



The origins of Italian human capital divides: new evidence from marriage signatures, ca. 1815

Marco Martinez¹

Received: 16 December 2022 / Accepted: 5 May 2023
© The Author(s) 2023

Abstract

This paper provides new evidence on the geography of literacy rates in pre-unitary Italy. I provide direct estimates based on a novel and balanced random sample of marriage certificates in 1815. The new figures are disaggregated by gender, area, and skill levels, and are compared to the relevant alternative estimates available. Literacy rates are generally low, and North versus South difference in 1815 literacy was as low as 13.7 percentage points, with the Southern literacy rate of about 50% that of Northern Italy. The North–South gap is much smaller for women than for men, and the average female literacy rate across Italy is a lower 9%. Literacy rates in Central Italy were almost identical to that of Southern Italy, arguably because the schooling systems of Central and Southern Italy were more elitist than the Northern Italian one. This evidence suggests that, although partially present also before 1815, the wide magnitude of North–South gaps in literacy which characterized the country on the eve of the political unification (1861) originated after the Napoleonic period. Primary school centralization reforms might have helped women to rapidly improve literacy rates, leading to a first, regionally unequal, ‘Silent Revolution’ (Cappelli and Vasta in *Cliometrica* 15:1–27, 2020a).

Keywords Literacy rates · Random sampling · Marriage certificates · Human capital

JEL Classification N01 · N33 · J24

✉ Marco Martinez
m.martinez@santannapisa.it

¹ Institute of Economics, Sant’Anna School of Advanced Studies, Piazza Martiri della Libertà 33, 56127 Pisa, Italy

1 Introduction

Traditional views of nineteenth-century Italian economic history consider the country before the Unification as a land of widespread illiteracy and unskilled agricultural labor, especially in the South.¹ Economic historians have provided convincing evidence that in terms of economic performance and living standards, Italy was already characterized by a significant dualism at the time of the Unification, whether measuring it through real wages (Federico et al. 2019a, b; Malanima and Daniele 2007), life expectancies and heights (A’Hearn and Vecchi 2017), social capital and culture (Putnam et al. 1994), GDP (Felice 2014; Malanima and Daniele 2007; Daniele and Malanima 2011), Human Development Index (Felice and Vasta 2015), or income inequality (Felice 2019).²

The first unitary census already showed evident North–South literacy gaps: in 1861, the census literacy rates were 27.28% in the North, 17.19% in the Centre, and 10.14% in the South (MAIC 1866b, XXII–XXV). Literacy rates have proved important in explaining the evolution of key economic parameters after the Unification (manufacturing: Basile and Ciccarelli 2018; schooling: Cappelli and Quiroga Valle 2020; labor productivity: Ciccarelli and Fachin 2017; rate of invention: Nuvolari and Vasta 2017; real wages: Federico et al. 2019a, b). But the lack of systematic data for the period before the political unification (1861) makes it difficult to attribute the role of specific events to the origins of the Italian regional divides. Gathering evidence on the origins of the divides would also be important to understand the nature of the underlying mechanisms of further divergence.

A conventionally held view is that from the mid-eighteenth century onward, Italian absolutist rulers attempted, with some success, to start institutional reforms inspired by the Enlightenment, and that Napoleon built upon their efforts by initiating a process of economic growth (Dincecco and Federico 2021). Nonetheless, the cost of French dominance to Italian states in terms of taxes, trade, and potentially even military conscription was high (Dincecco et al. 2011). This brings into light the question of whether the benefits of institutional reforms outweighed the costs. Dincecco and Federico (2021, 23) suggest that the Napoleonic reforms were important to ‘clear out the institutional undergrowth left over from the *Ancien Régime*, but that they were not in themselves enough to provide Italy with modern growth-enhancing institutions’ and that the real turning point was the reforms enacted in Piedmont after the military defeat by Austria in 1848–49 and later extended to the rest of Italy. Other studies have attributed an important role for the origin of literacy divides in the late eighteenth-century to the Habsburg Austrian institutional reforms in the Northern region of Lombardy (Toscani 1993), or to the different responses to public schooling reforms after the Restoration of 1815 (Lupo 2005, 141–175), or to

¹ This view of a stagnant Italian economy after the end of the seventeenth century became prevalent after the publication of the essay by Cipolla on the ‘Decline of Italy’ (Cipolla 1952) and of Braudel’s *La Méditerranée* (Braudel 1949), and was later supported by Cipolla (1989). Aymard (1978) offers a revisionist approach that is discussed in Zamagni (1993, 6–12) and Malanima (2005).

² The social capital explanation that attributes the origins of the North–South divides to as far as the Middle Ages has been questioned with alternative (Cappelli 2017) and complementary (A’Hearn 1998; Mariella 2022) explanations.

the particularly conservative reaction of the Bourbons in the Kingdom of the Two Sicilies after the various insurrections of 1820–21, 1830–31, and especially of 1848 (Felice 2013, 30–38). In this respect, Postigliola and Rota (2020, 2021) document a significant effect of the duration of the French period—longer in Northern Italy than in Southern Italy—on 1861 literacy rates obtained from unitary censuses and on social capital. These explanations are not necessarily in conflict with each other. However, to make advancements in this important debate, there is an urgent need for new systematic data collection efforts aiming at measuring the extent of regional divides before the unification.

In an attempt to overcome the data limitations that pervade pre-unitary Italian sources, this paper offers new and direct estimates about the literacy rates at a crucially important moment in time: right after the first State-driven schooling reforms were enacted in Austrian Lombardy and Borbonic Kingdom of Naples, that is, after the 1770s, but before the Napoleonic educational reforms of the 1802–1815 could show their effects. In this relatively neglected period of Italian history, we can see the premises for the earliest State-driven human capital efforts undertaken in Italy, and related path dependency in later educational and social outcomes.

I offer a systematic account of literacy rates from all the sources that, to my knowledge, are available to comprehensively study literacy rates in pre-unitary Italy. The main data source is a novel and representative cross-sectional sample of 1216 individual-level marriage certificate signatures dating to 1815. I randomly select villages with marriage certificates in the Antenati (www.antenati.san.beniculturali.it) online genealogical archives following Schofield (1973), who, 50 years ago, first attempted to scale up the literacy information gathered from more than 200 local parishes to the level of England. *Antenati* gathers birth, death, and marriage records from Italian State archives. The sample (henceforth ‘*Antenati*’ sample) is constructed to have representative literacy rates at the level of Italian provinces, regions and macroarea (North, Centre, and South), similarly to Rossi et al. (2001) for household budgets and Shaw-Taylor and Wrigley (2014) for occupational structure. I account for under-sampling issues with weighting and regression approaches. This allows for a safely balanced sample that avoids the underrepresentation of some areas, which also divides the literacy rates of each area into three main occupational sectors (services, industry, and agriculture). From marriage certificates, I collect information about the ability to sign of spouses, but also about their occupations, and about parents’ occupations. As an additional step, I collect further data from selected parts of Italy. The first is a novel dataset containing all the available signatures, rather than a sample thereof, for a large area of the provinces of Udine, Salerno, and Abruzzo Ulteriore II. The second is a sample of eighteenth-century marriage signatures (from *Processetti* matrimoniali) for the city of Naples. I also collect literacy rates from studies who focused on specific parts of Italy in and around 1815 (also referred to as ‘local studies’). Finally, I consider unitary censuses and related secondary literature.³

³ Local studies include Scirocco (1987), Toscani (1987), Piseri (2002), Toscani (1993), Brambilla (1991), Milanese (1991), and Ferraresi (1991). I do not compare my study only with studies that either focused on later periods (study only with studies that either focused on later periods i.e., Griseri 1973)

The *Antenati* sample of signatures across Italy well reconstructs the literacy that was documented by local studies and indirect approaches in the period of interest. The levels of literacy rates are low throughout Italy, and not only in the South. Overall, I find a sizeable North versus South difference in literacy rates of 13.7 percentage points, whereas the previous estimate by Ciccarelli and Weisdorf (2019) was 20 percentage points. However, in terms of levels, the level of literacy in the South was about 47% of the North one. This estimate is similar to Ciccarelli and Weisdorf (2019) since in our estimates, the levels of literacy rates are lower both in the North and in the South. Compared with the previous literature, probably the most important result is the particularly low level of literacy rates of women (9.2%). The wealthiest parts of Southern Italy had particularly high literacy rates: Naples and its province had literacy rates comparable to those of Lombardy in the North. The rich information of marriage certificates also enables us to look at new dimensions of illiteracy. For instance, the urban–rural difference was 18 percentage points and the gender difference was of 22 percentage points for an average male literacy of 31%. The North–South gap was particularly high among peasants, and the literacy rate of Northern Italian peasants was similar to the average literacy in Central and Southern Italy.

This paper contributes to the following streams of research. First, it advances the literature on the schooling system and literacy in Italy through a direct approach to the study of pre-unitary origins of Italian North–South divides. This complements the existing studies that derived information from small areas of Italy or unitary censuses with new and rich individual-level data, which also disaggregate literacy rates by skill levels, at a time when first state-driven educational efforts started.⁴ Second, this study provides an operative methodology to safely derive countrywide estimates of literacy rates from marriage certificates. This is useful for scholars who wish to learn about supply of human capital in other countries at their early stages of industrialization, when often no census data is available (e.g., Meier zu Selhausen et al. 2018; Álvarez and Palencia 2018).⁵ I also corroborate the empirical validity of this exercise with a particularly rich comparison with other sources. Third, by placing the results in the historical context of early nineteenth-century schooling reforms,

Footnote 3 (Continued)

for Piedmont in 1831–1856), earlier periods (i.e., Pellizzari (1987) for mid eighteenth-century Kingdom of Naples) or on areas not collected in this sample as Sardinia (Pruneri 2011). Considered official statistics are the estimates of Ciccarelli and Weisdorf (2019) for the period 1821–1881, and the 1861, 1871, and 1881 unitary censuses: MAIC (1866a, 388–411), MAIC (1875, 4–239), MAIC (1883, 217–564). I consider individuals aged 18 or more in the 1861 census, rather than aged more than six years old as Cappelli and Quiroga Valle (2021), to better match the age structure of *Antenati* spouses. I also consider other proxies of literacy after the unification derived from Ministero dell’Istruzione (1890).

⁴ Examples of such studies for Italy are Ciccarelli and Weisdorf (2019), Postigliola and Rota (2020) and Bozzano, Cappelli and Vasta (2022).

⁵ The inherent difficulty in collecting often handwritten data from marriage registers scattered over the country has discouraged the use of marriage certificates to derive direct, countrywide estimates of pre-industrial literacy rates from marriage signatures, a notable exception being Schofield (1973).

this paper also indirectly relates to the literature on the role of nineteenth-century school reforms for human capital formation and modern economic growth.⁶

The remainder of the paper proceeds as follows. Section 2 introduces the reader to the context of the history Italian education over the period 1770–1861. Section 3 describes the sources. Section 4 describes the research design and discusses its validity. Section 5 presents the main results about literacy rates and places them in context. Section 6 contains a conclusion.

2 The evolution of literacy, 1770–1861: historical background

This section covers a brief history of the educational policies enacted in Italy from the 1770 to 1861 and can be useful for the reader to adequately place the historical results of this paper in context.⁷ Until the 1770s, the absolutist governments of the many Italian states did not conceive education as a state affair. Most schools were privately established where the demand was the highest. In the North, primary schools were demanded especially by small landowners, manufacturers, and the ones in need of skills to emigrate (Toscani 1993). Teaching was catered to individual students rather than classroom-based and teachers were ecclesiastical (Genovesi 2010). The so-called enlightened despots gradually became aware of the importance of economic growth for political legitimacy and military power. Even the government of catholic rulers as Maria Theresa of Habsburg-Teschen reduced the feudal rights (i.e., through taxation), and ecclesiastical privileges (i.e., in education provision). Relatedly, the members of the Company of Jesus were expelled from various European States.

Jesuits had been central providers of higher education in Italy since the foundation of the first Jesuit College in Messina in 1604. In many states, the Jesuits continued to influence the intellectual life even after the expulsion, as in the reformist Habsburg Empire. But the Jesuit expulsion spurred a general tendency to find alternatives to religious education also at the level of primary education. A variety of solutions were offered, including the decision of the Papal State to keep the education entirely in the hands of ecclesiastical teachers; intermediate forms that handed education to lay orders, such as the *Scolopi* in the Grand Duchy of Tuscany; the first attempts to have a public education system in the Kingdom of Naples of Ferdinand IV of Bourbon and in the Lombardy ruled by emperor Maria Theresa of Habsburg-Teschen. In the Kingdom of Naples, the diffusion of new collective teaching methods (normal schools) and new public primary schools was effective in metropolitan Naples and neighboring areas (Terzi 2001; Zazo 1927), but less so in rest of the kingdom. According to Lupo (2005), the unsuccessful diffusion of primary schools

⁶ Starting from the pioneering work of David Mitch (1986), many studies adopted quantitative approaches assess various effects of the early schooling reforms, emphasizing the role of state-driven schooling centralization and political economy considerations. Recent examples include Cappelli and Vasta (2020b) for Italy; Montalbo (2020) for France; Tapia and Martinez-Galarraga (2018) for Spain; Cvrček and Zajicek (2019) for Austria; Cinnirella and Hornung (2016) for Prussia; Andersson and Berger (2019) for Sweden; Milner (2021) for England; Chen et al. (2015) for China.

⁷ See Piseri (2017, 21–60) for an excellent introduction.

elsewhere in the South was due to insurmountable obstacles posed by local noblemen and clerics (Lupo 2005). This explains the exceptionally high levels of literacy rates in Campania, where Naples is. The reforms enacted under Maria Theresa were also successful in urban centers, but less so in rural centers (Toscani 1993). Importantly, Maria Theresa also suppressed the religious orders to fund public and—ideally—mandatory primary schools. The political government of her descendant Joseph II of Habsburg-Lorena specifically targeted areas with fewer private schools. Toscani (1993) observes that in the agrarian region of *Bassa Padana*, the schools were now many, but the demand for schooling remained low, as did literacy.

