



# Surnames in Local Newspapers and Social Mobility

Massimo Baldini<sup>1</sup> · Andrea Barigazzi<sup>1</sup> 

Accepted: 5 August 2024 / Published online: 22 August 2024  
© The Author(s) 2024

## Abstract

This study aims to investigate social mobility by introducing an innovative data source: surnames in newspapers. The core hypothesis posits that frequent mentions in newspapers are indicative of greater social relevance. Through the collection of local newspaper data in the municipality of Modena, Italy, from 1921 to 2011, we examine variations in the relative representation of surnames in newspapers with respect to their presence in registry data. The results suggest that surnames in newspapers are not a random sample of the population, supporting the assumption that they reflect social significance. Surnames belonging to privileged groups exhibit a higher representation compared to other social groups. This higher relative representation seems to be transmitted to future generations and converges over time toward the mean, with variations depending on the considered high-status group. This kind of analysis could contribute to identifying different mobility patterns at the local level and represents a useful alternative when established data sources, such as income, education, and occupational data, are not available.

**Keywords** Newspapers · Social mobility · Surnames · Elites

**JEL Classification** J62 · N34

## 1 Introduction

The study of social mobility between generations has always stimulated academic debate (see among others Ganzeboom et al., 1991). However, the difficulty in finding data linking generations over time has held back for a long time the development of research on this subject. It is only in recent decades that the empirical literature on intergenerational mobility

---

✉ Andrea Barigazzi  
andrea.barigazzi@unimore.it

Massimo Baldini  
massimo.baldini@unimore.it

<sup>1</sup> Department of Economics “Marco Biagi”, University of Modena and Reggio Emilia, Via Jacopo Berengario 51, Modena 41121, Italy

has made considerable progress, mainly due to the availability of new and better data<sup>1</sup> linking adjacent generations (Solon, 2018). Social mobility over long periods has received less attention from scholars because of data limitations (Solon, 1999; Black & Devereux, 2010). This has pushed in parallel to the search for alternative methods and data to overcome this obstacle. In particular, there is a growing interest in the exploitation of surnames in studying social mobility, used as pseudo-links connecting households of different generations (Clark, 2014, Barone & Mocetti, 2021). The present work builds on this literature by introducing a novel data source for the analysis of social mobility: newspapers.

The premise is to verify if information collected from newspapers can be used as a possible complement to existing measures of social mobility. We aim to test two basic assumptions: whether the frequency of surnames in newspapers provides independent informative content and is not a random sample from the population, and whether the surnames that appear most frequently in local newspapers have greater social relevance. If these hypotheses are confirmed, we can then study the changes over time in the surnames of elite groups in newspapers, expecting that stable representation indicates low social mobility, while significant fluctuations would suggest high mobility and social changes in relation to different historical phases.

To do so, we investigate the relative representation (Clark and Cummings, 2014) of surnames in newspaper articles compared to their presence in the population. If privileged groups (nobility, cultural elites, etc.) are more frequently represented in newspapers than in population registers, this confirms the hypothesis that newspaper surnames are not randomly selected but carry traits of social relevance<sup>2</sup>. Therefore, if this over-representation remains stable over time, it indicates the transmission of power across generations.

Several features derived from these data can be used to analyse social mobility over time, bearing in mind that the informational content of a surname can vary greatly depending on the context and type of the article in which it appears. For example, many of the surnames that appear in the political section of a newspaper correspond to people who have power in collective decision-making processes; in crime-related articles, the surnames of both the victims and the criminals (often of lower socio-economic status) appear; those who advertise their products/services may be considered richer than the average member of the reference population, just like those who spend themselves on charity work or who can afford to buy space for an obituary.

To ascertain whether there have been variations in the relative representation (RR) of surnames over the years, the surnames mentioned in local newspapers in the municipality of Modena, Italy, from 1921 to 2011 are taken into account, together with registry data on the spread of surnames in the general population. By way of explanation, if there is an over-representation of certain surname groups in pages dealing with local politics compared to their registry distribution, then that surname group will most likely belong to individuals/

---

<sup>1</sup> These types of data mainly refer to (i) cross-sections with retrospective surveys on the social status of parents, (ii) panel surveys with detailed and repeated information on the social and economic status of household components followed over extended periods of time, (iii) fiscal/administrative data linking economic information between parents and children (Mooi-Reci, 2020).

<sup>2</sup> The hypothesis that the likelihood of being quoted in newspapers is associated with higher social status has already been examined in other studies. For instance, the literature studying obituaries emphasizes that they reveal the influence of dominant Western elites, groups linked by common origins, and educational patterns (Matsuda, 1996; Fowler & Bielsa, 2007). Ban et al. (2019) demonstrate that the volume of press coverage devoted to political actors or offices helps to indicate their actual power.

families exercising political power at the local level. If this over-representation persists over time, then power is transmitted from one generation to the next.

As with other papers that exploit the information content of surnames, it is not possible to state with certainty that a surname observed today is part of the direct descendants of the same surname found in previous decades. But, in line with Barone and Mocetti (2021), by focusing on the local rather than the national level, more precise links between generations are produced. Moreover, the strong heterogeneity and localism of Italian surnames further reinforce the quality of these pseudo-links. Additionally, when adjusting for rare surnames in the population, the strength of these connections becomes even more apparent.

To the best of our knowledge, the contribution of this work to the existing literature is threefold. We are the first to propose the use of surnames appearing in newspapers to study social mobility and power transmission of elites. Moreover, while the literature on long-run social mobility has traditionally focused on income, education, and occupation outcomes (Solon, 2018), we implement an analysis that is rather concerned with the transmission of authority, power, and social influence. Finally, the employment of newspapers as a source of historical data has mainly been developed in the United States and the United Kingdom, while in Europe, due to the lower presence of accessible digitised archives, it is a tool with wide scope for scientific research.

Our findings indicate that surnames in newspapers provide independent informative content and are not randomly extracted from the population. We find support for the hypothesis that surnames appearing in newspapers have particular social significance; those associated with higher social status groups are more represented compared to the rest of the population. Surnames linked to socially significant groups in the past tend to maintain a greater relative representation in newspapers over time, highlighting the transmission of social status across generations. This transmission regresses toward the mean at varying rates depending on the social group examined.

The rest of the paper is organised as follows. In Sect. 2, a brief review of the research on newspapers and on surnames (in particular in the study of social mobility) is proposed. Section 3 explains the data used. Section 4 reports the method of analysis and the observed results, while two robustness checks are proposed in Sect. 5. Finally, Sect. 6 concludes.

## 2 Literature Review

### 2.1 The Use of Surnames in Economic Research and the Study of Social Mobility

A person's surname can tell a lot about the social status of the ancestors and the geographical origin of the family (e.g., the surnames *Le Boulanger* in France, *Becker* in the United States, and *Bäcker* in Germany link their owners to their ancestors' occupation as bakers). Furthermore, the information conveyed by surnames may influence the way a person is judged, a fact that economic research has exploited to study discrimination in various fields, for example in the labour or in the rental markets. For instance, Pascual et al. (2015) find that popular surnames receive a higher and more positive evaluation than uncommon and infrequent ones in the labour market. Noble-sounding surnames more frequently hold managerial positions than other surnames (Silberzahn & Uhlmann, 2013). A person with an African American-sounding name is less likely to be called back for a job interview than

a person with a white-sounding name (Bertrand & Mullainathan, 2004), and anonymous application procedures increase the chances of both women and individuals of non-Western origin of advancing to the interview stage (Åslund & Skans, 2012). Many studies have also found signs of discrimination in rental housing markets, where foreign-sounding names that have shown interest in the ads are less likely to be contacted by flat owners (Ahmed & Hammarstedt, 2008).

