5.5	3.5	7.3	4.9	2.0	0.9	6.5	68.6	9397
	MAG	0.3	9.3	4.4	9.3	7.4	3.5	0.9	8.9	55.9	3359
	MAR	0.7	8.0	3.5	5.9	7.8	3.8	1.4	10.3	58.6	3465
	MOL	0.6	1.5	1.2	2.8	1.3	0.7	0.7	4.1	87.0	4163
	TAR	0.9	7.0	5.0	5.9	5.2	5.0	1.0	12.3	57.7	9719
	TRI	0.5	15.5	4.2	9.7	9.4	4.3	0.9	12.0	43.6	1856
Total 2013–2017		0.9	5.8	3.1	5.2	4.0	2.4	0.8	9.6	68.2	87,558

Table 4 Origin–Destination arrays for first-year university students in 1955–56

A-Sicily

Origin	Destination								
	EMILIA-ROMAGNA	LATIUM	LOMBARDY	PIEDMONT	TUSCANY	VENETO	OTHER	SICILY	Total
	%	%	%	%	%	%	%	%	
AGRIGENTO	0.7	2.9	2.9	1.4	1.4	0.7	3.7	85.8	268
CALTANIS-SETTA		3.5	2.8	1.4	2.8		6.4	82.7	139
CATANIA		0.9	1.3	0.5		0.2	0.8	96.0	935
MESSINA	0.1	0.8	0.8	0.8	0.4	0.1	2.6	94.1	720
PALERMO		0.9	1.3	0.3	0.1	0.2	1.5	95.5	916
RAGUSA	0.4	3.4	1.4	0.9	0.4	0.4	1.9	90.5	202
SYRACUSE	0.2	2.8	1.9	2.8	0.8		3.3	87.8	353
TRAPANI	0.3	1.8	2.2	1.8	0.3	1.1	8.1	84.0	269
Not specified	0.7	2.9	0.7	2.1		1.4	2.9	89.0	137
Total	0.1	1.5	1.5	1.0	0.4	0.3	2.5	92.3	3939

B-Sardinia

Origin	Destination								
	EMILIA-ROMAGNA	LATIUM	LOMBARDY	PIEDMONT	TUSCANY	VENETO	OTHER	SARDINIA	Total
	%	%	%	%	%	%	%	%	
CAGLIARI	0.1	0.3	0.4	0.1	0.3	0.0	1.6	97.1	694
NUORO	0.0	6.4	0.9	1.8	1.8	1.8	0.9	86.2	109
SASSARI	0.0	3.9	4.3	0.7	7.1	1.1	5.3	77.7	282
Total	0.1	1.8	1.5	0.5	2.2	0.5	2.5	91.0	1085

C-Apulia

Origin	Destination									
	CAMPANIA	EMILIA-ROMAGNA	LATIUM	LOMBARDY	PIEDMONT	TUSCANY	VENETO	OTHER	APULIA	Total
	%	%	%	%	%	%	%	%	%	
BARI	1.8	0.4	2.0	2.5	0.9	0.4	0.2	0.8	91.1	1116
BRINDISI	5.6	1.5	7.1	4.1	1.0	2.5	0.5	0.5	77.2	197
FOGGIA	20.9	2.8	12.0	3.7	0.5	2.1	1.6	2.8	53.8	435
LECCE	10.0	3.3	13.8	3.8	2.1	3.1	2.6	1.4	59.8	420
TARANTO	7.9	1.7	8.6	3.3	2.0	1.3	1.3	1.3	72.6	303
Total	7.6	1.5	7.0	3.2	1.2	1.4	1.0	1.3	75.8	2471