The first phase of French domination (1796–1799) was too short-lived for the invaded states to adopt a consistent set of reforms in line with the revolutionary ideologies, although the French imposed very high tributes to all conquered regions (Dincecco and Federico 2021, 7). The second phase of French domination (1801–1815) saw a gradual incorporation of all countries of pre-unitary Italy, excluding Sicily and Sardinia, to what would soon become the French empire in 1804.⁸ This time, the domination was also political, and was exercised directly by Napoleon Bonaparte, Bonaparte family ties, or close allies. The Napoleonic government introduced radical reforms to various aspects of the civil and administrative system in all the Empire, creating what is sometimes referred to as ‘administrative monarchy’ (Benigno 2014). The reforms came at the cost of higher taxation relative to *ancien régimes* because it also became necessary to fund the French and Italian military forces (Dincecco and Federico 2021, 7–8). This radical institutional change was also reflected in the design of public schooling policies: from 1802 on, various laws and decrees aimed at centralizing the education around the State, which determined the syllabi, imposed free and mandatory attendance to primary schools, prepared and funded teachers, and directly controlled private, often ecclesiastical, schools. The administrative uniformity of the Napoleonic period implied that the public schooling reforms of the time were eventually imposed throughout continental Italy. And yet Vigo (1971) stressed how only the states that were not under the direct control of Napoleon—the Kingdom of Naples and especially the Cisalpine Republic, which included Lombardy—could have the necessary public autonomy to create a coherent set of legislation that was tailored to the local issues. Yet, in Lombardy public expenditure in schooling varied considerably by department (Piseri 2017, 105–119, 139–142, 222–236). However, in the Kingdom of Naples even the larger autonomy was insufficient: the first set of reforms of 1806 proved to be ineffective until central authorities gained reach over the more remote provinces with the *Decreto Organico della Pubblica istruzione*.⁹

The most evident effect of these reforms was to centralize and boost public education provision, although with important regional variations. Indeed, institutional centralization did not mean that funding was also completely centralized. The central government only had the duty to inspect the schools, probably because of the crisis of the public finances that followed the collapse of the absolute monarchies.

⁸ Postigliola and Rota (2020) provides a precise timeline of the incorporations in the French Empire.

⁹ For a detailed description of every legislative measure enacted in the Kingdom of Naples in the Napoleonic decade (1806–1815), see Lupo, Gargano, and Marra (2014).

Moreover, while the number of schools increased, this did not necessarily go hand in hand with a shift in the public perception of primary schools as having different features from ‘ancient’ schools, influenced by religious beliefs, local élites, and parental choices.¹⁰

After the Restoration (1815) every Italian state at least formally reverted to the administrative systems of *ancien régimes*, with the notable exception of the newly formed Kingdom of Two Sicilies composed of the Kingdom of Naples and the island of Sicily (see Appendix Fig. 9a). Arguably to reduce popular support for revolutionary ideas, taxation was kept low (Dincecco and Federico 2021). As in many other fields such as taxation of land (Felice 2013), the Borbonic king remained idle in terms of creating new primary schools: Gargano (2012) documents a fall in the number of *public* male schools (from 2498 to 822) and a fall in public female schools (from 1556 to 1142). The number of *private* schools also saw a decrease. Alternatives to public schools (i.e., evening schools, adult schools) were concretely diffused only after 1861 (Vigo 2017). Meanwhile, the monarchic response to the revolts of 1820–1821 was repressive and gave back a prominent role to the clergy in public education, in line with the conservative attitude that other Italian States had in 1815 (Lupo 2005).

In Northern and Central Italy, except in the Papal State, the role of the State in the public education system was gradually re-introduced. Indeed, the new absolutist governments could not restore the set of the privileges that their aristocratic title or ecclesiastical status entitled them to before the Napoleonic period. As a result, already in 1818, the Austrian Empire had extended the public schooling system to *Lombardo-Veneto*, for boys and girls alike. In *Savoia* (today’s Piedmont), the schools remained in the hands of the clergy until the Boncompagni law 1848. The Boncompagni Law re-affirmed the central role of the state in the education system by controlling also private and ecclesiastical schools, and was the steppingstone of the 1859 Casati Law, later imposed in unified Italy. Instead, the legislative efforts of the Bourbons in the South were increasingly dedicated to more *élite* secondary and higher education, and the response to the 1848 insurgency was again repressive (Lupo et al. 2014).

3 Sources

The main original source of this study are pre-unitary marriage certificates derived from civil records. In Italy, historical civil records are scattered across more than a hundred different archives, making it difficult to conduct comprehensive studies using these sources. However, a government project, *Antenati*, started to scan civil registration records from the State archives (Antenati 2023). In this study, I transcribe a

¹⁰ Banti (1996, 55–97) and Macry (2002) show how noblemen, landowners, small business owners, and professionals in the North and in the South found ways to keep the family wealth concentrated that resembled the ones prohibited by Napoleonic Code on succession (in particular, the *majorat*) up to the late nineteenth century. It is possible to imagine that families retained a similar unchanged attitude towards education despite the institutional reforms.

random sample of scanned certificates from the *Antenati* genealogical platform. Of the 137 Italian State archives, 85 contain marriage certificates for the period of interest. Of these, 31 are still awaiting digitization, so I considered the certificates from the 54 State archives. The fact that the digitization effort is still ongoing could raise concerns about whether it is not premature to use this source. Yet, most archives with pending digitization are minor ones or sub-branches of larger archives that also contain marriage certificates for their area (i.e., Assisi, Chieti, Pistoia, Massa, Frosinone). The data is structurally unavailable in large areas, either because the certificates do not have signatures (i.e., in Genova) or because marriage certificates are located elsewhere, such as in parishes and civic courts (i.e., Milan, Nuoro).¹¹

These instances will not be solved even when the digitization of State archives is complete: even having the full sample from *Antenati* would not give a fully representative picture of Italy, if only because the entire regions of today's Latium except the southernmost provinces, Sardinia, Umbria, and the city of Milan are not covered. This raises the possibility that the sample is not representative of Italy as a whole. This would be the case if, for example, the part of Italy not covered by the sample is disproportionately urban or rural, as urban and rural areas are very different in terms of literacy rates. When the composition of covered and non-covered parts of Italy is compared in terms of urban–rural shares in 1815 and literacy rates in 1861, the composition is very similar both in terms of urban–rural shares and of literacy in 1861 (see Tables 8 and 9).¹² When the composition differs, the composition of the sample is much closer to the overall one than the part of Italy which could not be sampled is. The composition is also similar by macroarea (North, Centre, and South). The resulting North–South gap would have appeared smaller if the sample had been biased upward for Northern Italy and downward for Southern Italy. This evidence provides some reassurance that no significant sampling bias, along these dimensions, exists between or within macroareas. As a further sensitivity check, I looked at how literacy rates change by macroarea if assuming that the areas not covered in the sample have the same literacy of the bordering provinces covered on *Antenati*, and the resulting literacy rates are very close to the literacy rates of the *Antenati* sample (see Table 10). Further research can later be extended to the remaining archives, and the sampling approach presented here serves this purpose, although issues with data preservation make it clear that for some areas there will still be no marriage certificates with signatures in the future (see next section,

¹¹ To avoid underestimations caused by the absence of the large urban areas of Genoa and Milan, large urban areas are proxied by reasonably similar cities for which certificates are available, such as Turin and Venice, and missing observations from urban areas are proxied by similar observations from neighboring provinces (see Appendix Sect. A3 for further details).

¹² In the *Antenati* sample, the North of Italy is composed of the regions of Liguria, Piedmont, Lombardy, Venetia, and Emilia-Romagna. The Centre of Italy is composed of the regions of Tuscany and Marche. The South of Italy is composed of the provinces of Terra di Lavoro, Abruzzo Ulteriore I, Abruzzo Ulteriore II, Abruzzo Citra, Basilicata, Calabria Ultra I, Calabria Ultra II, Catania, Girgenti, Messina, Molise, Napoli, Noto, Palermo, Principato Citeriore, Principato Ulteriore, Terra d'Otranto, Terra di Bari, and Trapani. Appendix Table 13 shows covered and non-covered parts of Italy.

Figs. 3, and 7 in Appendix 1 for the spatial coverage of the *Antenati* source).¹³ I will refer to the sample of marriage certificates drawn from this source as the ‘*Antenati* sample.’ Besides marriage certificates, the paper also uses datasets derived from the published census reports of 1861, 1871, and 1881, and of other official documents.¹⁴

The ability to sign marriage certificates is used as the proxy for literacy rates. The ability to sign can be indicative of functional literacy adjusted to the standards of the early nineteenth-century Italy.¹⁵ Marriage certificates in Italian civil records also provide information about occupations, ages, and places of residence of spouses, their parents, and of the witnesses. Moreover, the signatures of the spouses in marriage certificates suffer less selection bias than other sources, such as the signatures of witnesses in death records or the signatures of sailors’ boarding certificates (Rediker 1989) because every strata of the population married,¹⁶ and spouses did not have particular reasons to learn how to sign compared to the rest of the population.

The earliest attempt to measure signature-literacy is probably Sargant (1867), although important examples are also Cipolla (1969) and Schofield (1968). And yet Astoul (1992, 182–185) raised concerns about measuring literacy rates with marriage signatures (*le methode ‘Maggiolo’*) in the French *Midi* (see also Fleury and Valmary 1957). His main concern was that often, parish priests were not keeping appropriate track of all marriages and of their signatures.¹⁷ Poussou (1993, 214–215) also admitted that the results can be erroneous in some parishes and called for a check of the methods by collecting more signatures in specific areas.¹⁸ Additionally, comparing literacy rates derived from signatures with stated literacies in censuses or conscript literacies may not be legitimate if they measure different dimensions of literacy. In Italy, literacy inferred from marriage records in 1872 matches census

¹³ Looking at Fig. 7, one may argue that the *Antenati* source overrepresents urban areas. To ensure that the *Antenati* sample accurately represents the population, the sampling is therefore conducted with the help of population weights (see Sect. 4 and Appendix A3).

¹⁴ Occupations were coded according to the HISCO and HISCLASS classifications of Van Leeuwen (2002) and Van Leeuwen and Maas (2011). The official sources I use are MAIC (1866a, b); MAIC (1875); MAIC (1883); MAIC (1900), and Ministero dell’Istruzione (1890). I also refer to the elaborations of census data of Ciccarelli and Weisdorf (2019).

¹⁵ If a person can sign a certificate, this does not mean that she can access book production or the popular press (see Marchesini 1992). There are sources that enable measurement of more complex literacy skills, such as receipts of payments but these sources are scattered and hardly systematic (see Petrucci 1987). I classified trembling signatures as ‘poor handwriting.’

¹⁶ In 1851 there were 151,227 religious workers, 130,003 mariners, and 260,209 beggars (*mendicanti*) out of a population estimated to be 10,729,539 (ASI 1853, 68–71). Religious workers and mariners are likely to be more literate than the average population, whereas beggars are less likely to be literate. Because the two groups are similar in size, we should not expect significant differences in literacy between included and excluded individuals.

¹⁷ ‘*Ces deux exemples illustrent la negligence de maints cures quin ne rédigent pas les actes au fur et à mesure, ce qui les conduit à une multitude d’oublis [...] il est bien evident que les paroissiens ne pouvaient pas signer un register où les actes n’étaient pas rédigés!*’ (Astoul 1992, 185).

¹⁸ For similar attempts in other countries, see Pellizzari (2000) and Houston (2014) for a review.

self-reported literacy for the 1871 census very closely.¹⁹ This suggests that the use of marriage certificates for estimating literacy rates is warranted in this particular historical setting. In 1815, Italy was still pre-industrial, and thus being able to read and write was probably less valued as an important skill by the population than when signing was needed to access better occupational positions, as in more industrialized countries. However, even in 1815, the ability to sign accurately predicted the socioeconomic status of spouses (see Figs. 13 and 14 in Appendix 2).²⁰

I focus on 1815 also because it is the only year providing us with a comprehensive geographical picture of literacy based on signatures. Indeed, civil registries were abolished in most Central and Northern Italy after 1815. Parish registers regained the responsibility of recording births, deaths, and marriages, as with the *status animarum* and recorded information according to the Napoleonic format, but the most records are scattered and difficult to access. Civil registrations were instead maintained in the Kingdom of the Two Sicilies and in the Duchy of Modena and Reggio; for the latter, only remnants of Southern Liguria were digitized before 1860. Civil Registration was also introduced in Sicily in 1820. This short description of the sources covered by Antenati clearly suggests the advantage of focusing on 1815 because this would allow to combine information from Napoleonic records for Central and Northern Italy and Restoration records for Southern Italy. Figure 1 shows an example of a Napoleonic marriage certificate. Only the format of Restoration civil records is standardized and based on a pre-printed format.

I collect Restoration-period records for Southern regions for years as close as possible to 1815 (usually 1816) after the end of the Napoleonic period. For Northern Italy, with only certificates dating back to the Napoleonic period, I collect information that is as close as possible to 1815 (usually 1814) to exclude years where the novelty of the civil registration system may have affected the number of certificates, and to have a sample with comparable years. Excluding second marriages, the mean age in the sample is 26.76 years old for grooms and 23.25 years old for brides, in line with Rettaroli (1990).

¹⁹ A'Hearn et al. (2022, 119) document that 'province-level correlations between census literacy [in 1871] and marriage register signature rates, school enrolment rates 10 years later, and conscript literacy test results are in the range +0.93 and +0.96 for both men and women.' In the first three unitary Italian censuses 1861, 1871, and 1881 the enumerators asked if individuals were able to at least read, to read and write, or were illiterate. Each family head received *schede del censimento* from enumerators (Mastroluca and Verrascina 2012). The illiterate family head had to find a literate individual among neighbors to fill out the forms.

²⁰ Figure 13 shows that, when broken down by occupational category, some occupations have exceptionally high literacy rates. This is especially true for unskilled workers in the Centre. Figure 14 shows that this is primarily due to Central Italy's smaller sample size compared to North and Southern Italy. The sample cannot be stratified by occupations because they cannot be observed beforehand when collecting certificates, as explained in greater detail in Sect. 4.1. Thus, in order to account for unavoidable sampling imbalances like this one and safely disaggregate literacy rates by occupation categories, I propose a regression approach (Sect. 4.2).

4 Research design

4.1 Sampling strategy

I adopt a clustered random sampling scheme (Som 2020, 115–124). The clusters are ‘villages.’ Villages are often municipalities, but in some cases, they correspond to smaller areas or even parishes. I randomly sample villages within the list of available villages on each of the 49 State archives of *Antenati*, and collect marriage certificates in the selected villages. To illustrate the approach, I will use Naples as an example. On *Antenati*, for Naples, there are 18 neighborhoods in the city and 71 villages in the province on the *Antenati* platform. I randomly sample neighborhoods and villages from the list of neighborhoods and villages. Then, to determine the number of certificates to be collected in each village, I consider the fact that according to sample size calculations (see Appendix A2), the minimum sample size is 1216 certificates. The population figure of the province of Naples in 1815 is 667,389 (see Mariella et al. 2020 and supplemental material). The overall population of the portions of Italy covered in this sample is 16,916,109. Converting the population share of Naples to the corresponding marriage rate using the best available figures of marriages per inhabitant (see Appendix A3), Naples must have at least $0.0372 \times 1216 = 45.34$ certificates.