As an alternative method of measuring intergenerational social mobility, the use of surnames has received increasing interest. Collado et al. (2012), having data on the socioeconomic status of individuals in two Spanish regions at the end of the 19th and the end of the 20th century, use pseudo-links obtained from surnames to link ancestors to descendants. They find that having a high level of education and belonging to a high socioeconomic group is still influenced by the socioeconomic status of the great-great-grandparents. Guell et al. (2015) estimate a correlation of educational outcomes of 0.60 for Catalonia in Spain. Clark (2014) finds that in a wide set of examined countries, multigenerational mobility follows an autoregressive process with a high and time-invariant persistence rate. Barone and Mocetti (2021) go even further back in time, exploiting a unique dataset linking the status of Florentines in 1427 (derived from a tax census) to that of their likely descendants (with the same surnames) in 2011. They find that the elasticity of earnings is positive and statistically significant between generations nearly 600 years apart. They also show evidence of even stronger transmission of real wealth and persistence in some elite professions. None of these papers that study mobility through surnames, however, has made use of surnames taken from newspapers.

## 2.2 Newspapers

Newspapers can be the source of an incredibly vast amount of easily accessible historical information (Hansen, 2004). In fact, their pages contain the historical memory collected day by day of major political, cultural, economic, scientific, and other events (Tosh, 2010). Newspaper data can be applied to study a wide range of topics, but so far only some of them have been explored. Some examples refer to the healthcare sector (Markel et al., 2007; Ager et al., 2022), the study of collective action (for a review, see Earl et al., 2004), or news sentiment (Gentzkow et al., 2006; Shapiro et al., 2020; Aguilar et al., 2021).

As for research in economics, Hanlon and Beach (2022) highlight that, before the arrival of large-scale digitized historical newspaper databases, the literature has predominantly focused on the use of newspapers: (i) as a way of measuring a certain type of treatment in order to construct a key explanatory variable; (ii) as the basis of many price series in studies on financial and commodity markets. They also emphasise that the bulk of existing economic studies have focused on data from the United States or the United Kingdom, where most of the digital archives are located. A very promising research perspective is therefore to expand the use of newspapers outside this geographical area.

### 3 Data

The analysis incorporates different data sources. The primary one is based on a dataset covering surnames mentioned in local newspapers in the municipality of Modena, Italy, spanning from 1921 to 2011. Since digitized archives of these newspapers are not available, data are collected manually by extracting surnames from physical articles and recording them in a database categorized by year, section (for further details, see Table 1), and newspaper title. This entailed the need to restrict the survey to specific time intervals. Arbitrarily, surnames are gathered for one year per decade (1921, 1931, ..., 2011), considering the first three days of each month and the entire month of March. A robustness check with a different selection of days is provided in Sect. 5.

It is important to collect data from different newspapers since media coverage is not independent of the biases of the newspapers themselves and the preferences of the readers (Gentzkow & Shapiro, 2010; Larcinese et al., 2011). There have been three local newspapers active in the territory of Modena in the reference period. Including all three newspapers reduces this bias.

- *La Gazzetta di Modena*<sup>3</sup>, the main daily newspaper in the city of Modena, active throughout the time span analyzed. It is a moderate, generalist daily, not politically aligned (except of course during the fascist period).
- *L'Unità (local edition of Modena)*, available for the years 1951-1961-1971-1981-1991 of our sample. Historical left-wing daily linked to the Italian Communist Party. Over the years it has gradually embraced more moderate and reformist positions following the evolution of its reference party.
- *Il Resto del Carlino (local edition of Modena)*, used for the years between 1961 and 2011. It is a generalist daily, historically affiliated with the agrarians of Emilia and the sugar industrialists during the first decades of the XX century and usually supporting center-right parties after the Second World War.

**Table 1** Collected surnames by year and section

Section	1921	1931	1941	1951	1961	1971	1981	1991	2001	2011	Total
<i>Advertising</i>	355	662	363	754	755	311	184	324	325	77	4,110
<i>Business</i>	205	373	181	284	197	403	277	63	394	177	2,554
<i>Charity</i>	359	998	552	434	361	212	73	0	0	0	2,989
<i>Public Events &amp; Civil Society</i>	212	211	548	665	446	964	762	854	2,334	2,271	9,267
<i>Crime News &amp; Court Reporting</i>	529	294	413	902	101	402	262	151	268	230	3,552
<i>Obituaries</i>	55	43	384	245	82	282	152	135	287	212	1,877
<i>Local Politics</i>	531	467	476	578	671	790	804	1,228	1,447	1,605	8,597
<i>Religion</i>	0	32	20	58	49	40	63	79	87	72	500
<i>Science &amp; Technology</i>	74	366	111	148	404	313	258	224	395	347	2,640
<b>Total</b>	<b>2,320</b>	<b>3,446</b>	<b>3,048</b>	<b>4,068</b>	<b>3,066</b>	<b>3,717</b>	<b>2,835</b>	<b>3,058</b>	<b>5,537</b>	<b>4,991</b>	<b>36,086</b>

Note: The *Crime News & Court Reporting* section is a combination of the subsections' *thief*, *victim*, *accidents*, and *judges-lawyers-police*

<sup>3</sup> The newspaper has gone through several name changes: *Gazzetta dell'Emilia* (from the foundation to 1947); *La Gazzetta di Modena* (1947–1953); *Gazzetta dell'Emilia* (1953–1967); *Gazzetta di Modena* (1968–1977); *Nuova Gazzetta di Modena* (1981-present).

In order to distinguish the mentions according to the different spheres of influence of individuals and the type of article, surnames are classified into pre-established sections: Advertising, Business, Charity, Public Events & Civil Society, Crime News & Court Reporting, Obituaries, Local Politics, Religion, Science & Technology. Surnames appearing in articles that refer to non-local (for example national politics) and sports events are excluded. Table 1 shows the number of mentions per year and per section. Additional information and statistics on this and the other datasets used in the paper can be found in the Appendix tables.

The second dataset concerns the distribution of surnames within the population. This type of information is crucial as it allows us to compare how frequently a surname is mentioned in newspapers relative to its presence in the general population. We have the surnames of household heads in 1936, and of all residents in Modena in 1981 and 2001 from the Italian population census<sup>4</sup> 8th, 12th, and 14th.

The subsequent phase involves the identification of surnames belonging to elite groups. Among these groups, we include surnames associated with the Modenese nobility, surnames of professors working at the University of Modena, surnames of students attending the 'Licei' high school, and surnames from the Jewish community.

The 'Licei' dataset contains information regarding the surnames of students who attended Licei in the municipality of Modena. In Italy, the educational system offers three main types of high school pathways: Licei, technical institutes, and vocational schools. While the latter two pathways directly prepare students for entry into the workforce, Licei are primarily designed to provide preparation for further university studies. Being able to afford secondary and tertiary education indicates a certain level of economic well-being, and it is particularly anticipated for past years when access to university education was a privilege reserved for a few, and only more affluent families could afford to support their children's studies.

Noble surnames refer to the nearly 300 Modenese families that held noble titles at the beginning of the 20th century (Pietramellara, 1902). In Italy, noble titles have not been recognized since 1948 following the enactment of the Italian Constitution. However, nobility had for centuries indicated a social class distinguished by privileges and attributes, and it is reasonable to expect that these prerogatives (especially better economic and asset conditions) continued, at least in part, even after the cessation of noble title recognition in 1948.