Table 5 Origin–Destination arrays for first-year university students in 1967–68

A- Sicily

Origin	Destination								
	EMILIA-ROMAGNA	LATIUM	LOMBARDY	PIEDMONT	TUSCANY	VENETO	OTHER	SICILY	Total
	%	%	%	%	%	%	%	%	
AGRIGENTO	1.2	2.0	1.8	1.3	3.3	0.5	5.6	84.3	782
CALTANIS-SETTA	1.0	2.7	2.3	1.9	1.6	1.6	1.0	87.9	486
CATANIA	0.0	0.9	0.9	0.6	0.4	0.3	1.3	95.5	2637
MESSINA	0.3	1.3	1.3	1.3	0.2	0.4	1.7	93.5	2150
PALERMO	0.1	0.8	0.9	0.4	0.4	0.3	2.4	94.7	1858
RAGUSA	0.8	3.2	3.4	1.6	1.4	1.0	1.1	87.5	625
SYRACUSE	0.5	2.2	0.9	2.3	1.6	0.3	1.0	91.3	1050
TRAPANI	0.7	3.4	2.9	2.0	4.9	1.4	4.2	80.6	716
Not specified	0.6	3.4	1.2	0.6	1.2	0.3	0.9	91.7	324
Total	0.4	1.7	1.4	1.1	1.1	0.5	2.0	91.8	10,628

B- Sardinia

Origin	Destination								
	EMILIA-ROMAGNA	LATIUM	LOMBARDY	PIEDMONT	TUSCANY	VENETO	OTHER	SARDINIA	Total
	%	%	%	%	%	%	%	%	
CAGLIARI	0.2	1.8	0.7	0.4	3.1	0.5	1.4	91.8	1689
NUORO	0.5	5.5	2.1	0.9	7.3	0.9	2.5	80.4	439
SASSARI	0.7	5.9	2.5	2.5	11.2	1.6	4.3	71.5	769
Total	0.4	3.5	1.4	1.0	5.9	0.9	2.3	84.7	2897

C- Apulia

Origin	Destination									
	CAMPANIA	EMILIA-ROMAGNA	LATIUM	LOMBARDY	PIEDMONT	TUSCANY	VENETO	OTHER	APULIA	Total
	%	%	%	%	%	%	%	%	%	
BARI	0.9	0.5	1.2	0.9	0.9	0.6	0.3	0.6	94.1	3287
BRINDISI	1.1	2.6	3.2	1.1	1.9	3.2	0.4	1.9	84.8	756
FOGGIA	7.2	3.9	5.9	3.4	3.6	2.0	1.5	12.2	60.3	1546
LECCE	2.0	2.8	4.4	2.4	1.1	4.1	1.4	2.3	79.5	1935
TARANTO	2.3	2.3	4.2	1.1	2.3	4.0	1.6	1.7	80.5	1347
Total	2.5	2.1	3.4	1.7	1.7	2.3	0.9	3.2	82.2	8871

Table 6 Tables of abbreviations

A- Abbreviations of target regions	
Destination regions	**Abbreviations**
Abruzzo	ABR
Basilicata	BAS
Calabria	CAL
Campania	CAM
Emilia-Romagna	EMI
Friuli Venezia Giulia	FRI
Lazio	LAZ
Liguria	LIG
Lombardy	LOM
Marche	MAR
Molise	MOL
Piedmont	PIE
Apulia	PUG
Sardinia	SAR
Sicily	SIC
Tuscany	TOS
Trentino Alto Adige	TRE
Umbria	UMB
Valle D'Aosta	VAL
Veneto	VEN
B-Abbreviations of Sicilian areas	
Sicilian areas	**Abbreviations**
Agrigento	AGRI
Canicattì	CANI
Castelvetrano	CAST
Catania	CATA
Messina	MESS
Palermo	PALE
Ragusa	RAGU
Syracuse	SIRA
Trapani	TRAP
Vittoria	VITT
C-Abbreviations of Sardinian areas	
Sardinian areas	**Abbreviations**
Cagliari	CAGL
Nuoro	NUOR
Olbia-Tempio	OLBI
Oristano	ORIS
Sassari	SASS
D-Abbreviations of Apulian areas	
Apulian areas	**Abbreviations**
Altamura	ALTS
Andria	ANDR
Bari	BARI
Barletta	BARL

Table 6 (continued)

D-Abbreviations of Apulian areas	
Apulian areas	**Abbreviations**
Bitonto	BITO
Brindisi	BRIN
Casarano	CASA
Conversano	CONV
Foggia	FOGG
Lecce	LECC
Maglie	MAGL
Martina Franca	MART
Molfetta	MOLF
Taranto	TARA
Tricase	TRIC

Author contributions
All the authors equally contributed.