One may wonder why, instead of directly sampling marriage certificates, marriage certificates should be drawn at random within villages. The reason is that on the platform, the most disaggregated unit of observation is pictures containing marriage certificates, rather than certificates themselves, and there is no one-to-one relation between pictures and certificates: depending on the selected village, a picture might contain several marriage certificates (e.g., Fig. 1) or only a part of a single marriage certificate that develops over more pages (e.g., Fig. 2). A clustered random sampling scheme solves the resulting risk of oversampling areas with more than one marriage certificate per picture because it samples units (villages) that are more aggregated than marriage certificates.²¹ In the sampling procedure, I use two main weights to post-stratify the obtained sample in an iterative fashion to have enough observations in each gender, region, and urban–rural cell.²² The first weight is the one indicating that, for example, at least 45 certificates should be collected in Naples; it is composed

²¹ Schofield (1973) cannot observe beforehand how many marriage certificates are contained in each set of parish records, and therefore how to construct a weighted sample of the population, so inference is valid only at the parish-level. Here I can make a step forward because observing the total number of marriages within an archive allows to obtain a balanced sample of marriage certificates at the archive level. Also Cappelli and Vasta (2020b) use a stratified sample of socioeconomic and institutional features of municipalities to study the impact of the Daneo-Credaro schooling reform of 1911 on municipal-level literacy. Another advantage of the clustered sampling scheme is that it solves the issue of there being new certificates introduced on the platform as time goes by, because it considers villages as sampling units, and if a village is available on the Antenati platform it means that the collection is complete for that village. This means that even though the online source is updated over time, my sampling frame does not change as it is based on the list of 5,874 villages with full marriage certificates available in the platform at a specific point in time (December 2022).

²² For a complete illustration of the procedure, including technical details, a flowchart, and the example for Naples, see Appendix Fig. 9.

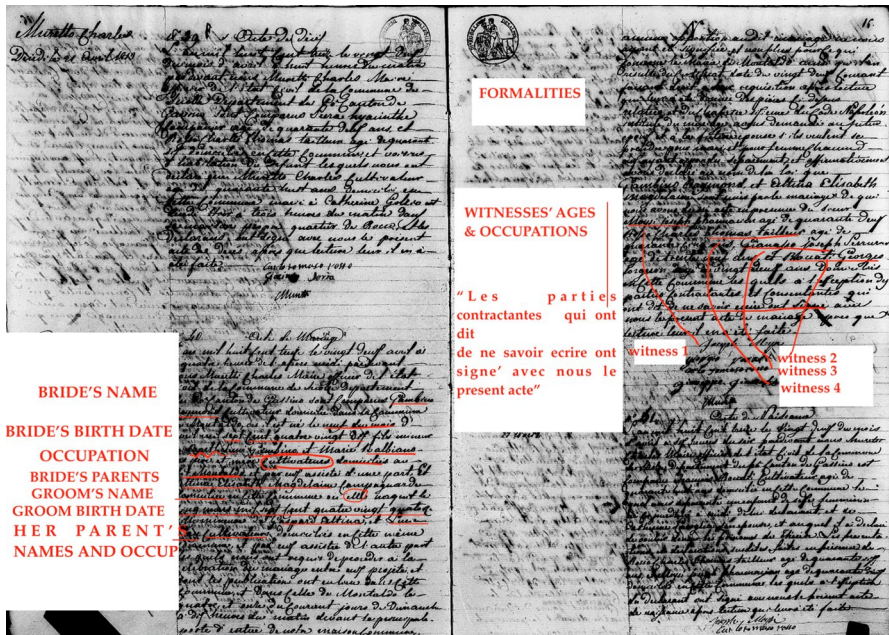


Fig. 1 Marriage certificate from Sciolze, Piedmont (1814). *Content:* The groom’s name is Raimondo Gambino, born in 20 April 1790 (so he is 24 years old by the day of marriage, the 22 April 1814); his occupation is *coltivatore* (peasant) and so are his parents. The bride Altina Elisabetta was born in 05 March 1794, so she is 20 years old at marriage. She is a peasant and so are their parents. The witnesses’ names are Giuseppe Maissi, Giorgio Bloccati, Giuseppe Giovanni Stoffa, and Carlo Tommaso Rosso. Only the witnesses could sign

of a population weight, which is the total sample size divided by the regional population weight of the area adjusted by their marriage rate, and an archival weight, which indicates the minimum number of certificates to collect in each archive over the total of the region. The weight corrects for between-region variability in archive coverage with regional and provincial population estimates for 1815 drawn from Mariella et al. (2020), whose data reconstruction can be considered the most updated and complete.

When the Antenati archive does not include all provinces within a region, the first weight stratifies by the population of unitary regions, and when the data collection is more complete, the first weight stratifies by the smaller level of pre-unitary regions (see Appendix A3 for more details). The second weight is an urban–rural weight and corrects for the share of urban areas in the region; the population was derived from the 1800 estimations of Malanima (2016). I introduce an indicator to distinguish cities from agricultural agglomerations (‘agro-towns’) and only include cities in the urban population (Blok 1969; King and Strachan 1978; Salvo and Zichichi 2003; Curtis 2013). I consider 9000 in 1800 as the minimum population threshold for larger towns in 1815, and 4000 as minimum threshold for smaller towns. I also assume that when town size in 1800 is smaller than 5000, it is 4000. For Naples 62% of certificates should come from urban areas and 15% should come from rural

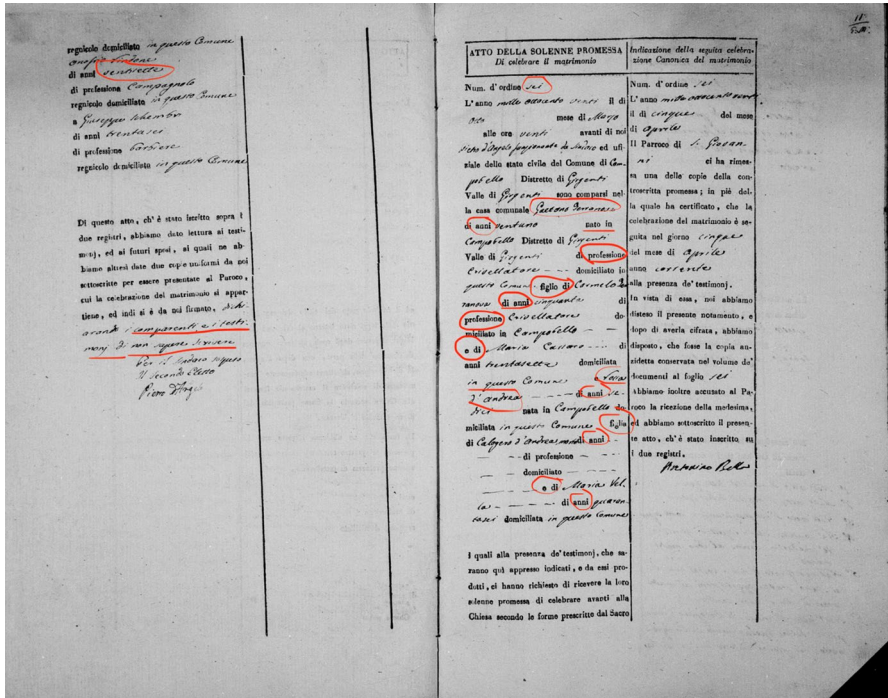


Fig. 2 Marriage certificate from Campobello, Sicily (1820). *Content:* The spouses, their spouses and the witnesses of marriage no. 5 (left side) have declared of not being able to write. In marriage no. 6 (right side), the groom is *Gaetano Ferraresi*, aged 21, a *crivellatore* (miner), residing in Campobello, son of *Carmelo Ferraresi*, aged 50, a *crivellatore*, and *Maria Cassaro*, aged 37, of unspecified occupation. The bride is *Rosa d'Andrea*, aged 16, resident in Campobello, of unspecified occupation, son of *Calogero d'Andrea Cassaro*, deceased, and *Maria Vella*, aged 46, also of unspecified occupation

areas, so I collect certificates in urban and rural villages respecting the proportion.²³ Sensitivity analyses suggest that the results and coverage in Northern Italy would be dramatically underestimated if not using weights, because in Northern Italy, the real population is larger than the number of marriage certificates on the Antenati platform would suggest. The results are instead quite similar when considering the 15–30 marriage cohort and the 30–40 marriage cohort separately instead of the entire one (see Appendix Tables 17 and 18).

A clustered random sampling also has the advantage of allowing for analyses that better represent the specific village than randomly picking certificates from different villages. It also allows to integrate the research with qualitative information by village. Also, the process of random selection can be clearly documented because it is based on the list of potential villages to select that was created before starting to sample. The coverage is uniform across the areas of Italy that are contained in

²³ If there are insufficient observations in urban or rural cells to represent the sampling weights in the Antenati archive, the missing observations are proxied by those from the closest available province (see Table 7).

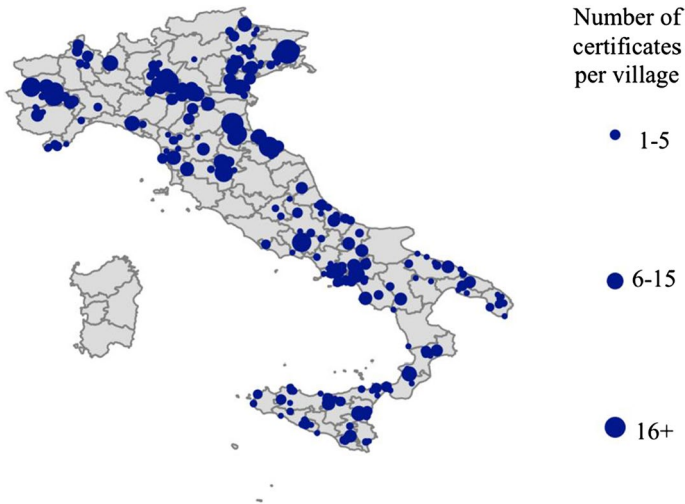


Fig. 3 Spatial coverage of randomly sampled clusters, or ‘villages’

digitized archives (see Fig. 3). After the post-stratification, each region or province, and each urban and rural area has a balanced number of certificates (see Table 11 in Appendix 1). Although three regions are missing, the universe of interest for which this sample is representative at the level of region and province is most of continental Italy (see Figs. 7, 8, and 9 in Appendix 1). Table 1 summarizes literacy rates with averages across important dimensions such as macroarea, gender, urban–rural, and occupation. The results will be critically discussed with a regression approach in the next subsection.

4.2 Regression-based adjustments

I use a regression approach to further correct for sampling imbalances. Regressions rely on more assumptions than the simple averages of Table 1. But a regression has the advantage of allowing for flexible balancing of the sample along dimensions that could not be stratified on when collecting the data, such as occupations. Skills and occupations are arguably crucial to explain literacy even after controlling for gender and urban–rural dimensions. For example, suppose that in 1815, the real population of women working as weavers in urban Northern Italy is 20%. The sampling approach cannot stratify on occupations, and as a result, it might accidentally contain only 5% of brides that work as weavers in urban Northern Italy, despite collecting the correct proportion of certificates given the population and urban–rural share of the selected area.

As the literacy rates of weavers tend to be low, women’s literacy rates in urban Northern Italy would consequently be overestimated. As an additional indirect effect, with other things equal, this sample bias would drive up literacy rates for the entire North, for urban areas, and for grooms. To mitigate this concern and safely add the dimension of literacy rates by occupation, I adapt the approach of Clark (2005), who introduced it to estimate real wages in England, to a context where the

Table 1 Literacy rates by gender, location, and occupation category (in %) from the post-stratified *Antenati* sample, c.a. 1815

Area	Summary		Gender		Urban–Rural		Occupation (HISCLASS)				GPI (%)
	N	Average literacy	Male	Female	Urban	Rural	High status	Craftsmen	Farmers	Unskilled	
North	1306 (1.22)	26.19 (1.92)	40.21 (1.28)	12.12 (3.03)	44.03 (1.28)	21.56 (4.23)	58.82 (3.20)	36.56 (1.38)	16.64 (3.80)	23.20	30.14
Centre	294 (2.14)	15.99 (1.98)	22.45 (2.42)	9.52 (5.63)	36.49 (1.94)	9.09 (9.60)	46.43 (6.41)	28.00 (1.76)	6.57 (22.36)	50.00	42.40
South	830 (1.15)	12.53 (1.98)	20.48 (1.03)	4.58 (3.10)	17.33 (1.22)	11.47 (5.18)	47.87 (2.95)	16.02 (1.11)	4.81 (8.19)	15.00	22.36
Italy	2,430 (0.81)	20.29 (1.33)	31.33 (0.83)	9.23 (2.15)	34.76 (0.84)	16.61 (3.11)	53.49 (2.16)	28.17 (0.89)	11.70 (3.44)	23.18	29.46
N	2430	2430	1216	1214	492	1,938	258	433	1,299	151	

*The occupation categories correspond (also in Appendix Tables 11, 12, 13) to Higher Status: HISCLASS 1–5, equivalent to ‘Services’; Industry: HISCLASS 6, 7, and 9, and 11; Agriculture: HISCLASS 8, 10; Unskilled: HISCLASS 12

**Macroareas are defined as follows (also in the following figures and tables). North: Liguria, Emilia-Romagna, Piedmont, Lombardy, and Venetia. Centre: Marche and Tuscany. South: Abruzzi, Basilicata, Calabria (excluding Calabria Citra), Campania, Molise, Puglia (excluding Capitanata), Sicily (excluding Caltanissetta)

***GPI stands for Gender Parity Index: the ratio of female over male literacy rates

variability is within many dimensions, but in a single cross section (area, gender, urban–rural, skill), rather than predominantly over time, and where the dependent variable is binary.²⁴ I estimate literacy rates in a (logit) regression framework controlling for skill, gender, urban/rural and regional dummies as follows:

$$\text{Lit}_i = \alpha + \sum_{i=1}^N \beta_i \text{Skill}_i + \sum_{i=1}^N \gamma_i \text{Gender}_i + \sum_{j=1}^J \delta_j \text{URBRATE}_j + \sum_{j=1}^J \theta_j \text{LOC}_j + \varepsilon_i \quad (1)$$

where Lit_i is the ability to sign of the i -th observation; Skill_i is a categorical variable with three values: primary, secondary and tertiary derived from HISCLASS 12; Gender_i is a gender dummy; URBRATE_j is a categorical variable indicating whether individual i resides in an urban or in a rural area within area j ; and LOC_j is a categorical variable with a value for each province, region or macroarea considered. As it was not possible to stratify the sample by skill levels without sampling all marriage certificates, in order to consider also the skill dimension in the regression, I aggregate the results by macroarea (North, Centre and South) to have enough observations for each skill level. To obtain literacy rates in each cell (e.g., for skilled women employed in the urban part of Northern Italy) I use the odds of the logit regression to predict the average marginal effect of being literate versus illiterate across each dimension (e.g., whether working in agriculture, industry, or services), fixing the categories of the remaining covariates (e.g., for women located in urban Northern Italy). Table 2 shows the resulting literacy rates. The resulting estimates are very disaggregated by macroarea, skill, urban–rural, and gender dimensions. The estimates suggest that Southern and Central Italy were very similar in terms of literacy rates also across occupations and Northern Italy instead had higher literacy rates also in rural areas. In order to have a more aggregate picture and to compare the regression estimates with the simple average estimates, as a next step I aggregate literacy rates within each of the dimensions and across gender, urban–rural, and skill dimensions. This gives us literacy rates of, say, of all women, or of all urban dwellers, or of all individuals in Northern Italy, rather than those of more detailed categories such as women in the North, employed in services in urban sectors. To do so, I aggregate the coefficients of Table 2 by North, Centre, and South using population, urban–rural, and skill weights. The population and urban–rural weights are to the weights used to post-stratify the sample. As some provinces and regions were not covered in the sample, I use weights that only include covered areas. The skill weight is obtained from Chilosi and Ciccarelli (2022).²⁵ The regression approach yields estimates reassuringly similar to those resulting from the simple average approach across the macroarea and

²⁴ Federico et al. (2019a) adopted the approach of Clark to obtain a representative sample of real wages for Italy. They already refined Clark’s approach to mitigate unbalanced samples at given points in time by adding regional population weights. In this paper I further disaggregate the estimates along the gender, urbanization, and skill dimensions.