The group of university professors is represented by the surnames of those who were engaged in teaching and research at the University of Modena starting from 1921. Alongside a better economic and income condition, it is presumable that this social group had greater influence on the socio-cultural landscape of Modena.

Finally, the city of Modena has hosted a community of Italian Jews since the Middle Ages. The history of this social group has undergone very heterogeneous phases. During the Este Duchy, although recognized as citizens, Jews were ghettoized, preventing any cohabitation with Catholic Christians and barring them from elections to municipal councils (Ghelfi et al., 1995). This disadvantaged condition manifested in significant impoverishment of the Jewish community, with about 30% of them receiving subsidies at the end of the 18th century (Fregni, 1993). It was only from 1859, with the annexation of the Este Duchy to the Kingdom of Sardinia and the ensuing creation of the Kingdom of Italy, that the condition of the Jewish community saw a significant and rapid improvement. With the abolition of discriminatory laws and recognition of civil and political capacity, there was an enhancement

<sup>4</sup> We access the surnames through the registry data of the resident Modenese population, which are collected in conjunction with the population census.

in the well-being also of the Modenese Jewish community, which aligns with the broader context of the transformation of Jewish conditions in Italy. As reported by Levi (1997), at the beginning of the 20th century, the Jewish population held higher occupational positions compared to the rest of the Italian population: 8% worked in agriculture versus 55% of the Italian population; 29% were employed in industry and over 40% in commerce, compared to 30% and just over 5% of the rest of the population; 23% were involved in free professions or public administration compared to 9% of the rest of the population. Higher average living standards are also reflected in longevity studies, which indicate that after emancipation, the Modenese Jewish community experienced longer life expectancy and lower mortality rates compared to the rest of the population (Marach, 1999). This phenomenon fits into the general trend in Italy of higher life expectancy for the Jewish community compared to the rest of the population (Della Pergola, 1970). Further, already in 1861 only 5.8% of adult Jews were illiterate, compared to 64.5% of the total Italian population (Carocci, 2005). By 1927, illiterates among the Jews had disappeared while they were still 27% of all residents in Italy. This favorable situation, which began with Italian unification, was radically disrupted during the period of racial persecution and with the introduction of Fascist racial laws in the 1930s. The Jewish group, however, differs from the other three elite groups we identify. While it is reasonable to assume that most members of the other groups were well-off, certainly the group of Jews had a component in poor economic conditions. Nonetheless, there is strong evidence that another portion occupied positions typical of the middle and upper middle class, and some of them did hold prominent political and economic positions. It should also be noted that in the studies of social mobility through surnames, the use of typically Jewish surnames to indicate elite groups has already been applied (Clark, 2014).

More detailed information on these surname groups is provided in Table A8.

## 4 Methods of Analysis and Results

The fundamental premise of this study is that the information collected from newspapers can be used as a valuable complement to existing measures of social mobility. To test this hypothesis, a series of steps need to be followed. First, it is important to:

- i) Determine whether the frequency of surnames in newspapers provides independent informative content and is not a random extraction from the population.
- ii) Verify whether the surnames that most frequently appear in local newspapers have significant social relevance.

If these two steps are fulfilled, then the change over time in the way surnames are present in newspapers can provide information on social mobility. The second phase involves analyzing the mobility of elite surname groups in newspapers over time. We expect these social groups to have higher representation in newspapers. Specifically, the more stable and higher the representation of these surname groups over time, the lower social mobility, as the same family surnames tend to be consistently present in newspapers. Conversely, the greater the fluctuation in surname distribution across decades, the more significant social change and mobility. At the same time, we expect that the presence of these groups in newspapers will

change based on different historical phases and the varying conditions these groups have experienced over time.

In all these phases, the key statistic on which we rely is the Relative Representation (Clark & Cummins, 2014). This is the ratio between the share of a surname, or groups of surnames present in a context where it is presumed that privileged positions are transmitted, and the share of those same surnames in the population during the specified period. Clark and Cummins (2014) for example compute the RR of rare noble surnames among students at Oxford or Cambridge as follows:

$$\text{Relative Representation of } x = \frac{\text{Share of } x \text{ in privileged context}}{\text{Share of } x \text{ in general population}},$$

where  $x$  is a noble surname or a group of noble surnames, and the privileged context is the group of all students of Oxford or Cambridge. In our case,  $x$  is a surname belonging to one of our four elite groups (nobles, Licei students, university professors, and Jewish) and the privileged context of the numerator is the total set of surnames present in the newspapers in a given year.

We have only access to the frequency distribution of surnames in the general population from registry data for the years 1936, 1981, and 2001. This necessitates using the same registry for multiple decades as the denominator for calculating relative representation. This challenge is mitigated by the fact that the population composition changes slowly over time. Specifically, for the years 1921, 1931, 1941, and 1951, we rely on the 1936 registry. For the years 1961, 1971, and 1981, we use the 1981 registry, while for the years 1991, 2001, and 2011, we refer to the 2001 registry.

Therefore, the relative representation of noble surnames in 1921 newspapers is:

$$RR \text{ of noble surnames in newspaper}_{1921} = \frac{\text{Share of noble surnames in newspaper}_{1921}}{\text{Share of noble surnames in registry}_{1936}}$$

#### 4.1 Checking Newspaper Data

The first check on the newspaper data concerns their potential independent informational content. It is natural for the most common surnames in the general population to be widely represented in newspapers. However, the question is whether the distribution of these surnames in newspapers represents a random sample of the general population. In this case, their change over time cannot provide information on social mobility.

To address this issue, we compute the relative representation of each surname and assess whether these values exhibit a correlation with their respective relative frequencies in the general population (Table 2). If a strongly positive correlation is observed, it indicates that the surnames in newspapers are merely a random reflection of the general population composition and do not provide autonomous information. For surnames reported in newspapers in 1921, the correlation coefficient between their RR and their relative frequency in the population is -0.23. In 1981, the same correlation is -0.13. For the 200 most common surnames in the 1936 registry, the correlation in 1921 is exactly 0 (Table A9 in appendix). For the 200 most frequent surnames in the 1981 registry data, the correlation is -0.04. In 1921,

**Table 2** Correlation between the RR of surnames in newspapers and their frequency distribution in general population

	RR 1921	RR 1931	RR 1941	RR 1951	RR 1961	RR 1971	RR 1981	RR 1991	RR 2001	RR 2011
<i>Frequency distribution registry 1936</i>	-0.23	-0.22	-0.19	-0.24						
<i>Frequency distribution registry 1981</i>					-0.16	-0.16	-0.13			
<i>Frequency distribution registry 2001</i>								-0.21	-0.08	-0.14

Note: Frequency distributions in general population refer exclusively to surnames present in newspapers during the reference year. Therefore, the  $-0.23$  value for 1921 is the correlation between the RR of surnames present in 1921 newspapers and their frequency distribution in the 1936 registry

the 200 most common surnames in the 1936 registry exhibit an average RR of 0.57, which is close to the overall average (0.63). This trend persists in subsequent years. Therefore, we can conclude that the surnames in newspapers do not result from a mere random extraction from the underlying population.

The second control questions whether there is a basis for the assumption that surnames appearing in local newspapers possess particular social relevance. Naturally, not all surnames in this source belong to the local elite, but for a portion of them, this could be a plausible scenario. Once this point is established, observing the evolution of surname distribution in newspapers over time can provide insight into changes in the structure of elite groups. If the distribution of surnames belonging to privileged groups shows few changes between two time periods, it may suggest that influential positions have not undergone significant alterations, and vice versa.