Funding
This study was partially funded by the European Union – NextGenerationEU, in the framework of the GRINS-Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 – CUP B73C22001260006). 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. This paper has been also supported by 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 Stem in Higher Education & Women INequalitieS [SHE WINS], CUP I53D23004810006, Grant Assignment Decree No. 1060 adopted on 07/17/2023 by the Italian Ministry of Ministry of University and Research (MUR).

Availability of data and materials
Not applicable.

Declarations

Ethics approval
Not applicable.

Competing interests
Not applicable.

Received: 17 October 2023 Accepted: 17 September 2024
Published online: 09 October 2024

References

Aina, C., Mussida, C., & Lombardi, G. (2022). Are business and economics alike? *Italian Economic Journal*. https://doi.org/10.1007/s40797-021-00177-w

Anderberg, M. R. (2014). *Cluster analysis for applications: Probability and mathematical statistics: A series of monographs and textbooks* (Vol. 19). Academic Press.

Attanasio M, Enea M. (2020). La mobilità degli studenti universitari nell'ultimo decennio in Italia, in G. De Santis, E. Pirani, M. Porcu (eds), Rapporto sulla popolazione. L'istruzione in Italia, Il Mulino, Bologna, 43–58.

Attanasio, M., & Priulla, A. (2020). Chi rimane e chi se ne va? Un'analisi statistica della mobilità universitaria dal Mezzogiorno d'Italia. In M. Attanasio, O. Giambalvo, M. Porcu, & G. Ragozzini (Eds.), *Verso Nord* (pp. 27–63). Le nuove e vecchie rotte delle migrazioni universitarie.

Ballarino, G., Colombo, S., Panichella, N., & Piolatto, M. (2022). Human capital dynamics: The geographical mobility of high-school graduates towards university in Italy. *Regional Studies, 56*(6), 921–939. https://doi.org/10.1080/00343404.2021.1912723

Böcker, A. (1994). Chain migration over legally closed borders: Settled immigrants as bridgeheads and gatekeepers. *Journal of Social Sciences, 30*(2), 87–106.

Brooks, R., & Waters, J. (2010). Social networks and educational mobility: The experiences of UK students. *Globalisation, Societies and Education, 8*(1), 143–157. https://doi.org/10.1080/14767720903574132

Burt, R. S. (1992). Structural holes: The social structure of competition. *Harvard University Press*. https://doi.org/10.4159/9780674029095