²⁵ Suppose that, for instance, we are interested in observing the literacy rates of women in Northern Italy across urban–rural and skill dimensions. Calling a , b , c the literacy rates of women in urban areas and d , e , and f the literacy rates of women in rural areas across the three occupational sectors derived from table 2 and w the weights across the urban–rural and skill dimensions by gender, I aggregate as follows: $w_{\text{women, North}} = w_{\text{rur}} * (w_{\text{serv, North}} * a (= 33.9) * + w_{\text{ind, North}} * b (= 14.8) * + w_{\text{agr, North}} * c (= 4.9)) + w_{\text{urb}} * (w_{\text{serv, North}} * d (= 47.8) + w_{\text{ind, North}} * e (= 23.7) + w_{\text{agr, North}} * f (= 8.5))$.

gender dimensions, granting us the confidence that the results of the sampling procedure are robust to the important omitted variable of skills (compare Table 1 and Appendix Table 11). For this reason, regressions will only be used to add the important skill dimension to the picture (Tables 2 and 3).

Table 2 Literacy rates obtained through logit regressions (in %), 1815 c.a

Area	Sector	Men		Women	
		Urban	Rural	Urban	Rural
North	Industry	64.2 (15.6)	50.1 (14.2)	23.7 (15.9)	14.8 (14.8)
	Agriculture	34.8 (16.6)	23.0 (10.3)	8.5 (19.7)	4.9 (15.0)
	Services	84.1 (19.6)	74.7 (16.8)	47.8 (20.5)	33.9 (18.1)
Centre	Industry	45.1 (22.8)	31.6 (23.2)	12.5 (23.8)	7.4 (24.3)
	Agriculture	19.7 (23.2)	12.0 (20.7)	4.6 (26.1)	2.3 (24.1)
	Services	70.8 (25.1)	57.5 (24.2)	29.6 (26.5)	19.0 (25.9)
South	Industry	35.4 (17.9)	23.5 (16.6)	8.7 (20.3)	5.1 (19.3)
	Agriculture	14.0 (19.7)	8.4 (15.0)	2.8 (24.3)	1.6 (20.6)
	Services	61.7 (19.9)	47.5 (17.1)	21.6 (22.6)	13.5 (19.3)

*Each cell represents the predicted value of literacy rates at each model term's margin (gender, urban-rural, skill, area)

**Standard errors (in %) are in parentheses

Table 3 Regression estimates of Italian literacy rates by occupation (%), 1815 c.a

Macroarea	Occupational sector			
	N	Services	Industry	Agriculture
North	1306	56.7	34.9	15.5
Centre	294	41.2	21.8	8.3
South	830	33.3	16.2	5.8
Italy	2430	44.7	25.2	10.4
N		259	584	1299

*The occupation categories correspond to Services: HISCLASS 1-5, equivalent to 'Higher Status'; Industry: HISCLASS 6, 7, and 9, and 11; Agriculture: HISCLASS 8, 10, and 12

**The weights used to aggregate are the new urban-rural weight with population weights derived from Mariella, Postigliola and Rota (2021), and the skill weights in 1815 derived from Chilosi and Ciccarelli (2022, 1360)

The regression estimates differ from the simple average ones across the urban–rural dimension, particularly in central Italy.

The reason is of a technical nature: from Chilosi and Ciccarelli (2022) we can observe the proportion of skilled individuals by gender, but the skill weights are not disaggregated by urban and rural dimensions. Therefore, we cannot safely aggregate the literacy estimates of Table 2 along the urban–rural dimension. To see why, suppose we want to find the literacy rates of urban and rural Northern Italy. According to Chilosi and Ciccarelli (2022, 1360), in 1815 there are 64% individuals employed in the primary sector, 18% in industry, and 18% in services across Northern (and Central) Italy.

The sample has 60% individuals employed in the primary sector, 28% in industry, and 12% in services. If we could assume that the occupational shares were similar between urban and rural areas, we would use the overall weight of Chilosi and Ciccarelli (2022) also to look at urban–rural differences. But in urban areas of Northern Italy, our sample has only 28% individuals employed in the primary sector, 52% in industry, and 20% in services. Although these figures may be skewed due to sampling biases, they are more likely to reflect the rates of cities than the aggregate figures of Chilosi and Ciccarelli (2022), which include the vastly different occupational structure of rural areas. Using Ciccarelli and Chilosi's weights to compare the literacy rates of urban and rural individuals in the North would overestimate the proportion of peasants also in urban areas, thus underestimating literacy rates in urban areas, and overestimating it in rural areas. This is indeed what I observe (see Appendix Table 14), and for this reason I prefer to present the urban–rural literacy averages obtained directly from the population and urban–rural sample (Table 1).

5 Literacy rates in 1815

5.1 The new estimates

In this section I describe and interpret the new estimates of human capital in early Risorgimento Italy. The key results by macroarea, gender, and urban–rural are drawn from the simple averages of Table 1, and the results by occupation are drawn from the regression outputs of Tables 2 and 3. The North versus Central–South and gender gaps in literacy rates are wide. On average, spouses of Northern Italy signed 13.7% more certificates than Southern spouses (Table 1). Considering the lower bound of Southern literacy rates at 11.4% and the upper bound of Northern literacy rates at 27.4%, the difference becomes approximately 16 percentage points. The estimates of the regression approach are close to the estimates constructed using our stratified sample (compare Table 1 with Table 14 in Appendix). Importantly, the North–Centre gap is comparable in size to the North–South gap. Northern Italian spouses could sign marriage certificates 26.2% of the time, while their Central and Southern Italian counterparts did so 16% and 12.5% of the time on average, respectively.

The gap is mostly explained by the high literacy rates of Piedmont in the North and the low literacy rates of most Southern and Central Italy, except for Tuscany, Basilicata, and Campania (Figs. 4a and 9b). The provinces in Calabria, Apulia, and Basilicata have the lowest literacy rates in the South; the central regions of Tuscany and Marche closely match the literacy levels of southern regions; the literacy rate of the Emilia-Romagna region in the Centre–North, formerly half part of the Papal State is half that of the neighboring region of Lombardy, and similar to that of the province of Principato Citeriore (Salerno) in the South. Literacy gender gaps are consistently high throughout Italy, and lower only in the regions of Italy with lower average literacy rates (Fig. 4). 31.3% of grooms signed the marriage certificates, while only 9.2% of brides did (Table 1). The resulting literacy Gender Parity Index (GPI), the ratio of female to male literacy, is 29.5%, with a clear North–South pattern increasing as one moves North (see also Fig. 4b). The North–South ratio for female literacy rates in 1815 is 2.64. This is almost equivalent to 2.7, that is the value provided by Cappelli and Vasta (2020a) and based on elaborations from Ciccarelli and Weisdorf (2019)’s data for 1821. This ratio reached levels as high as 3.5 in 1860, suggesting that it widened markedly during the early nineteenth-century. Southern and Central Italy are surprisingly similar in terms of aggregate literacy rates, also across gender, urban–rural and occupational dimensions. In the North, the urban–rural literacy gap is very marked, while in the South it is almost nonexistent (see Table 1). This might explain why, over time, literacy rates increased in Northern Italy, while they did not in the South: the new model of school-based education focused first on urban agglomerations before diffusing in rural areas. The fact that urban areas in Northern Italy had higher literacy rates than urban areas in the South suggests that in Northern Italy, the process of schooling diffusion had already started and would have radically changed the scale of literacy diffusion over time. The wider spatial diffusion but lower overall level of literacy rates suggests that, besides large cities such as Naples, this diffusion still had to be initiated there.²⁶ The literacy rates of farmers (HISCLASS codes 8, 10, and 12) were at 10.4%, while workers in industry (HISCLASS codes 6, 7, 9, and 11) could sign in marriage certificates 25.2% of the time on average (Table 3). Among spouses of higher status (HISCLASS codes 1 to 5), mostly working in services, 44.7% of them could sign a marriage certificate. The literacy rate of Northern Italian farmers was twice the average literacy rate of Central and Southern Italy. This can be historically explained with the effort to diffuse primary schooling throughout Northern rural areas, and particularly in the Lombardy ruled by Joseph II of Habsburg-Lorena. Literacy rates in the North are especially high among farmers (15.5%) and among services workers (56.7%) compared to the rest of Italy.

I now compare the *Antenati* results with unitary census-based reprojected estimations of Ciccarelli and Weisdorf (see Fig. 5 and Table 4).²⁷ The *Antenati* figures date back to around 1815 for spouses aged 23 years old on average. Ciccarelli and

²⁶ See Basile, Ciccarelli, and Groote (2022) for the role of railways in the diffusion of literacy rates from urban to rural Italy.

²⁷ The conclusions of the comparison are unchanged when proxying the literacy of missing provinces in the *Antenati* sample with that of neighboring provinces (compare Table 4 with Table 10).

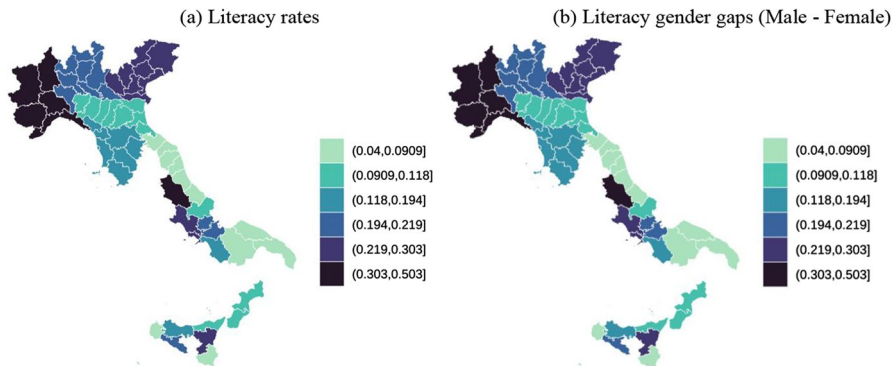


Fig. 4 Geography of literacy rates from the *Antenati* sample (by quantile), 1815 c.a

Weisdorf (2019) instead uses the cohort of individuals aged 90 to 100 years old in the 1881 census to infer the literacy rates of the group aged 30–40 in 1821. As a result, we should not expect the two numbers to be equal, as the *Antenati* estimates are direct and rely on marriage certificates rather than reprojecting census information. The most remarkable difference is that the estimates of Ciccarelli and Weisdorf are much higher for Northern and Central Italy than the 1815 *Antenati* sample estimates, while they are more similar for Southern Italy. In the North, literacy rates were 26.2% in the *Antenati* sample and 36.6% in Ciccarelli and Weisdorf (see Table 4 and Appendix Fig. 15). According to Ciccarelli and Weisdorf (2019), the North–South gap in levels in Italy was 20 percentage points, while this study found it to be 13.7 percentage points, thus smaller of 6.3 percentage points than previously estimated. The South represented 45.3% of the literacy of the North in Ciccarelli and Weisdorf (2019) and 47.7% in the *Antenati* sample (Table 4, panel B). This suggests that the size of the North–South gap was smaller than in Ciccarelli and Weisdorf (2019) because literacy rates were systematically lower in this study, rather than because the South represented a smaller share of the North’s literacy rate.

The difference in North–South literacy levels is especially pronounced for women. In 1815, the difference in female literacy rates between Northern and Southern Italy was as little as 7.5 percentage points and the female literacy rate in the South accounted for only 38% of the Northern literacy rate.²⁸ The high literacy gender gap that Ciccarelli and Weisdorf (2019) detected for women within Southern Italy appears now to be a more general feature of Italy at large. Female literacy rates in *Antenati* are only 9.2%, and they are almost double (16.2%) in the census reprojectations of Ciccarelli and Weisdorf (2019) (see also Appendix Fig. 15)²⁹ For males, instead, the literacy rates in *Antenati* appear lower but more in line with Ciccarelli and Weisdorf’s.

²⁸ The *Antenati* literacy rates for women are slightly higher in the North with the regression approach (13.8%), but the differences with Ciccarelli and Weisdorf (2019) are still remarkable (see Appendix Table 14). Figure 15 also shows the evolution over time of literacy rates by gender.

²⁹ Linking the *Antenati* 1815 estimates with the unitary censuses of 1861 and 1871, the trend appears to be similar to that derived from the 1881 census, but the levels of literacy rates are systematically lower.