To test the hypothesis that surnames in newspapers contain information about elite groups, we compare the relative representation of surnames that appear in newspapers and belong to a socially relevant group with the RR of the rest of the newspaper's surnames. For example, we expect that surnames that belong to the noble group have a RR in newspapers greater than the RR of the surnames that are present in the newspapers but are not included in the noble group.

Table 3 confirms these expectations. Surnames associated with socially relevant groups are more prominently represented in newspapers than the others, and this trend remains consistent over the analyzed years. For the four examined elite groups, their average relative representation is higher than that of the general population. In almost all cases, their relative representation is also above 1, a value indicating equitable distribution of surnames in newspapers and registers. For instance, noble surnames are almost four times more present in newspapers than in the registry in 1921, while non-noble surnames are more present in the general population than in newspapers. The same holds for university professors teaching in 1921, with a relative representation of 3.42. This value further increases for professors in 1931, with a RR seven times higher in newspapers than in the general population.

The low number of surnames belonging to a group may produce high standard errors indicating considerable data variability. Therefore, it is a positive sign that the results remain robust for the group of Licei students, where surname diversity and numerosity are higher.

An intriguing aspect also arises in the comparison of relative representation between typically Jewish surnames in Modena and the other surnames: in 1921 the RR of Jewish surnames is 4.31, almost eight times higher than that for non-Jewish surnames (0.59). In 1931,

**Table 3** RR of surnames groups of noble, professors, jewish and licei students in newspapers

Year	Noble		Professors		Jewish		Licei Students	
	No	Yes	No	Yes	No	Yes	No	Yes
1921	0.54 (0.43-0.65)	3.72 (1.38-6.05)	0.56 (0.44-0.67)	3.42 (0.67-6.17)	0.59 (0.47-0.71)	4.31 (0.98-7.64)	0.53 (0.41-0.66)	1.86 (1.16-2.55)
1931	0.63 (0.51-0.74)	1.87 (1.14-2.59)	0.57 (0.48-0.66)	7.27 (1.74-12.80)	0.64 (0.53-0.75)	4.35 (1.47-7.24)	0.56 (0.45-0.67)	1.98 (1.41-2.54)
1941	0.64 (0.50-0.77)	2.24 (0.43-4.04)	0.62 (0.49-0.74)	4.21 (-0.91-9.33)	0.67 (0.52-0.81)	1.22 (-0.32-2.75)	0.56 (0.42-0.70)	1.91 (1.29-2.48)
1951	0.55 (0.46-0.63)	1.63 (0.56-2.71)	0.74 (0.46-0.63)	3.28 (0.53-6.10)	0.57 (0.48-0.65)	1.56 (-0.50-3.61)	0.53 (0.43-0.61)	1.10 (0.84-1.36)
1961	0.65 (0.52-0.77)	1.70 (0.79-2.61)	0.59 (0.46-0.71)	4.44 (2.33-6.54)	0.66 (0.52-0.78)	2.29 (0.39-4.18)	0.62 (0.49-0.75)	1.77 (1.13-2.40)
1971	0.80 (0.65-0.93)	2.20 (0.49-3.90)	0.75 (0.61-0.90)	2.70 (1.81-3.58)	0.81 (0.66-0.94)	2.34 (-1.37-6.05)	0.77 (0.62-0.92)	1.42 (1.12-1.73)
1981	0.78 (0.58-0.98)	1.36 (0.45-2.26)	0.73 (0.53-0.92)	2.64 (0.57-4.71)	0.79 (0.58-0.98)	1.51 (-0.44-3.46)	0.77 (0.56-0.98)	1.13 (0.85-1.4)
1991	0.55 (0.47-0.63)	1.39 (0.55-2.23)	0.54 (0.46-0.62)	1.45 (0.91-1.98)	0.56 (0.47-0.63)	0.95 (-0.66-2.56)	0.52 (0.44-0.60)	1.42 (1.06-1.77)
2001	0.55 (0.48-0.62)	1.15 (0.69-1.61)	0.54 (0.48-0.61)	1.42 (0.80-2.03)	0.55 (0.48-0.62)	3.20 (0.29-6.11)	0.52 (0.45-0.59)	1.46 (1.10-1.81)
2011	0.55 (0.47-0.63)	0.95 (0.41-1.48)	0.53 (0.45-0.61)	2.08 (1.17-2.99)	0.56 (0.47-0.63)	1.82 (0.06-3.57)	0.53 (0.45-0.62)	1.26 (0.78-1.73)

Notes: 95% confidence interval in parentheses. Surnames collected in the subsection's thief, victim, accidents are not included. See Table A.8 in the Appendix for surname groups definitions and statistics

Jewish surnames maintain a significantly higher RR; however, in 1941, three years after the introduction of discriminatory racial laws against Jews in Italy by the fascist regime, their RR drops to 1.22. From 1951 onwards, with the return of democracy and the beginning of the economic expansion phase, the relative representation of Jewish surnames rises again and in 1961 surpasses the value of 2. Newspapers seem to mirror the changing fortunes of this social group, reflecting societal dynamics over time.

It must be considered that the section of the newspaper where these surnames appear can influence the results. A valid concern arises regarding the possibility that various articles may not necessarily depict the events of the most influential individuals, but rather unrelated incidents not necessarily linked to social status. This concern is particularly pertinent for the section dedicated to Crime News and Court Reporting. Therefore, we excluded surnames associated with articles on crimes, both as victims and perpetrators, from Table 3. Similarly, we excluded all incidents, whether they were road accidents, domestic, or workplace incidents.

For the aforementioned reasons, we anticipate that surnames linked to elite groups will have a lower or in-line RR compared to the rest of the population when considering only the subsection covering incidents or thefts in the Crime News and Court Reporting section. This expectation appears to be validated. When we narrow the analysis to surnames collected only in the subsection covering incidents or thefts, the relative representation of surnames belonging to high-status groups decreases significantly, aligning with that of other surnames (see Table A10).

An additional check can be conducted by examining a random group extracted from the population registries. We expect this group not to show a significantly different RR from the rest of the population. We extracted the first 500 surnames in alphabetical order from the 1936, 1981, and 2001 records, and observed values of their average RR exceeding 1 only in 1921 and 1941, while in other years, they were below this threshold (see Table A11).

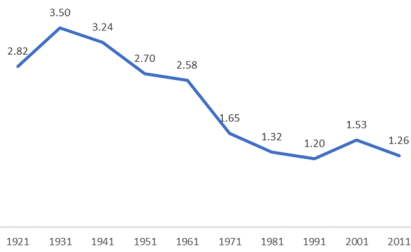
## 4.2 Mobility of Elite Surnames Groups in Newspapers

The preceding section has substantiated our hypothesis that surnames appearing in newspapers are not merely a random extraction from the population but rather contain elements of social relevance/distinction. The way their representation changes over time can thus provide insights into societal dynamics. The more stable the RR of surname groups over time, the lower social mobility, as the same individuals or families tend to be consistently present in newspapers. Conversely, the greater the fluctuation in surname distribution across decades, the more significant social change and mobility.

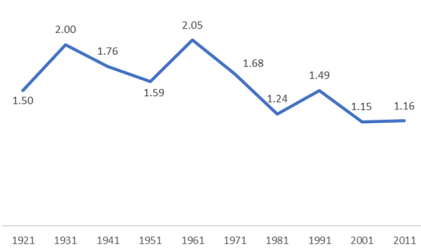
The data at our disposal allow us to differentiate by social group and section. We begin our analysis by focusing on the surnames of university professors and Licei students, as these provide more reliable estimates due to a larger number of observations and a greater variety of surnames. Furthermore, for these groups, we can distinguish by year, with information available for professors teaching in 1921, 1931, and so forth, as well as for students in the same years. For noble surnames and those belonging to the Jewish community, the numerosity and variety of surnames remain stable over time, given the absence of new entries or exits in these groups; for instance, new noble families are not named in the 1900s.