Cammelli, A. (2000). Contare gli studenti, Statistica e popolazione studentesca dall'Unità ad oggi. *Annali Di Storia Delle Università Italiane, 4*(2000), 9–23.
Cervia, S., & Biancheri, R. (2017). Women in science: The persistence of traditional gender roles. A case study on work–life interface. *European Educational Research Journal, 16*(2–3), 215–229.
Eurenius, A. M. (2020). A family affair: Evidence of chain migration during the mass emigration from the county of Halland in Sweden to the United States in the 1890s. *Population Studies, 74*(1), 103–118. https://doi.org/10.1080/00324728.2018.1559945
Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. John Wiley & Sons.
Fadda, N., Marinò, L., Pischedda, G., & Ezza, A. (2022). The effect of performance-oriented funding in higher education: Evidence from the staff recruitment budget in Italian higher education. *Higher Education, 83*, 1003–1019. https://doi.org/10.1007/s10734-021-00725-4
Genova, V. G., Tumminello, M., Attanasio, M., & Aiello, F. (2021). A network analysis of student mobility patterns from high school to master's. *Statistical Methods & Applications, 30*, 1445–1464. https://doi.org/10.1007/s10260-021-00592-4
Giocondo, R. (2022). *Il Dissidente Italiano*. Booksprint Edizioni.
Haug, S. (2008). Migration networks and migration decision-making. *Journal of Ethnic and Migration Studies, 34*(4), 585–605. https://doi.org/10.1080/13691830801961605
Hillman, N. W. (2016). Geography of college opportunity: The case of education deserts. *American Educational Research Journal, 53*(4), 987–1021. https://doi.org/10.3102/0002831216653204
Istat (2020). Il mercato del lavoro 2019. Rapporto. https://www.istat.it/it/files//2020/03/Nota-stampa-Rapprto-mercato-del-lavoro-2019.pdf
Kuschminder, K. (2016). Strong ties, weak ties: Exploring the role of networks in domestic worker migration from Ethiopia to the middle east. *Asian and Pacific Migration Journal, 25*(4), 401–421.
Liu, J. M., Ong, P. M., & Rosenstein, C. (1991). Dual chain migration: Post-1965 Filipino immigration to the United States. *The International Migration Review, 25*(3), 487–513.
Mabogunje, A. L. (1970). Systems approach to a theory of rural-urban migration. *Geographical Analysis, 2*(1), 1–18.
MacDonald, J. S., & MacDonald, L. D. (1964). Chain migration ethnic neighborhood formation and social networks. *The Milbank Memorial Fund Quarterly, 42*(1), 82–97.
MacDonald, J. S., & MacDonald, L. D. (1970). Italian Migration to Australia. Manifest functions of bureaucracy versus latent functions of informal networks. *Journal of Social History, 3*, 249–275.
Massey, D. S. (1990). Social structure, household strategies, and the cumulative causation of migration. *Population Index, 56*, 3–26.
Palloni, A., Massey, D. S., Ceballos, M., Espinosa, K., & Spittel, M. (2001). Social capital and international migration: A test using information on family networks. *American Journal of Sociology, 106*(5), 1262–1298.
Panichella, N. (2012). Le migrazioni interne nel secolo scorso: Vecchie e nuove forme a confronto. *Stato e Mercato, 95*(2), 255–281.
Pérez, P. A., & McDonough, P. M. (2008). Understanding Latina and Latino College Choice. A social capital and chain migration analysis. *Journal of Higher Education, 7*(3), 249–265.
Person, A. E., & Rosenbaum, J. E. (2006). Chain enrollment and college enclaves: Benefits and drawbacks of Latino college students' enrollment decisions. *New Directions for Community Colleges, 133*, 51–60.
Reyneri, E. (1998). Emigrazione ed Economia Sommersa. *Stato e Mercato, 53*(2), 287–317.
Ruiu, G., Fadda, N., Ezza, A., & Esposito, M. (2019). Exploring mobility of Italian Ph.Ds over the last decades. *Electronic Journal of Applied Statistical Analysis, 12*(4), 748–773.
Ruiu, G., Genova, V. G., Attanasio, M., Breschi, M. (2020), L'isola che se ne va. Un'analisi della mobilità degli studenti sardi, in Delogu, M., Meleddu , M. (a cura di), Riflessioni trasversali sul capitale umano, Forum Editrice, Udine, pp. 105–113.
Tocchioni, V., & Petrucci, A. (2021). Italian PhD students at the borders: The relationship between family background and international mobility. *Genus, 77*, 14. https://doi.org/10.1186/s41118-021-00127-5
Viesti, G., Pellegrini, G., Iuzzolino, G. (2011). Convergence among Italian Regions, 1861–2011. Quaderni di Storia Economica, 22, Bank of Italy. Downloadable: Retrieved from https://www.bancaditalia.it/pubblicazioni/quaderni-storia/2011-0022/Quaderno_storia_economica_n_22.pdf
Wilson, T. (1998). Weak ties, strong ties: Network principles in mexican migration. *Human Organization, 57*(4), 394–403.
Wolfe, D. A. (2005). The role of universities in regional development and cluster formation. In G. A. Jones, P. L. McCarney, & M. L. Skolnik (Eds.), *Creating knowledge, strengthening nations: The changing role of higher education* (pp. 167–194). University of Toronto Press.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.