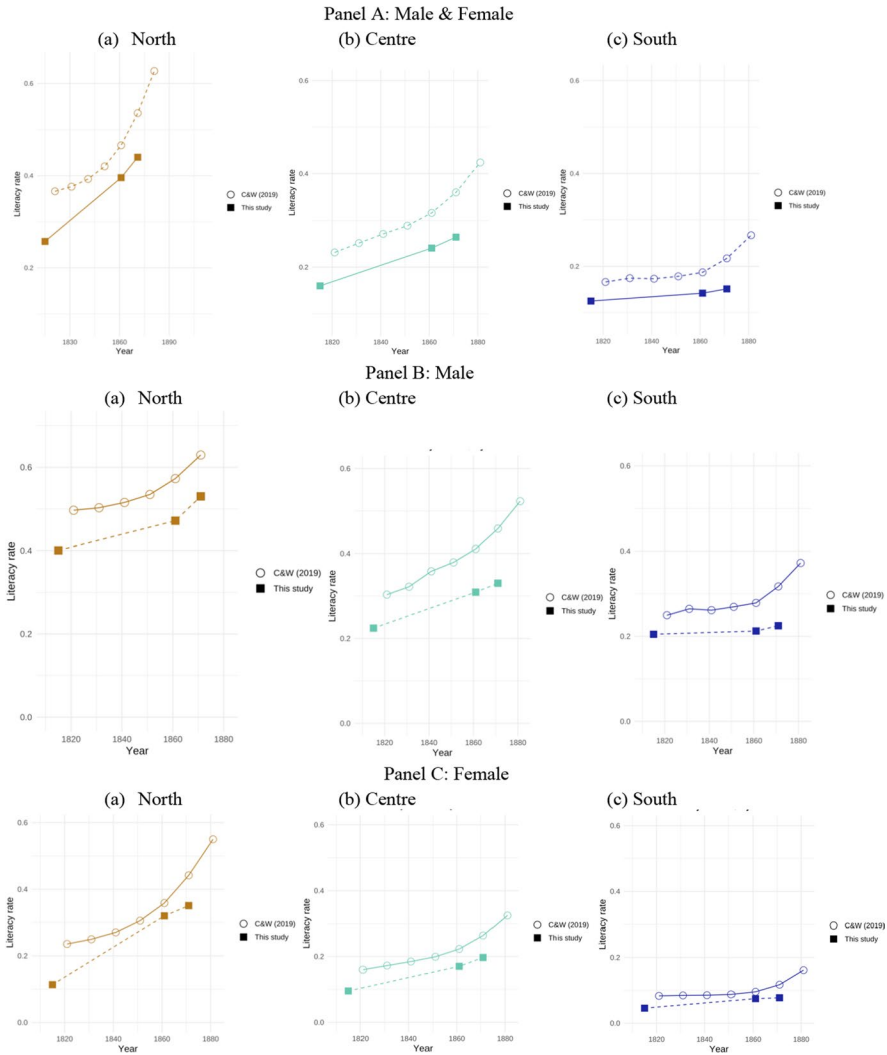


Fig. 5 Literacy rates by source, gender and macroarea. *Note:* ‘This study’ includes elaborations from the *Antenati* sample for 1815 and the literacy rates from national censuses of 1861 (MAIC 1866a, 388–411) and 1871 MAIC (1875, 4–239) for the years 1861 and 1871

Why were literacy rates more similar across Italy than previously estimated, and especially so for women? I put forth a tentative explanation that includes both historical and technical arguments. The *Antenati* estimates measure ‘pre-Napoleonic’ literacy rates for spouses who could attend primary schools just before the French reforms. At the time, North–South differences in State efforts to provide access to primary schooling were limited. Only parts of Lombardy in the North and urban Naples in the South already received systematic public schooling efforts. Institutional divides in educational provision were not as marked as after the French

Table 4 Comparison between literacy rates in the *Antenati* sample and in Ciccarelli and Weisdorf (2019)

Panel A: Literacy rates (in %)

	1815	1821
	This study	Ciccarelli and Weisdorf (2019)
<i>Male & Female</i>	Literacy rate	Literacy rate
Italy	20.3	26.3
North	26.2	36.6
Centre	16.0	23.2
South	12.5	16.6
<i>Male</i>		
Italy	31.3	36.3
North	40.2	49.7
Centre	22.4	30.3
South	20.5	25.0
<i>Female</i>		
Italy	9.2	16.2
North	12.1	23.5
Centre	9.5	16.0
South	4.6	8.3

Panel B: Comparative literacy gap estimates South/North (North=100)

	This study	Ciccarelli and Weisdorf (2019)
Male & Female	47.7	45.3
Male	51.0	50.3
Female	38.0	35.3

*This study's estimates are based on the *Antenati* sample with spouses aged 23 years old on average in 1815. Ciccarelli and Weisdorf (2019)'s estimates for 1821 are based on individuals aged 90 to 100 years old in 1881, and thus aged 30–40 in 1821

**The Ciccarelli and Weisdorf regional literacy rate estimates are aggregated across Italy and across macroarea using the population estimates of Mariella, Postigliola and Rota (2021). The *Antenati* sample is stratified according to population and urban–rural weights (see Sect. 4)

reforms, and primary schools relied on private institutions charging tuition fees both in the North and in the South.

Women suffered the most in this educational system, because it is plausible to argue that only wealthy families could invest in the education of both their sons and daughters, especially in teaching them to write and not just to read or to learn '*lavori femminili*' such as knitting and mending.³⁰ For women, such historical insights and

³⁰ Enquiries of the *Regno Italico* (Lombardy and Venetia) of 1805 suggest that only 1/4 of the very few schools (62) open to women taught them to write, and not just to read and to perform *lavori femminili* (Piseri 2017, 125–127).

the fact that the gap did not change much from 1815 to 1861 support the tentative explanation that unlike for men, pre-Napoleonic factors played a dominant role in explaining the high North–South literacy gaps of 1815 of women. The fact that in 1815 the literacy gender gap is especially marked in Piedmont where most primary schools were private (Chilosi 2007, 419) and less marked in Lombardy, where the Austrian government had already expanded public female schooling, supports this explanation (see Fig. 4b). The 1911 Daneo-Credaro school centralization reform raised female literacy rates in 1911 (Cappelli and Vasta 2020a).

Comparing the new 1815 estimates with unitary censuses of 1861 and 1871 (e.g., A'Hearn and Vecchi 2017, 178) suggests two possible perspectives on the history of education from the Napoleonic period to the Unification. The estimates of Ciccarelli and Weisdorf (2019) suggest a stagnation of literacy rates, while the new direct estimates suggest a sizeable increase during the Restoration.³¹ Although a preliminary conjecture, it is plausible to assume that the strong centralization feature of French reforms and the legacy of later Restoration reforms in the North led to a first ‘Silent Revolution’ for women, which lowered the literacy gender gap there considerably over the Risorgimento period and increased women’s literacy rates across Italy. With new public schools becoming open also to women, more of them could attend schools that taught them to write rather than *lavori femminili*. This increase in female literacy happened in the North and, although to a limited extent, also in the South.³² As a result, the North–South gap in female literacy rates slightly diminished over the Risorgimento period (see Table 5).

Schooling reforms during the Napoleonic and Risorgimento periods seem to be associated with a new, regionally unequal pattern of human capital accumulation for men, in contrast to a more uniform rise in female literacy rates. This disparity was primarily driven by a stagnation in literacy rates in the South (see Table 5): in 1815, the male literacy rate in the South was roughly half that of the North (51%), but it fell to a quarter of the North’s (23%) between the end of the Napoleonic period and the political unification of 1861.

Unlike the Daneo-Credaro reform, pre-unitary primary schooling reforms closed gender gaps in a very regionally unequal manner. The North was more exposed to pre-unitary reforms, and as a result, literacy rates increased more there than in the

³¹ The literacy rate in Italy in 1821 was, according to Ciccarelli and Weisdorf (2019), 26.3% (see Table 4, panel A), while on the eve of the unification it was 27% following Vecchi and A'Hearn (2017, 178).

³² Although not widely enforced (see Sect. 2), in the Kingdom of Naples the Napoleonic *Decreto Organico della Pubblica Istruzione* of 1806 imposed at least one female teacher to be present in every municipality to teach female pupils how to read and write: ‘*In ogni città o terra vi saranno uno o più maestri di leggere, scrivere, abaco, e principij di morale per fanciulli. Vi saranno anche delle maestre, che insegneranno alle fanciulle le sopradette cognizioni, e le prime arti donnesche*’ (Lupo, Gargano, and Marra 2014, 205) [‘There shall be one or more teachers of reading, writing, abacus, and principles of ethics in every town or land for all boys. There shall also be female teachers who shall teach to girls the disciplines mentioned above, and introductions to the arts of women.’ My own translation].

Table 5 North–South literacy gap from 1815 to 1861

	Southern Italian literacy rate (North = 100)		
	Antenati (1815)	Census (1861)	Difference (1861–1815)
Male & Female	47.7	36.0	– 11.7
Male	51.0	23.3	– 27.7
Female	38.0	45.0	7.0

Literacy rates in are drawn from individuals aged 18 or more years old from the 1861 census (MAIC, 1866a, 388–411). The *Antenati* literacy gaps are derived from Table 4. The shares include the North and South literacy rates only

South.³³ The spatially unequal pattern of post-Napoleonic primary school diffusion had a particularly strong impact on male literacy rates in 1861. Instead, the post-1815 regionally unequal centralization reforms in public schooling provision partially reduced female literacy gaps. The available sources do not allow us to compare the occupational structure of the *Antenati* sample with local studies for women as they do for men (see Appendix Table 14). But the *Antenati* sample, even though it is not balanced in its sample composition according to the occupational structure (and this is why I used regression approaches to disaggregate literacy rates by occupation in Sect. 4), still shows important North versus Centre–South differences in the occupational structure. Northern Italy has substantially more women employed in the secondary sector and less in the primary sector than the rest of Italy (see Fig. 6).

The socioeconomic structure of Northern Italy was still different from that of Southern Italy, suggesting that the different results are not driven by sample composition. A less agrarian occupational structure of women in the North did not translate into substantially higher female literacy rates in the North, possibly due to the institutional obstacles that a system based on private education implied at the time for all women except those originating from wealthy families. More technical aspects relate to the fact that cohorts in more recent censuses can explain the discrepancies between the results using the 1871 census and the estimates by Ciccarelli and Weisdorf, who used the 1881 census, but cannot entirely explain the difference that we observe between our sample and either the 1881 or 1871 censuses. Other explanations can relate to issues of census data retropolations: selective mortality, cohort effects, differences between stated (census) and proved literacy abilities (signatures), and age heaping (see the supplemental material for a technical discussion). Taking

³³ Bozzano, Cappelli and Vasta (2023, 37) show that the areas of Italy most affected by the French reforms, so Northern Italy and especially the North-West, aimed at increasing school attendance and thus gross enrollment rates rather than improving the efficacy of schooling in the Risorgimento period. As a result, by the unification and during the Liberal Age, North-Western Italy had the highest gap between gross enrolment rates and literacy rates, and enrolment rates far exceeded literacy rates. The gap between literacy and gross enrolment rate (GER) in both private and public primary educational institutions might have been even higher in the Risorgimento years. Already in 1850, the gross enrolment rate of the Kingdom of Sardinia (Piedmont, Liguria, and Sardinia) was as high as 42%, but literacy rates were lower than that (39.5%, according to Ciccarelli and Weisdorf 2019).

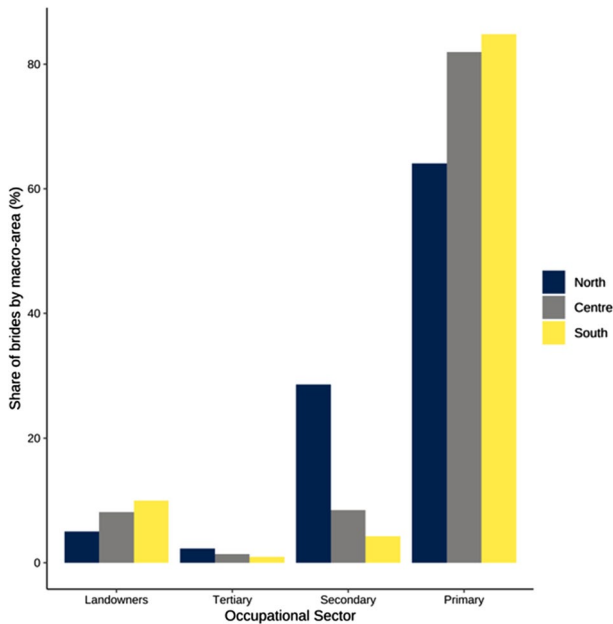


Fig. 6 Female occupational structure in the *Antenati* sample, by macroarea

the uncertainty that all such biases—which cannot be circumvented—imply for an even comparison, figures from marriage certificates and censuses are comparable, but highlight a different picture of the Italian North–South divides: the regional divides are less marked because literacy rates are generally low, and Central Italy has literacy rates more comparable to the South than to Northern Italy.

5.2 Comparison with local studies

As a last exercise, I compare the *Antenati* sample literacy rates with a set of local estimates, including new ones collected for this study. In this paper, a ‘local study’ is a study that contains quantitative estimates of literacy rates based on signatures in local areas such as villages and that can sum up to provinces or regions (see Table 5). Caution is necessary when comparing the estimates, as they often refer to selected areas of the selected regions or cities.³⁴

³⁴ ‘Local studies’ include Toscani (1987), Piseri (2002), Toscani (1993), Brambilla (1991), Milanesi (1991), Ferraresi (1991). Other studies estimated literacy rates in areas of Italy in the period, but to the best of my knowledge these are the only ones that cover the areas from 1800 to 1861 and for areas also covered in the *Antenati* sample. For instance, the local studies for Lombardy oversample the areas in the mountains, where literacy rates are higher and the number of landowners is larger; The local estimates for Piedmont and Marche are not very reliable as they only include a very selected part of the region (4 villages for Piedmont and 30 for Marche, concentrated around the regional capital). The estimates for Naples are very detailed but focus only on the central neighborhoods of the city.

Table 6 Comparison of considered studies and new estimates

References	Area	Source	Period	N vil- lages	N certifi- cates	Literacy male	Literacy female
This study	City of Naples	<i>Processetti Matrimoniali</i>	1750	1	81	0.29	0.10
This study	City of Naples	<i>Processetti Matrimoniali</i>	1775	1	148	0.34	0.12
Scirocco (1987)	City of Naples	Napoleonic Civil Records	1810	12	1922	0.57	0.25
This study	City of Naples	<i>Antenati</i> sample	1815	12	27	0.52	0.11
This study	City of Naples	<i>Antenati</i> all neighborhoods	1815	14	50	0.42	0.12
Toscani (1987)	Lombardy	Napoleonic Civil Records	1806–1810	176	14,072	0.43	0.08
Piseri (2002)	Lombardy	Napoleonic Civil Records	1800–1830	29	302	0.38	–
Toscani (1993)	Lombardy	Napoleonic Civil Records	1800–1830	341	13,625	0.49	–
Milanesi (1991)	Lombardy	Mantova State Archive	1806–1815	4	683	0.22	0.12
This study	Lombardy	<i>Antenati</i> sample	1810–1814	29	195	0.37	0.15
Brambilla (1991)	Marche	Macerata State Archive	1808–1814	30	–	0.19	0.06
This study	Marche	<i>Antenati</i> sample	1814	7	54	0.17	0.09
Ferraresi (1991)	Piedmont	Napoleonic Civil Records	1806–1810	4	656	0.25	0.13
This study	Piedmont	<i>Antenati</i> sample	1806–1810	17	161	0.59	0.09
Ciccarelli and Weisdorf (2019)	Udine province	1881 Census	1821	–	–	0.29	
This study	Udine province	<i>Antenati</i> sample	1810–1814	7	14	0.27	
This study	Udine province	<i>Antenati</i> full	1810–1814	88	4674	0.14	
Ciccarelli and Weisdorf (2019)	Salerno province	1881 Census	1821	–	–	0.18	
This study	Salerno province	<i>Antenati</i> sample	1816–1818	12	31	0.18	
This study	Salerno province	<i>Antenati</i> full	1816–1818	49	1950	0.21	
This study	Abruzzo Ultra II	<i>Antenati</i> sample	1816	6	18	0.50	0.00
This study	Abruzzo Ultra II	<i>Antenati</i> 50%	1816	63	472	0.26	0.04

*The term ‘local studies’ encompasses the studies on pre-unitary Italian literacy and alphabetization of X. Toscani, M. Piseri, E. Brambilla, A. Milanesi, A. Ferraresi. See also Bianchi (2019a, b) and Marzeddu (2019)

**For areas covered by more than one study (Lombardy) I considered the literacy rates only once per municipality. The ‘*Antenati* full’ uses the full set marriage certificates available around 1815 in the area between the rivers Isonzo and Livenza for Udine, and north of Eboli (included) for Salerno. The ‘*Antenati* 50%’ covers all marriage certificates from 63 of the 116 villages with marriage certificates of Abruzzo Ulteriore I. The ‘*Antenati* all neighborhoods’ covers sample of marriage certificates from all neighborhoods of the city of Naples. This sample is stratified by the number of certificates issued in each neighborhood and includes also the more peripheral neighborhoods of Miano, Marianella, and Stabilimento dell’Annunziata

Local studies provide precious information for three primary purposes. First, local studies can help to explain differences between previous studies. Literacy rates in Lombardy and Marche are broken down by occupational class, according to local studies. This allows us to compare the proportion of people employed in various sectors in *Antenati* and in local studies, as well as the literacy rates of each class. According to local studies, 46.2% of male spouses in Lombardy can sign, while 37% can sign according to the *Antenati* sample. The literacy rate by occupational sector is very close in *Antenati* and in local studies (see Appendix Table 12). Thanks to the disaggregation of literacy rates by occupational sector, we know that the difference can be attributed to the fact that in the local studies, more than 18% of individuals are landowners, compared to 7% in *Antenati*. The *Antenati* sample for Lombardy also includes 62% of workers in the primary sector, whereas local studies only include about 52%. The overrepresentation of landowners in Lombardy local studies, driven by an overrepresentation of hilly areas in Lombardy local studies, can explain why literacy rates in Lombardy are higher according to local studies (see supplemental material for further discussion). A similar logic can be applied to Marche, albeit with a smaller sample of villages in local studies.