From a broader perspective, examining all newspaper sections except for the categories “thief,” “victim,” and “accidents,” reveals changes in the relative representation of sur-

*University Professors 1921-1931-1941-1951*

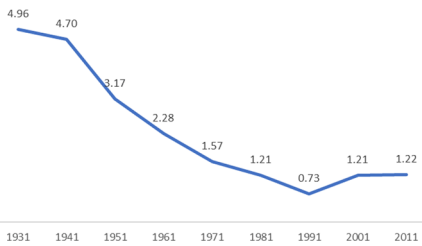


*Licei students 1921-1931-1941-1951*

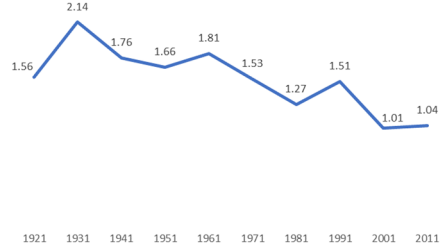


**Fig. 1** RR of university professors and Licei students who taught and studied in 1921-1931-1941-1951  
 Notes: Surnames collected in the subsection's thief, victim, accidents are not included. See Table A.8 in the Appendix for surname groups definitions and statistics

*University Professors 1931-1941*



*Licei students 1931-1941*



**Fig. 2** RR of university professors and Licei students who taught and studied in 1931 and 1941  
 Notes: Surnames collected in the subsection's thief, victim, accidents are not included. See Table A.8 in the Appendix for surname groups definitions and statistics

names of university professors and Licei students who taught and studied in the years 1921-1931-1941-1951, referring to the 1936 registry (Fig. 1). Both groups exhibit an RR greater than 1, which would occur if the proportion of surnames in the newspapers mirrored that of the general population. In 1931 and 1941, university professors are present in newspapers over 3 times more than in the registry data, while in 1921 and 1951, their presence is more than double their representation in the general population. Over time, the RR tends to decrease, stabilizing around values like 1.2 and 1.3. Conversely, Licei students maintain a constant RR over time, with a slight decrease in the last 40 years considered.

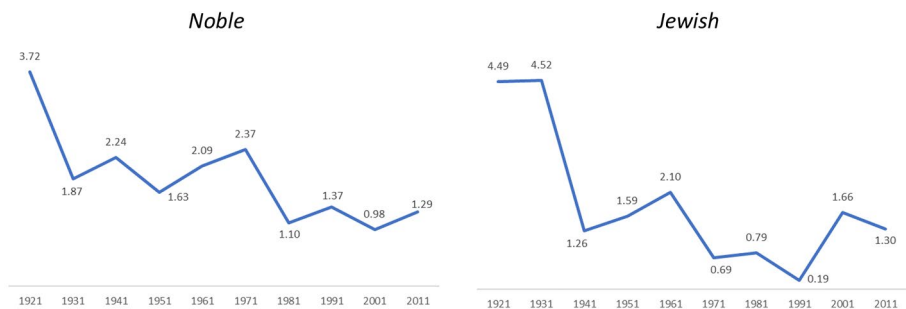
If we narrow the analysis further to the surnames present in the list of university professors and Licei students only in 1931 and 1941, the years closest to the 1936 registry, the RR of professors in those years becomes even more pronounced (Fig. 2). In newspapers from 1931 to 1941, they are present almost 5 times more than in the registry data, but over time, their RR decreases, converging towards the mean (RR=1) from the 1980s. Licei students, on the other hand, have a lower RR, just over 2 times their representation in the registry data in 1931 and 1.76 in 1941. A regression towards the mean is also observed for them, though it is slower and less pronounced.

As previously mentioned, the analysis of noble surnames and those belonging to the Jewish community is constrained by the stability of these data over time, marked by the absence of new entries or exits within these surname groups. The only distinction feasible is to consider, for years after '51, when the reference registry transitions to that of 1981 and 2001, only those surnames present in the 1936 registry. Consequently, a noble surname is considered in 1961 only if it was also present in the 1936 registry.

The results confirm a higher prevalence of noble surnames in newspapers compared to their representation in the general population in the early 1900s, particularly in 1921 (Fig. 3). This trend remains robust until the 1970s, converging towards the mean from the 1980s onward. Surnames of Jewish origin exhibit a higher relative representation in the pre-World War II years, sharply declining in 1941 following the introduction of racial laws in Italy and the country's alignment with Nazi Germany. Their RR experiences a slight increase in 1951 and 1961 but remains significantly lower compared to the first part of the examined period.

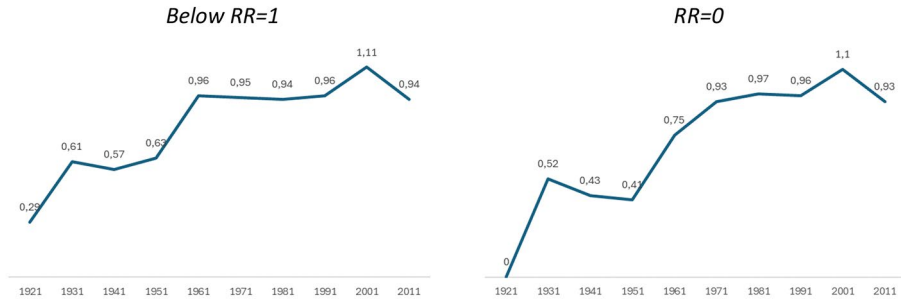
We have seen so far that surnames belonging to elite groups at the beginning of the period tend to be over-represented in newspapers, and that there is over time a reduction in their RR, consistent with the presence of social mobility. Similarly, for those poorly represented or even absent in newspapers in a given year, a convergence towards the mean is expected in the medium to long term. This phenomenon is particularly evident for surnames with a low RR but a widespread presence in the general population. Figure 4 provides a clear illustration of this concept: among the 200 most common surnames in 1936, we select the 161 with low RR in newspapers in 1921 (below the equitable distribution of 1 between surnames in newspapers and registers). Overall, their RR is 0.29 in 1921. An increase is observed in the following years, stabilizing around the threshold of equal distribution of 1 starting from the 1960s. Focusing the analysis on the 53 surnames without mentions in newspapers in 1921 reveals a similar trend. There is a consistent regression towards the mean converging to 1 by 1971, with a particularly pronounced increase in the 1960s and 1970s.

The analysis of data across newspaper sections reveals additional aspects of interest. For instance, we can investigate the mobility of surnames associated with university professors who taught in the years 1921-1931-1941-1951, specifically focusing on the section directly relevant to their field, namely "Science and Technology" (Fig. 5). Their relative representation is significantly high during the years directly relevant to them (from 1921



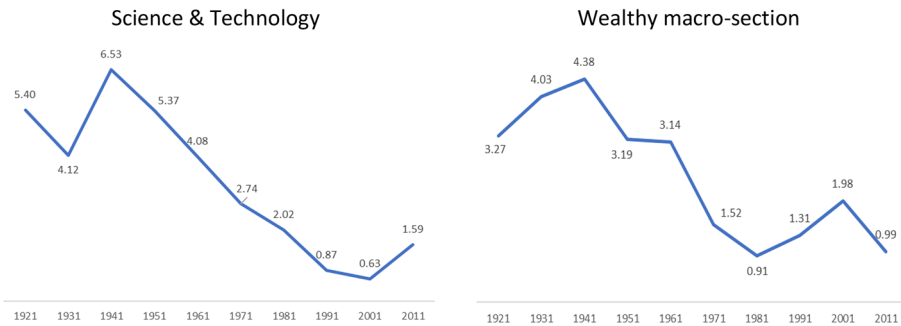
**Fig. 3** RR of noble and Jewish community surnames

Notes: Surnames collected in the subsection's thief, victim, accidents are not included. See Table A.8 in the Appendix for surname groups definitions and statistics



**Fig. 4** RR of common and underrepresented surnames in 1921 newspapers

Notes: Surnames collected in the subsection's thief, victim, accidents are not included. See Table A.8 in the Appendix for surname groups definitions and statistics



**Fig. 5** RR of university professors who taught in 1921-1931-1941-1951 in the Science & Technology section and Wealthy macro-section of newspapers

Notes: See Table A.8 in the Appendix for surname groups definitions and statistics

to 1951), surpassing that observed when considering all newspaper sections. This elevated representation persists in the subsequent years before regressing towards the mean from the 1990s onward.