Local studies can also be used to validate the representativeness of specific parts of the sample when there is reason to believe that the randomly selected villages do not accurately represent the unsampled ones. This bias is inherent in all random sampling methods, but it is frequently difficult to quantify. Local studies allow us to estimate how sensitive specific results are to chance imbalances. Consider Abruzzo Ulteriore II as an example. The province stands out as an outlier with literacy rates comparable to those of Naples (see Fig. 5). As a local study, I collected 50% of the signatures of the entire Abruzzo Ulteriore II whose resulting average literacy rate is 14.9%, lower than the *Antenati* sample in Abruzzo Ulteriore II by about 10 percentage points (see Table 6). The overestimation for Abruzzo Ulteriore II does not change the overall literacy rate estimates across macroarea.³⁵

This is reassuring because it shows that the sample estimations are robust even with the *Antenati* sample's possibly highest random imbalance.³⁶ A similar type of local study was conducted for the provinces of Salerno and Udine, where instead the *Antenati* sample correctly represents the population literacy rate.

Finally, local studies can provide us with more historically informative insights. For example, the city of Naples stands out because thanks to the additional information which can be obtained from the 1750 and 1775 *processetti matrimoniali* (see

³⁵ Without the 36 observations of Abruzzo Ulteriore II, the literacy rate of the rest of the South (830 observations) is 12.30%. If Abruzzo Ulteriore II had a literacy rate of 14.9% rather than 25% as in the *Antenati* sample, the South would have a literacy rate of $12.3 * \frac{796}{830} + 14.9 * \frac{36}{830} = 12.4\%$, just 0.13 percentage points lower than the *Antenati* sample estimate for the South.

³⁶ The imbalance discussed here is inevitable in random sampling as there is a 5% risk of selecting villages that do not represent the population average at the provincial level (see Appendix section A2). Local studies allow us to appreciate how large this bias is. One might wonder why the sample should not be adjusted if local studies reveal random imbalances, as in Abruzzo Ulteriore II. The main reason is that rerandomizing would imply explicitly selecting a limited number of certificates in such a way to obtain a predetermined literacy rate that we know to be correct. This would add new measurement and selection biases in other dimensions, such as skill levels. Moreover, for Abruzzo Ultra II, it is not necessary because it does not alter the macroarea estimations for the South.

Table 6 and Appendix A4). In 1750, literacy rates were already high for grooms in central Naples (28.76%). Literacy rates reached levels as high as 57% for grooms in 1810 and 1815, just before the Napoleonic reforms but after the effect of the Absolutist rulers' reforms, according to the full set of certificates collected by Alfonso Scirocco (1987) and 52% according to Antenati. This indicates that the city of Naples witnessed an important rise in literacy rates before 1815, which placed it at a comparable level to the literacy rates of Lombardy in the North as measured by the data-rich studies of Xenio Toscani. But it also indicates that, unlike in Lombardy, the effort in the Kingdom of Naples was placed on raising literacy rates in more élite parts of the country, such as Naples itself, while neglecting to diffuse this model of literacy expansion elsewhere.

6 Conclusion

The debate on the origins of the North–South divide is one of the most enduring and intriguing in Italian economic history. This paper contributes to this debate with direct estimates of literacy rates in the early nineteenth century, primarily using signatures drawn from a random sample of marriage certificates. The main finding is that the North–South gap in 1815 was smaller than previously assumed because Northern and Central Italian literacy rates were lower than expected. This suggests that, although pre-Napoleonic educational efforts were also important, and particularly for women, the large North–South human capital divide observed in unified Italy is to be traced back to the Napoleonic or Restoration reforms. The restored Southern Borbonic Kingdom's failure to build new primary schools in the run-up to unification may have exacerbated the differences in schooling. Meanwhile, the fact that women's literacy rates were as low as 9.2% across Italy in 1815 may suggest that the reorganization of schools after the French period was critical to closing the gender gap in literacy education before the unification, similarly to the Daneo-Credaro school centralization reform of 1911 (Cappelli and Vasta 2020a).

In 1815, literacy rates were particularly high in areas exposed to the public education reform, for example the Austrian-ruled Northern region of Lombardy, the Borbonic Southern region of Campania, and Naples in particular. This suggests a tentative association between the centralized decision-making in education by the government and high literacy rates despite different socioeconomic backgrounds. In explaining the divergence in literacy between the Northern and the Southern regions that occurred after the Unification, some scholars (Vasta 1999; Felice and Vasta 2015, 58; Cappelli and Vasta 2020b, 161) argue that the decentralized nature of the 1859 Casati Law limited the diffusion of literacy.³⁷ Within this wider debate, the pre-unitary evidence indirectly suggests that even if partially enforced, school centralization reforms in 1861 would have been beneficial to leveling out literacy divides across unified Italy. The new evidence from 1815, by showing that literacy rates were higher where school centralization reforms were even only partially enacted, indirectly supports the hypothesis that also a centralized Casati Law would

³⁷ Zamagni (1993) instead argues that the Casati Law positively affected literacy rates.

have contributed in leveling the divides after the Unification, rather than exacerbating them.

It is still unclear whether school centralization across pre-unitary Italy would have been sufficient to level the divide. Italy was a heterogeneous country, and other facts about how school reforms were implemented besides centralization arguably mattered. For instance, the longer exposure to the Napoleonic reforms in the Northern States led to higher unitary literacy rates there (Postigliola and Rota 2020). The fact that before the French reforms, women's literacy rates were low not just in the South, but also in the North and in the Centre, suggests that the heterogeneously adopted French reforms contributed to raising female literacy rates only in the North and Centre of Italy. The finding of this paper that Central and Southern Italian literacy rates were almost identical in 1815 supports Bozzano, Cappelli and Vasta (2023), who show that before the Unification, the North adopted a model of educational provision completely at odds with that of the Centre and South. The goal of the Northern Italian schooling system was to increase enrollment rates for all, at the cost of a lower quality of education, while the Centre and the South, despite centralizing *de iure* schooling provision, gradually invested more resources in an elitist model of high-quality schooling. It is not too far-fetched either to speculate that the burden of financing schooling left to each municipality may have contributed to widening the divide: Northern Italian municipalities dedicated more funds to education than Southern ones, and this can explain the widening North–South divide after the French period.

Disentangling the role of the Napoleonic domination with that of later administration of pre-unitary states in order to understand the respective roles in the literacy rates and divides would be a valuable direction of future research. Another important way forward would be to focus on the interaction of demand factors behind literacy rates with institutional reforms such as the Restoration ones. For instance, future works may expand this sample by placing more emphasis on occupations as a proxy for the demand for skills, and may also link literacy rates with the geography of primary school provision to study how the demand for literacy was practically met in this era. Mapping the relationship between female school provision and literacy outcomes would be particularly valuable in learning more about the first 'Silent Revolution' that was provisionally observed for women.

Appendices

Appendix 1. Further illustration of the sampling strategy

A1. Additional information on the sample coverage

See Figs. 7, 8, 9, 10, 11, 12, Tables 7, 8, 9, 10.

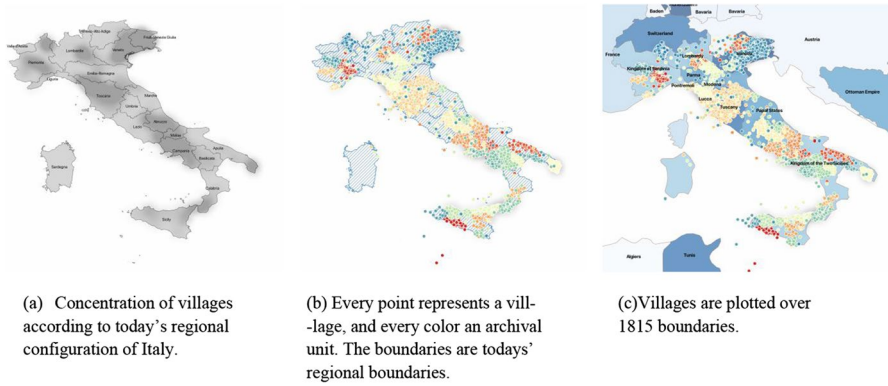


Fig. 7 Concentration of villages in the *Antenati* platform

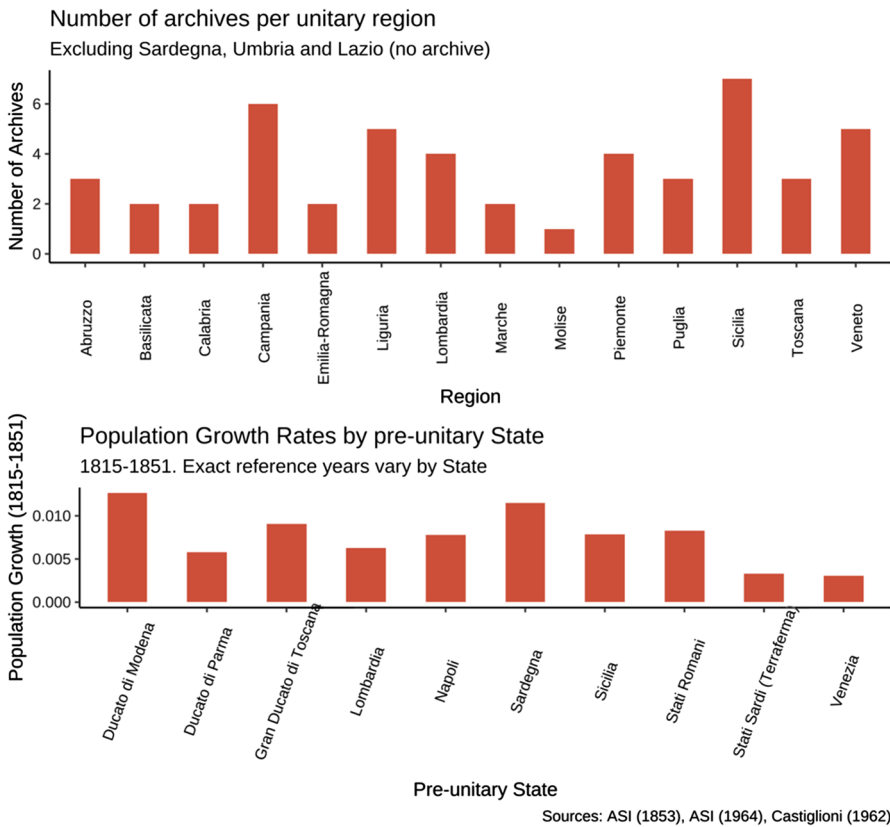


Fig. 8 Number of State archives by region and population growth in pre-unitary States (1815–1851)

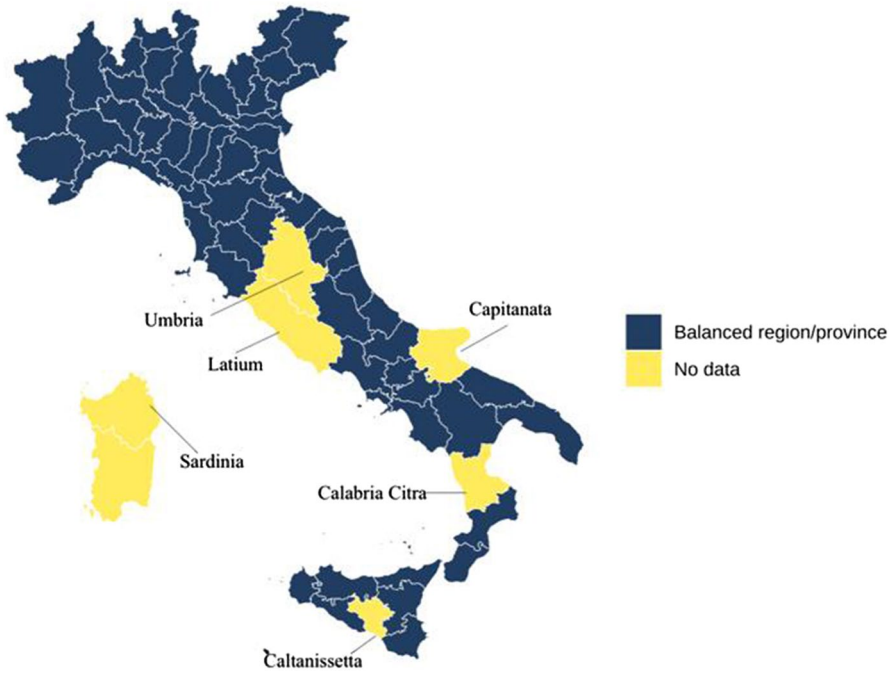
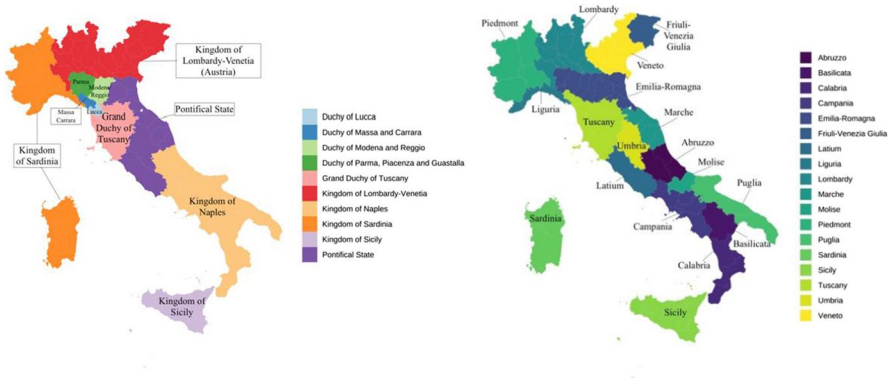