As one would expect, professions associated with privilege, such as university professors, are anticipated to correlate with a better economic status. Therefore, it is not surprising to observe that professors from 1921-1931-1941-1951 exhibit a high Relative Representation in the macro-section labeled “wealthy<sup>5</sup>” during those respective years (Fig. 5). Although their representation diminishes over time in this macro-section, it consistently remains above 1.

In the final phase, we propose a comparison between the results obtained from the newspaper dataset and those obtained from the Licei dataset as alternative indicators of elite groups. The information pertaining to students who attended Licei is particularly relevant in the context of one of the most established indicators in the study of social mobility: education. This stems from the expectation that the surnames of individuals belonging to

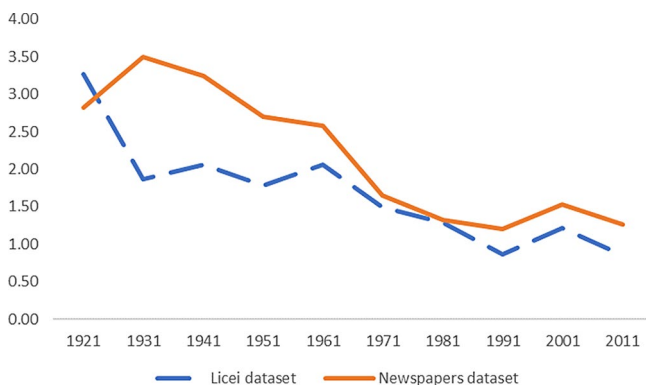
<sup>5</sup> The “wealthy” macro section refers to the combination of the Advertising, Business, Obituaries sections, and the subsection Judge-Lawyer-Police.

privileged social groups will exhibit a higher relative representation within the registries of Licei compared to their distribution in the general population. Therefore, the Licei dataset, in this case, is not used as an elite group to study its relative representation in newspapers, but directly as a privileged context (similar to newspapers in our above analysis or to Oxford and Cambridge students in Clark and Cummins (2014)). If  $x$  is the surname of a university professor, we compare the  $\frac{\text{share of } x \text{ in newspapers}}{\text{share of } x \text{ in general population}}$  to the  $\frac{\text{share of } x \text{ among Licei students}}{\text{share of } x \text{ in general population}}$ . If we expect Licei to be, like newspapers, places where there is an overrepresentation of elite social groups, then we can compare how the RR of these elite groups changes over time in newspapers compared to how it changes in Licei. The more the trend of RR for these groups is similar in newspapers and Licei, the more robust the results obtained will be.

We consider the surnames of professors from the years 1921, 1931, 1941, and 1951. Besides studying their RR in newspapers, we also analyze it in the Licei dataset. The surnames associated with professors have a higher representation in Licei and in the newspapers compared to their presence in the general population (Fig. 6). Although the two curves do not directly overlap, they show a similar trend. In newspapers, there is a notable emphasis on the RR of the elite group, particularly in 1931 and 1941.

## 5 Robustness Check

The first check concerns the use of surnames; indeed, it is not possible to state with certainty that a surname observed today is part of the direct descendants of the same surname found in previous decades. We do not follow the same individuals over time, but the same surnames. One method to strengthen these pseudo-connections is to limit the investigation to the rarest surnames. We repeat the analyses previously performed by excluding the 200 most common surnames in the 1936 registry and the 600 most common surnames in the 1981 and 2001 registries. This allows us to have a maximum of 20 families with the same surname in 1936 and a maximum of 50 individuals with the same surname in 1981 and 2001 (consistent with an average family size of 2.5 people). It is also important to consider that the vast majority of surnames belong to only one family in 1936 (62.17%, see Table A5 in the appendix). If we add surnames belonging to a maximum of 2 or 3 families, we cover almost 80% of the



**Fig. 6** RR of university professors who taught in 1921-1931-1941-1951 in newspaper and Licei dataset  
Notes: Surnames collected in the subsection's thief, victim, accidents are not included. See Table A.8 in the Appendix for surname groups definitions and statistics

total population. The same applies to 1981 and 2001; considering surnames belonging to a maximum of 5 individuals cover 80% of the entire population. Table 4 reproduces the setup of Table 3, showing an RR of rare surnames among elite groups in line with or even higher than that observed for all surnames. Similar trends, but with higher RRs for rare surnames, are also confirmed for other analyses (see Figure A1 in the appendix, reproducing Fig. 2).

The second robustness check concerns newspapers. Throughout the 20th century, just as society has changed, so have newspapers, their format, the number of pages, the events they choose to highlight, and consequently, their informational content. Table 1 and especially Table A1 in the appendix highlight this aspect. For instance, there are significant variations in the “Public Events & Civil Society” section, which has gained increasing space from the 1970s and even more so in the 2000s. To ensure robust estimates, information provided in local newspapers should ideally remain relatively constant. Otherwise, observed changes may reflect a decrease (or increase) in the space given to certain articles relative to others. For this reason, we conduct a verification of the results that take these considerations into account by reweighting surnames.

If the incidence of a section on the total observations in a particular year exceeds 22.75%, then that citation is weighted 0.5 instead of 1. If it exceeds 40%, it is weighted 0.33. This approach does not completely shield us from the highlighted problem; some sections are abandoned over time and cannot be recovered (see Charity from 1971 onwards). However, it allows us to obtain a more stable dynamic. Table A13 and Figure A2 reiterate the analyses in Table 1; Fig. 2 but with the new weighting. The results are almost entirely overlapping with those obtained previously. It is worth considering that changes in newspapers are closely linked to changes in society. Therefore, if editors decide to give more space to certain articles over others, it may be a result of a societal shift, with readers being more interested in specific events and those mentioned in these articles. Thus, the fundamental hypothesis of this article is not lost, namely that those who appear more frequently in newspapers likely have a higher influence in society.

The third robustness check also concerns newspapers. Having access to three different newspapers allows us, at least partially, to reduce distortions caused by the fact that media coverage is not independent of the biases of the newspapers themselves and the preferences of their readers. However, we do not have all three sources for all the years examined, which could introduce additional biases. Therefore, we repeat the configuration of Table 3, but this time using only the “Gazzetta di Modena” as the basis for our data, the newspaper available for all the years considered. The more the results obtained with this setup deviate from the main results, the stronger the observed distortions will be. The results shown in Table A14 in the appendix indicate that the outcomes are consistent with those observed in Table 3, but the RR of surnames belonging to elite groups is slightly higher when considering only the “Gazzetta di Modena.” This might imply that the findings for the years 1921, 1931, and 1941, when we have only one source, could be partially overestimated compared to when we have all three newspapers.

The fourth control concerns the selection of days and months for data collection. Since digitized archives of these newspapers are not available, data were collected manually by extracting surnames from physical articles. This required restricting the survey to specific time intervals. We arbitrarily decided to consider the first three days of each month and the entire month of March. Collecting data for all days of a single month allows for a sufficiently large dataset in terms of observations, while also collecting information in other

**Table 4** RR of rare surnames groups of noble, professors, Jewish and Licei students in newspapers

Year	Noble		Professors		Jewish		Licei Students	
	No	Yes	No	Yes	No	Yes	No	Yes
1921	0.71 (0.56–0.86)	4.36 (1.49–7.24)	0.73 (0.58–0.88)	4.31 (0.62–7.99)	0.77 (0.61–0.94)	4.49 (1.02–7.97)	0.71 (0.55–0.87)	2.34 (1.34–3.33)
1931	0.82 (0.67–0.97)	2.14 (1.25–3.03)	0.75 (0.63–0.86)	9.60 (1.91–17.29)	0.83 (0.68–0.98)	4.52 (1.51–7.53)	0.75 (0.60–0.89)	2.39 (1.61–3.17)
1941	0.83 (0.65–1.01)	2.60 (0.37–4.83)	0.81 (0.64–0.98)	5.31 (-1.54–12.17)	0.87 (0.68–1.06)	1.26 (-0.35–2.87)	0.75 (0.55–0.93)	2.31 (1.50–3.07)
1951	0.72 (0.60–0.83)	1.82 (0.49–3.14)	0.70 (0.59–0.82)	3.85 (0.55–7.23)	0.74 (0.62–0.85)	1.59 (-0.57–3.74)	0.70 (0.58–0.82)	1.29 (0.92–1.66)
1961	0.68 (0.53–0.81)	2.15 (0.71–3.58)	0.62 (0.48–0.75)	5.88 (2.84–8.91)	0.68 (0.54–0.82)	2.43 (0.41–4.46)	0.65 (0.51–0.78)	2.32 (1.17–3.46)
1971	0.84 (0.68–0.99)	2.90 (0.20–5.59)	0.80 (0.64–0.95)	3.37 (2.09–4.65)	0.85 (0.69–1.00)	2.47 (-1.52–6.46)	0.81 (0.65–0.97)	1.78 (1.26–2.29)
1981	0.82 (0.60–1.04)	1.46 (0.03–2.89)	0.77 (0.56–0.98)	3.34 (0.30–6.38)	0.82 (0.60–1.04)	1.58 (-0.52–3.67)	0.81 (0.58–1.03)	1.22 (0.75–1.67)
1991	0.56 (0.47–0.64)	1.50 (0.21–2.79)	0.55 (0.46–0.64)	1.61 (0.82–2.40)	0.56 (0.47–0.64)	0.96 (-0.80–2.71)	0.54 (0.45–0.62)	1.61 (1.04–2.18)
2001	0.56 (0.48–0.63)	1.08 (0.43–1.74)	0.55 (0.47–0.62)	1.49 (0.57–2.40)	0.55 (0.48–0.62)	3.35 (0.18–6.51)	0.53 (0.46–0.60)	1.65 (1.09–2.21)
2011	0.56 (0.47–0.64)	1.06 (0.23–1.90)	0.54 (0.45–0.63)	2.61 (1.26–3.97)	0.56 (0.47–0.64)	1.94 (0.03–3.84)	0.55 (0.46–0.63)	1.39 (0.61–2.16)

Notes: 95% confidence interval in parentheses. Surnames collected in the subsection's thief, victim, accidents are not included. See Table A.8 in the Appendix for surname groups definitions and statistics

months helps balance the potential repetition of individual events that may recur multiple times. However, this does not eliminate the possibility of bias due to events resonating over several days. Therefore, for the years 1921 and 1931, we collect additional data using two different approaches. In the first approach, we collect data for the entire month of October and for three randomly chosen days in each of the other months of the year. In the second approach, we use the “constructed week” method (Riffe et al., 1993), which involves randomly selecting one day for each day of the week (Monday, Tuesday, ..., Sunday) from the study period (in our case, the year). Specifically, we gather 2 constructed weeks for both 1921 and 1931. The closer the results obtained with these two approaches are to the main results, the more robust the latter will be. Table A15, which replicates the setting of Table 3, shows results that, while partially varying - as can be expected - from those recorded by collecting data in March and the first three days of each month, are still strongly consistent with what was previously observed.

## 6 Conclusions

In this study, we propose the use of surnames in local newspapers as a novel data source to explore the social mobility of elites and family generations. The key idea is that individuals frequently cited in newspapers hold social importance, and we aim to investigate whether this higher visibility persists over time.

We first verified that surnames in newspapers are not randomly extracted from the population but reflect traits of social relevance. Subsequently, we analyzed the mobility of surnames in newspapers over time. Surnames belonging to privileged groups exhibit a higher representation compared to other social groups, and this higher RR seems to be transmitted to future generations in a differentiated manner depending on the considered high-status group. In the case of surnames of university professors and Licei students, their representation tends to converge towards the mean, particularly from the ‘80s onwards. The same holds true for noble surnames, while for surnames belonging to the Jewish community, the pivotal moment of change is situated in ‘41. In all cases, there is a process of regression towards the mean in the RR value over the decades, signaling the presence of social mobility. We compared the results obtained from the newspaper dataset with those of Licei students, noting a relatively similar trend between the two information sources. However, newspapers appear to overstate the representation of relevant social groups compared to the Licei one. We also performed controls for rare surnames, reweighting the newspaper dataset, using a single newspaper as the information source, and conducting a different selection of days and months for data collection. All these checks confirm the same basic trends.

Given the nature of the data used, our study presents some limitations. Firstly, we focus on a geographically limited context. In a medium-small reality as Modena, social networks can be tighter and more cohesive compared to metropolitan areas, facilitating access to opportunities for those well-integrated and creating greater barriers for newcomers. As a result, the transmission of socioeconomic status may be more pronounced, with a greater influence of initial conditions. Furthermore, mobility in and out of small contexts can have a significant impact. Although Modena is distinguished by high socioeconomic well-being and substantial growth in prosperity since World War II, individuals seeking greater or different opportunities might migrate to larger cities. The social relevance of a family in the

past might not be captured in our study, not because descendants no longer hold that influence, but because they have relocated elsewhere.

Secondly, the lack of access to different newspaper sources for all the years can lead to distortions because each newspaper has its own editorial biases, attracts readers with specific preferences, and may choose to give greater prominence to certain events or individuals based on these preferences. If only one newspaper is available in certain years, the news and information collected might disproportionately reflect the viewpoints and priorities of that single source. This could result in a less balanced and more partial representation compared to when multiple newspapers offer diverse perspectives. Although we propose a robustness check, we cannot be certain that the lack of multiple sources in certain years does not lead to less accurate results.

Also relying on surnames as pseudo-links to connect families across generations presents some limitations. This method is already widely used and accepted in the literature (Collado et al., 2012; Clark, 2014; Clark & Cummins, 2014; Clark & Cummins, 2015; Barone & Mocetti, 2021). Additionally, focusing on the local level yields more precise links between generations, and the high heterogeneity of Italian surnames enhances the quality of these pseudo-links: the rarer the surname, the higher the probability that it belongs to the same family. However, we acknowledge that it is not possible to distinguish between individuals sharing the same surname who may not belong to the same family. In such cases, these individuals are treated collectively, and the citations of the same surname collected in different years might actually belong to distinct families.