Fig. 9 Covered regions and provinces in the *Antenati* sample after including sampling weights



a. Italian States after the Congress of Vienna (1815)

b. Italian regions at 1881 boundaries

Fig. 10 Description of the pre-unitary States (A) and unitary regions of Italy (B)

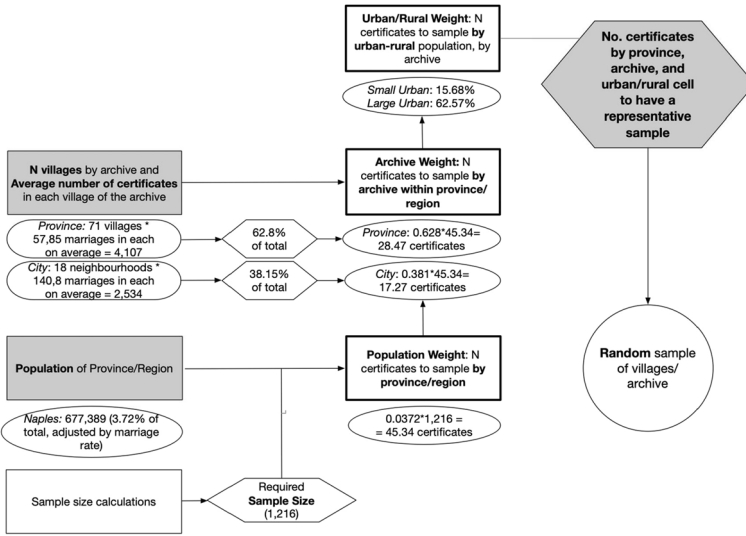


Fig. 11 Infographic illustrating the sampling and post-stratification approach. *Note:* Round boxes provide example figures for the province of Naples. Unlike most other cases, even though the State archive is just one, the province of Naples was treated as composed by two separate entities: city and province. The city of Naples is entirely urban. In the province, according to Malanima (2016), in 1800 the total urban population size was 124,000, of which Afragola, Torre Annunziata, and Pozzuoli counting as large cities with more or as much as 9,000 inhabitants, and totaling 35,000 inhabitants

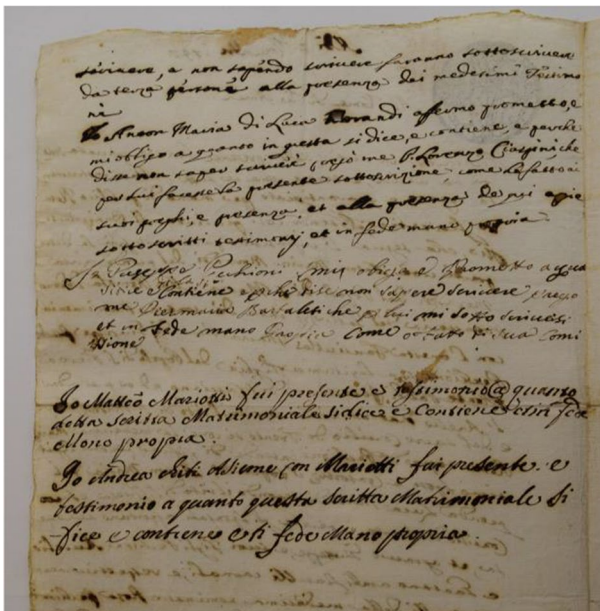


Fig. 12 Second page of a marriage document from Santa Maria La Romola, Tuscany, 1767

Table 7 Sample composition in terms of marriage certificates, by population of the area and by city size

Region-province	Collected certificates			Required cell size		
	Small urban	Large urban	Rural	Small urban	Large urban	Rural
Abruzzo Citra	1	2	14	0.8	2.2	14.1
Abruzzo Ultra I	2	0	9	2.1	0.0	9.2
Abruzzo Ultra II	<i>1</i>	1	15	0.9	1.3	14.8
Basilicata	1	1	12	1	1.1	11.6
Calabria Ultra I	<i>1</i>	1	17	0.7	1.0	17.1
Calabria Ultra II	0	1	16	0.0	0.8	15.5
Catania	2	8	12	1.5	8.1	12.3
Emilia-Romagna	4	23	93	4.2	23	93.1
Girgenti	0	5	10	0.0	4.9	9.7
Liguria	3	7	32	3.1	6.7	32.3
Lombardia	7	33	155	7.2	33.1	154.8
Marche	7	7	40	7	7.3	40.3
Messina	2	7	10	1.6	7.4	10.1
Molise	3	0	16	3.1	0.0	16.2
Napoli	7	29	10	7.3	29	10.1
Noto	<i>1</i>	2	12	0.9	2.0	11.5
Palermo	0	16	15	0.4	16.3	15.3
Piemonte	9	14	138	9.1	13.9	138.2
Principato Citeriore	1	1	29	0.7	0.7	28.5
Principato Ulteriore	0	2	34	0.0	2.0	34.2
Terra d'Otranto	2	5	15	2.1	4.7	15.3
Terra di Bari	0	7	18	0.0	7.3	17.5
Terra di Lavoro	4	2	27	4.2	1.9	27.2
Toscana	7	16	65	6.5	16.3	65.0
Trapani	0	6	5	0.4	5.5	5.0
Veneto	3	31	102	3.3	31.0	102.3
Total	68	227	921	67.8	227.5	921.2

*Each cell contains a number of certificates. Small urban areas have a population of less than 9,000 in 1800, and large urban areas have a population of more or as much as 9,000 in 1800

**Group 1 contains the certificates collected in the post-stratified sample. The marriage certificates were collected according to the required cell size rounded to the nearest integer

***Group 2 contains the minimum number of certificates required in each to have a representative sample

****Observations marked in italics are those that unavailable and were therefore collected from (1) similar urban agglomerations (Venice for Milan in Lombardy and for Genoa in Liguria) and (2) the closest available province: Abruzzo Citra for Small Urban in Abruzzo Ultra II (Penne); Calabria Citra for Calabria Ultra I (Amantea); Terra di Bari for Small Urban observations in Terra d'Otranto (Fasano); Avola in the province of Noto, which was not *città demaniale*

Table 8 Urban–rural weights in the parts of Italy covered and not covered by the *Antenati* sample, by macroarea

Macroarea	Covered by <i>Antenati</i>		Urban–rural shares				Sample	
	Population	Population 1815 (%)	Population by area (%)	Small urban (%)	Large urban (%)	Rural (%)	Small urban (%)	Rural (%)
Italy	Yes	13,347,462	86.0	5.5	18.5	76.0	5.6	18.7
	No	2,180,793	14.0	12.1	13.9	74.0	–	–
	Total Italy	15,528,255	100	6.4	17.9	75.7	–	–
North	Yes	6,953,275	100	4.1	16.3	79.5	4.0	16.5
	No	6,953,275	100	4.1	16.3	79.5	–	–
	Total area	6,953,275	100	4.1	16.3	79.5	–	–
Centre	Yes	1,659,825	61.6	9.3	16.7	74.0	9.9	16.2
	No	1,036,135	38.4	17.0	21.2	61.8	–	–
	Total area	2,695,960	100	10.9	16.4	72.7	–	–
South	Yes	4,734,362	80.5	6.2	22.3	71.5	6.7	22.9
	No	1,144,658	19.5	7.6	7.3	85.1	–	–
	Total area	5,879,020	100	6.5	19.4	74.1	–	–

*The provinces not included in the *Antenati* sample are: Latium and Umbria for the Centre, and Cagliari, Sassari, Calabria Citeriore, Caltanissetta, and Capitanata for the South

**The population figures are based on the 1821 provincial population Mariella et al. (2020) retropolated to 1815

***The ‘urban–rural shares’ are the planned urban and rural shares according to population data. The ‘sample’ contains the urban–rural shares present in the collected sample

****‘Large urban’ areas are urban areas of population of 9,000 or more based on the 1800 population of Malanima (2016). ‘Small urban areas are urban areas with population below 9,000 in 1800. Agro-towns are considered rural areas

Table 9 Italian literacy rates grouped macroarea and by the part of Italy covered and not covered on *Antenati*

Panel A: Literacy in 1861

Macroarea	Covered by <i>Antenati</i>	Literacy in 1861 (in %)			% of census population
		Male	Female	Male & Female	
Italy	Yes	33.9	19.6	26.6	90.5
	No	21.4	6.8	14.0	9.5
	Total Italy	32.7	18.3	25.5	100.0
North	Yes	47.2	32.03	39.57	100.00
	No	–	–	–	–
	Total area	47.2	32.03	39.57	100.00
Centre	Yes	32.0	18.01	25.14	84.4
	No	25.0	11.66	18.54	15.6
	Total area	30.9	17.03	24.11	100.0
South	Yes	21.5	7.9	14.5	84.7
	No	20.1	5.2	12.5	15.3
	Total area	21.3	7.5	14.2	100.0

*The literacy rates are based on individuals aged eighteen years old or above in 1864. Individuals able to read and write and individuals able only to read are both considered literate

**The provinces not included in the *Antenati* sample are: Latium and Umbria for the Centre, and Cagliari, Sassari, Calabria Citeriore, Caltanissetta, and Capitanata for the South

Table 10 Comparison between literacy rates in the *Antenati* sample and in Ciccarelli and Weisdorf (2019), assuming the literacy of the bordering provinces to missing areas of Italy

Panel A: Literacy rates (in %)		
	1815	1821
	This study	Ciccarelli and Weisdorf (2019)
<i>Male & Female</i>	Literacy rate	Literacy rate
Italy	19.4	26.3
North	26.2	36.6
Centre	15.6	23.2
South	12.6	16.6
<i>Male:</i>		
Italy	29.5	36.3
North	40.2	49.7
Centre	21.7	30.3
South	19.5	25.0
<i>Female:</i>		
Italy	8.5	16.2
North	12.1	23.5
Centre	9.2	16.0
South	3.9	8.3
Panel B: Comparative literacy gap estimates South/North (North = 100)		
	This study	Ciccarelli and Weisdorf (2019)
Male & Female	48.0	45.3
Male	48.5	50.3
Female	32.2	35.3

*This study's estimates are based on the *Antenati* sample with spouses aged 23 years old on average in 1815. Ciccarelli and Weisdorf (2019)'s estimates for 1821 are based on individuals aged 90 to 100 years old in 1881, and thus aged 30–40 in 1821

**The Ciccarelli and Weisdorf regional literacy rate estimates are aggregated across Italy and across macroarea using the population estimates of Mariella, Postigliola, and Rota (2021). The *Antenati* sample is stratified according to population and urban–rural weights (see Sect. 4)

***The *Antenati* sample estimates of this Table are the literacy estimates of Table 4 to which is summed the literacy of the missing provinces and regions, constructed assuming the literacy of the closest bordering provinces, weighted by the number of certificates of each neighboring province. In particular, *Umbria* is obtained using the literacy of the regions of Marche, Toscana, and Abruzzo; Sardinia is obtained using Sicily; *Caltanissetta* is obtained using the literacy of the provinces of Girgenti, Noto, Catania, and Palermo; *Calabria Citra* is obtained using the literacy of the provinces of Calabria Ultra I and Calabria Ultra II; *Capitanata* is obtained using the literacy rates of Terra di Bari, Molise, Principato Ulteriore, and Basilicata

A2. Sample size calculation

To calculate the minimum sample size, I assume a standard margin of error of 5% and a probability of true positive of 90% ($z_{\alpha}=1.645$), so that there is a 90% chance of estimating a literacy rate that is at most 5% distant from the true one. To calculate the coefficient of variation, or standardized standard deviation, of literacy, I rely on the closest available estimates at the time, the 1821 literacy rates of Ciccarelli and Weisdorf (2019). Their regional estimates are used to compute the mean (22.65%) and the standard error (14.38%) of the literacy rate. The resulting coefficient of variation is almost 0.65.

Following the calculations of Rossi et al. (2001, 911), the minimum sample size is $n = \left[\frac{r^2}{z_{\alpha}^2 \gamma^2} + \frac{1}{N} \right]^{-1} = 457,67$ marriage certificates, where $r = 5\%$ is the maximum allowable relative error between the estimate and the true value, $z_{\alpha} = 1.645$ is the critical value, γ is the coefficient of variation derived from Ciccarelli and Weisdorf (2019), and $\frac{1}{N}$ is an upward adjustment that increases the sample size more if the population size of marriage certificates (N) is expected to be lower. The number of certificates that gathered from the set of available archives on *Antenati* is estimated to be 86,864. This figure is obtained by multiplying the number of villages in each archive with marriage certificates of *Antenati* by the corresponding mean number of certificates that were gathered from a random sample of eight villages from each archive, also averaged over three benchmark years around 1815 to account for seasonality. The required sample size is not sensitive to changes in the estimated population of marriage certificates: Assuming a total number of marriage certificates of 50,000 the minimum sample size is of 456, and for 100,000 marriage certificates it is of 459. This number does not consider the fact that the structure of the data imposes to consider villages, instead of certificates or individuals, as sampling units. Correcting for intra-cluster correlation through the intra-cluster variance obtained from a pilot study of four sixteen villages drawn from four different archives and adjusting for the corresponding design effect of 2.68 (see supplemental material for details), the minimum sample size for the part of Italy covered on *Antenati* becomes 1216.

A3. More details on the sampling strategy

This section outlines the method adopted to obtain a representative sample of marriage certificates, starting at the village level and scaling up the estimates to the regional level (see also Fig. 8). I consider the provinces of Naples and the region of Lombardy to illustrate the procedure. The process goes as follows.

First, I count the average number of marriage certificates that there are in 1815 in each neighborhood or village: 140.8 marriages per neighborhood in Naples, and 57.9 marriages in the province.³⁸ Second, I look at the number of archives and the number of villages that each archive contains in the province of interest: is there more than one archive? In the case of Naples, the city and provincial data are collected by the same state archive, but the number of marriage certificates varies considerably in the two

³⁸ When the number of villages was very large, as in the case of Udine, I took a random sample of them, covering at least 25% of the considered villages.

areas, so I treat them as separate archival units.³⁹ The first weight, or ‘archive weight,’ indicates the proportion of certificates to be collected in each archive within a province or region and is based on the total number of marriage certificates contained in each. A province may only have one archive from which to obtain marriage certificates. In this case, the archive weight is equal to one.⁴⁰ If a province or region has more than one archival unit, the issue of how many observations to draw from each archive to accurately represent the true population composition arises. In Naples, there are two archival units, both corresponding to the province of Naples: the province, and the city of Naples. Naples has 18 neighborhoods with an average of 140,8 certificates in 1815. In the province of Naples there are 71 villages, with an average number of certificates in 1815 of 57.85. The archival weight divides the total number of marriages in the city of Naples by the total number of marriages in the entire province of Naples, including its province. Accordingly, the city of Naples accounts for $\frac{18 \times 140.8}{18 \times 140.8 + 57.85 \times 71} = 0.381$, or 38.1%, of the total number of marriages in the province of Naples. Third, I look at the population sources (primarily Mariella et al. 2020) to know how many individuals there were overall in the province of Naples in 1815 (677,389). I could simply divide the population of Naples by the population of the part of Italy covered in the sample. But we sample marriage certificates rather than directly individuals, and the rate of marriage varies across Italy. In 1853, there were just 107 individuals per marriage in Venetia and up to 143.09 individuals per marriage in the Grand Duchy of Tuscany (ASI 1853, 84). This regional variability in marriage patterns would not be accounted for by simply balancing the within-province representativeness of each archive.