It would be interesting to replicate this exercise with other local contexts to verify if the outcomes are confirmed. According to Acciari et al. (2022), northern Italian territories are characterized by substantial economic mobility; therefore, a comparison with a local reality in different parts of Italy could be very interesting to separate the influence of common historical changes from that of local factors. Moreover, it would be impactful to repeat this approach with digital newspaper archives and using automated data collection techniques.

In conclusion, we suggest that this type of analysis can contribute to identifying different mobility patterns at the local level and represent a useful alternative when established data sources, such as income, education, or data on occupational class, are not available.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11205-024-03415-7>.

**Funding** Open access funding provided by Università degli Studi di Modena e Reggio Emilia within the CRUI-CARE Agreement.

**Data Availability** Data available upon request.

## Declarations

**Competing Interests** The authors report there are no competing interests to declare.

**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/>.

## References

- Acciari, P., Polo, A., & Violante, G. (2022). And yet it moves: Intergenerational mobility in Italy. *American Economic Journal: Applied Economics*, 14(3), 118–63.
- Ager, P., Eriksson, K., Karger, E., Nencka, P., & Thomasson, M. A. (2022). School closures during the 1918 flu pandemic. *Review of Economics and Statistics*, 1–28.
- Aguilar, P., Ghirelli, C., Pacce, M., & Urtasun, A. (2021). Can news help measure economic sentiment? An application in COVID-19 times. *Economics Letters*, 199, 109730.
- Ahmed, A. M., & Hammarstedt, M. (2008). Discrimination in the rental housing market: A field experiment on the internet. *Journal of Urban Economics*, 64(2), 362372.
- Åslund, O., & Skans, O. N. (2012). Do anonymous job application procedures level the playing field? *ILR Review*, 65(1), 82–107.
- Ban, P., Fouirnaies, A., Hall, A. B., & Snyder, J. M. (2019). How newspapers reveal political power. *Political Science Research and Methods*, 7(4), 661–678.
- Barone, G., & Mocetti, S. (2021). Intergenerational mobility in the very long run: Florence 1427–2011. *The Review of Economic Studies*, 88(4), 1863–1891.
- Beach, B., & Hanlon, W. W. (2022). *Historical newspaper data: A researcher's guide and toolkit (no. w30135)*. National Bureau of Economic Research.
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4), 991–1013.
- Black, S. E., & Devereux, P. J. (2010). *Recent developments in intergenerational mobility (no. w15889)*. Institute of the Study of Labor.
- Bonilauri, F., & Maugeri, D. (1999). *Le comunità ebraiche di Modena e Carpi*, Firenze, Giuntina.
- Carocci, G. (2005). *Storia degli ebrei in Italia. Dall'emancipazione a oggi*, Rome, Netwon and Compton.
- Clark, G. (2014). *The son also rises*. Princeton University Press.
- Clark, G., & Cummins, N. (2014). Surnames and social mobility in England, 1170–2012. *Human Nature*, 25, 517–537.
- Clark, G., & Cummins, N. (2015). Intergenerational wealth mobility in England, 1858–2012: Surnames and social mobility. *The Economic Journal*, 125, 61–85.
- Collado, M. D., Ortuño-Ortín, I., & Romeu, A. (2012). *Long-run intergenerational social mobility and the distribution of surnames, manuscript*. Universidad de Alicante.
- Della Pergola, S. (1970). Osservazioni sulla longevità degli ebrei in Italia, in *Rassegna Mensile di Israel*, vol. XXXVI, n. 7-8-9, pp. 165–182.
- Earl, J., Martin, A., McCarthy, J. D., & Soule, S. A. (2004). The use of newspaper data in the study of collective action. *Annu Rev Sociol*, 30, 65–80.
- Fowler, B., & Bielsa, E. (2007). The lives we choose to remember: A quantitative analysis of newspaper obituaries. *The Sociological Review*, 55(2), 203–226.
- Fregni, E. (1993). La comunità ebraica di Modena nelle carte del suo archivio storico (sec. XVII-XVIII), in *Vita e cultura ebraica nello stato estense*, pp. 299–316.
- Ganzeboom, H. B., Treiman, D. J., & Ultee, W. C. (1991). Comparative intergenerational stratification research: Three generations and beyond. *Annual Review of Sociology*, 17(1), 277–302.
- Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1), 35–71.
- Gentzkow, M., Glaeser, E. L., & Goldin, C. (2006). The rise of the fourth estate. How newspapers became informative and why it mattered. *Corruption and reform: Lessons from America's economic history* (pp. 187–230). University of Chicago Press.
- Ghelfi, C., Baracchi, O., & Braglia, M. (1995). *G. La comunità ebraica di Modena*.
- Guell, M., Rodriguez Mora, S., & Telmer, C. I. (2015). The information content of surnames, the evolution of intergenerational mobility, and assortative mating. *The Review of Economic Studies*, 82(n. 2), pp693–735.
- Hansen, P. (2004). Writing business history without an archive. newspapers as sources for business history – possibilities and limitations, in Gadd, C.-J., Granér, S. and Jonsson, S. (Eds.), *Markets and Embeddedness, Essays in Honour of Ulf Olsson*, pp. 99–120.
- Larcinese, V., Puglisi, R., & Snyder Jr, J. M. (2011). Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers. *Journal of Public Economics*, 95(9–10), 1178–1189.
- Levi, F. (1997). *Gli ebrei nella vita economica nell'800*, in *Storia d'Italia, Annali 11, gli ebrei in Italia* (pp. 1171–1214). II, Einaudi.
- Marach, I. M. (1999). *Demografia E composizione familiare nella comunità Ebraica Di Modena fra '800 e '900*, in *Le Comunità* (pp. 177–184). ebraiche a Modena e a Carpi.

- Markel, H., Lipman, H. B., Navarro, J. A., Sloan, A., Michalsen, J. R., Stern, A. M., & Cetron, M. S. (2007). Nonpharmaceutical interventions implemented by US cities during the 1918–1919 influenza pandemic. *Jama*, *298*(6), 644–654.
- Matsuda, M. (1996). *The memory of the Modern*. Oxford University Press.
- Mooi-Reci, I. (2020). Intergenerational mobility research: Current challenges and future directions. *Australian Economic Review*, *53*(2), 241–246.
- Pascual, A., Guéguen, N., Vallée, B., Lourel, M., & Cosnefroy, O. (2015). First name popularity as predictor of employability. *Names*, *63*(1), 30–36.
- Pietramellara, M. (1902). *Blasonario generale italiano, ossia Descrizione degli stemmi delle famiglie nobili e titolate d'Italia*. Dispensa 6: del modenese. Direzione, Roma.
- Riffe, D., Aust, C. F., & Lacy, S. R. (1993). The effectiveness of random, consecutive day and constructed week sampling in newspaper content analysis. *Journalism Quarterly*, *70*(1), 133–139.
- Shapiro, A., Buckman, S. R., Sudhof, H., & Wilson, M., D. J. (2020). News sentiment in the time of COVID-19. *FRBSF Economic Letter*, *8*(1), 5–10.
- Silberzahn, R., & Uhlmann, E. L. (2013). It pays to be Herr Kaiser: Germans with noble-sounding surnames more often work as managers than as employees. *Psychological Science*, *24*(12), 2437–2444.
- Solon, G. (1999). Intergenerational mobility in the labor market. *Handbook of labor economics* (Vol. 3, pp. 1761–1800). Elsevier.
- Solon, G. (2018). What do we know so far about multigenerational mobility? *The Economic Journal*, *128*(612), F340–F352.
- Tosh, J. (2010). *The pursuit of history: Aims, methods, and new directions in the study of modern history* (5th ed.). Pearson Longman.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.