As a result, I convert both the Naples and Italian populations to the corresponding marriage rates. In the numerator, I multiply the population of Naples by the number of marriages per inhabitant (0.0072) estimated by ASI (1853, 83) for the Kingdom of Naples, and in the denominator, I multiply the population of each province by the corresponding rate of marriage per inhabitant of the province to obtain a weighted average population of Italy proxied by marriage certificates.⁴¹ The resulting ratio is the ‘population weight’ $\left(\frac{677,389 \times 0.0072}{130,849} \right) = 0.0372$.⁴² The minimum sample size in terms of marriage certificates is 1216, so we need a total of $1216 \times (0.0372) = 45.34$ certificates in the entire province of Naples. Combining the

³⁹ I also treated other large urban areas separately from their province (Palermo and Turin).

⁴⁰ The provinces and regions that do not have the archive weight because they only have one reference archive are Abruzzo Citra, Abruzzo Ulteriore I, Abruzzo Ulteriore II, Calabria Ultra I, Calabria Ultra II, Catania, Agrigento, Messina, Molise, Ragusa, Salerno, Terra di Bari, Terra di Lavoro, and Trapani.

⁴¹ The archival weight was also adjusted for the number of marriages per inhabitant in the pre-unitary kingdom from ASI (1853, 83). For instance, for the city of Naples, the percentage of certificates, adjusted by the marriage rate of the Kingdom of Naples (138.85), is $\frac{(140.8 \times 18)}{11,426,389} \times 138.85 = 3.07\%$, where 11,426,389 is the product of all average marriage certificates by all villages by the number of individuals per marriage in the corresponding pre-unitary state of the province. The corresponding percentage of certificates in the provincial area of Naples is 5.69%. The resulting refined version of the archive weight for the city of Naples is $\frac{3.07}{5.06 + 3.07} = 0.3504$.

⁴² The population weight is similar when using the population directly rather than weighting the population by the marriage rates, but it is less accurate because we are interested in sampling marriage certificates that represent the population shares. For Naples, the resulting population weight would be 4% rather than 3.72%.

archive and the population weight, we need $17.30 = 45.34 \times 0.381$ certificates from the city, and $28.19 = 45.34 \times 0.628$ certificates from the province.

Finally, I use the 1800 population figures of Malanima (2016) to account for the proportion of urban and rural population in both Naples and its province (see supplemental material for more details). In 1800, 62.6% of the population in the entire province of Naples lived in large urban areas with 9000 or more inhabitants. 15.7% of the population lived in urban areas with populations ranging from 4000 to 9000 people, while the remainder lives in rural areas. The city of Naples is entirely urban, and the provincial portion of Naples only accounts for 9.3% of the province's large urban population, which includes the city of Naples. Accordingly, $62.6\% \times 9.3\%$ of the large urban certificates are drawn from the province, and $45.34 \times 62.6\% \times 91.7\%$ (26) are drawn from the city.⁴³ The 13.9% certificates (7) from small urban areas and the 30.9% certificates (15) from rural areas are collected from the province of Naples.

The process is very similar in the region of Lombardy, and for this reason it is applied recursively over the remaining provinces and regions. The only difference is that Lombardy, like other Northern regions, has fewer archives than its population suggests: the archives of Como, Brescia, and Mantova have provided the scanned certificates, but the archives of Milano and Bergamo did not. The population weight ensures that the region is well represented in terms of the total number of certificates, but sampling biases may arise if the available archives are systematically different from the unsampled ones. For Lombardy, this is partially the case because excluding Milan means excluding the most dynamic, and probably literate area in Lombardy. It is also the case for the city of Genoa, which cannot be sampled in *Antenati*. For at least two reasons, Milan and Genoa are proxied by Venice in the main estimates for Northern Italy. First, the Northern literacy rates are highest when proxying both Milan and Genoa through Venice. This leads to an upper bound on the observed North–South gap of about 13.6%, compared to a North–South gap of about 12.5% that would result from using Turin as a proxy for Milan and Genoa (Table 12). Second, the certificates from Venice are better preserved than those from Turin. They also date back to 1815 rather than to 1810, and they only represent the urban core of Venice, whereas the observations for the city of Turin also derive from neighboring rural villages, such as Cirié.

The bias is thus accounted for by proxying Milan and Genoa through Venice (see Table 1). This choice is supported by three sensitivity checks. First, proxying Milan and Genoa by cities of the corresponding region with populations above 9000 inhabitants, such as Como, Cremona, and Mantova for Lombardy and Savona for Genoa (Appendix Table 11); second, proxying Milan and Genoa with observations drawn from the city of Turin (Appendix Table 12); third, proxying Milan with observations drawn from Turin, and Genoa with observations drawn from Venice (Appendix

⁴³ The province of Naples is a special case because the size of the archival unit of the city of Naples is such that the urban-rural weight is partially in contrast with the archival weight. In fact, the archival weight told us to sample only 17.15 certificates from the city of Naples, and the large urban weight told us to sample 26 certificates from the city. We should follow the urban-rural weight so as not to collect too many large urban observations from the province compared to what the urban population of the province represents.

Table 13). The interval of Northern Italian literacy rates ranges from 24.80% when other Lombardy and Liguria urban areas are used (Table 11) to 26.18% when both cities are proxied through Venice (Table 1).

A4. Other sources beyond civil registers to infer literacy in pre-unitary Italy

As discussed in Sect. 3, civil registers, particularly marriage certificates, provide a comprehensive overview of Italian literacy rates in 1815. However, there may be additional sources that can supplement civil registration records. Before marrying, spouses had to submit documents to the priest in preparation for the upcoming marriage. Marriage documents included declarations of not being already married, or about being no legal opposition to the marriage. Such documents were legally binding, so spouses were often required to sign them, even when the marriage certificates were only signed by the priests, as before the Napoleonic reforms. For this reason, the *processetti matrimoniali* are an important alternative source to marriage certificates in Italy for the pre-Napoleonic period. However, almost every parish in Northern and Central Italian states did not require the signatures of the spouses, so the *processetti* are a useful source to study pre-Napoleonic literacy rates only in the Southern Kingdom of Naples where instead signatures are more frequent. Evidence is hard to come by also in the South as every parish keeps the information in its own archive and the preservation quality varies considerably. The Diocesan State Archive of Naples is an exception as it contains the *processetti* relating to marriages of the entire city of Naples and is accessible up to the year 1800 not included. I thus collect the signatures of 224 marriage *processetti* of brides and grooms whose surname started with the letter ‘B’ in 1750 and with letters ‘B’ and ‘L’ in 1775 in Naples.⁴⁴

For the rest of Italy, the evidence is even more scant as only few parishes decided to ask the spouses to sign certificates of equivalent to *processetti*, and most individual parishes keep their own archive: for the province of Florence, out of more than 200 parishes gathered in the *Archivio Arcivescovile*, only 6 had *processetti* for the eighteenth century, and only one parish (Santa Maria La Romola) had signatures of 26 spouses. The signatures only covered the period 1767–1797 (Fig. 11).

Appendix 2. Further results and robustness checks

See Tables 11, 12, 13, 14, 15, 16, 17, 18, Figs. 13, 14, 15.

⁴⁴ I checked whether Northern and Central Italian archives contained signatures of spouses before Napoleon by contacting or visiting the Diocesan archives of Milan, Como, Brescia, Bologna, Florence, Pisa, and Lucca. Northern and Central Italian parish archives sometimes contain signatures for legal processes regarding divorces. However, there are just few such processes, and they are very selected because at the time it was only possible to ask for divorce due to very severe situations of violence, and the spouses were often wealthy.

Table 11 Estimates of Italian literacy rates by macroarea, gender, urban-rural and occupation category (%), with Milan and Genoa proxied by towns of their region with 9,000 or more inhabitants

	N	Average literacy	Male	Female	Urban	Rural	High status	Craftsmen	Farmers	Unskilled	GPI (%)
North	1306 (1.20)	24.80 (1.90)	38.53 (1.23)	11.04 (3.05)	38.19 (1.27)	21.58 (4.27)	57.46 (3.11)	35.44 (1.37)	16.42 (3.62)	18.26	28.65
Centre	294 (2.14)	15.99 (3.45)	22.45 (2.43)	9.52 (5.63)	36.49 (1.94)	9.10 (9.60)	46.43 (6.41)	28.00 (1.76)	6.57 (22.36)	50.00	42.40
South	830 (1.15)	12.53 (1.98)	20.48 (1.03)	4.58 (3.10)	17.33 (1.22)	11.47 (5.18)	47.87 (2.95)	16.02 (1.11)	4.81 (8.19)	15.00	22.36
Italy	2,430 (0.80)	19.55 (1.32)	30.44 (0.81)	8.65 (2.12)	31.38 (0.84)	16.65 (3.12)	52.73 (2.13)	27.77 (0.89)	11.59 (3.32)	19.15	28.41
N	2430	2436	1216	1214	478	1952	256	443	1,3031,303	141	

Table 12 Estimates of Italian literacy rates by macroarea, gender, urban–rural and occupation category (%), with Milan and Genoa proxied by Turin

Area	Summary		Gender		Urban–Rural		Occupation (HISCLASS)			GPI (%)	
	N	Average literacy	Male	Female	Urban	Rural	High status	Craftsmen	Farmers		Unskilled
North	1306 (1.20)	24.96 (1.91)	39.30 (1.21)	10.58 (2.97)	38.06 (1.28)	21.58 (4.31)	58.33 (3.22)	35.60 (1.37)	16.53 (3.56)	17.95	26.92
Centre	294 (2.14)	15.99 (3.45)	22.45 (2.43)	9.52 (5.63)	36.49 (1.94)	9.10 (9.60)	46.43 (6.41)	28.00 (1.76)	6.57 (22.36)	50.00	42.40
South	830 (1.15)	12.53 (1.98)	20.48 (1.03)	4.58 (3.10)	17.33 (1.22)	11.47 (5.18)	47.87 (2.95)	16.02 (1.11)	4.81 (8.19)	15.00	22.36
Italy	2436 (0.81)	19.63 (1.32)	39.84 (0.80)	8.40 (2.10)	31.50 (0.85)	16.62 (3.14)	53.15 (2.16)	27.57 (0.89)	11.68 (3.28)	18.88	21.08
N	2430	2430	1216	1214	492	1938	254	428	1310	143	

Table 13 Estimates of Italian literacy rates by macroarea, gender, urban-rural and occupation category (%), with Milan proxied by Turin and Genoa proxied by Venice

Area	Summary		Gender		Urban-Rural		Occupation (HISCLASS)				GPI (%)
	N	Average literacy	Male	Female	Urban	Rural	High Status	Craftsmen	Farmers	Unskilled	
North	1306 (1.21)	25.73 (1.92)	40.06 (1.24)	11.35 (3.02)	41.79 (1.28)	21.58 (4.31)	58.33 (3.21)	36.00 (1.38)	16.64 (3.78)	21.67	28.33
Centre	294 (2.14)	15.99 (3.45)	22.45 (2.43)	9.52 (5.63)	36.49 (1.94)	9.10 (9.60)	46.43 (6.41)	28.00 (1.76)	6.57 (22.36)	50.00	42.40
South	830 (1.15)	12.53 (1.98)	20.48 (1.03)	4.58 (3.10)	17.33 (1.22)	11.47 (5.18)	47.87 (2.95)	16.02 (1.11)	4.81 (8.19)	15.00	22.36
Italy	2430 (0.81)	20.04 (1.33)	31.25 (0.81)	8.81 (2.13)	33.54 (0.85)	16.62 (3.14)	53.15 (2.16)	27.84 (0.89)	11.72 (3.43)	21.92	28.19
N	2430	2430	1216	1214	492	1938	254	431	1305	146	

Table 14 Estimates of Italian literacy rates by macroarea, gender, urban–rural and occupation category (%), logit regression approach

Macroarea	Summary		Gender		Urban–Rural		Occupation (HISCLASS)			GPI
	N	Average literacy	Male	Female	Urban	Rural	Services	Industry	Agriculture	
North	1306	26.4	39.6	13.2	33.6	24.5	56.7	34.9	15.5	33.4
Centre	294	17.3	26.1	8.5	21.8	15.0	41.2	21.8	8.3	29.0
South	830	9.4	15.4	3.5	13.0	8.4	33.3	16.2	5.8	22.4
Italy	2430	18.2	27.8	8.6	23.4	16.5	44.7	25.2	10.4	3.09
N	2430	2430	1216	1214	492	1938	259	584	1299	

*The occupation categories correspond to Services: HISCLASS 1–5, equivalent to ‘Higher Status’; Industry: HISCLASS 6, 7, and 9, and 11; Agriculture: HISCLASS 8, 10, and 12

**The weights used to aggregate are the new urban–rural weight with the population weights derived from Mariella, Postiglione and Rota 2021, and the skill weights in 1815 derived from Chilosi and Ciccarelli (2022, 1360)

***GPI stands for Gender Parity Index: the ratio of female over male literacy rates

Table 15 Male literacy rates from marriage certificates and local studies, by occupation category

Region-province	Landowners		Primary		Secondary		Tertiary	
	Local	Sample	Local	Sample	Local	Sample	Local	Sample
Lombardia	0.65	0.67	0.19	0.20	0.52	0.65	0.60	0.53
Marche	0.22	0.00	0.07	0.10	0.37	0.22	0.48	1.00

*Comparison of literacy rates among occupational structures estimated in various local studies and in this sample, at the regional level. I converted the HISCO and HISCLASS classification into primary, secondary, tertiary, and landowners sectors to follow the classification of most local studies

Table 16 Share of male spouses in the different occupational groups, in local studies and in this sample

Region-province	N		% Landowners		% Primary		% Secondary		% Tertiary	
	Local	Sample	Local	Sample	Local	Sample	Local	Sample	Local	Sample
Lombardia	1162	195	18.59	6.62	51.64	62.43	19.10	20.44	10.67	10.50
Marche	786	54	30.41	1.92	25.70	76.92	27.61	17.30	18.68	3.85

*I converted the HISCO and HISCLASS classification into primary, secondary, tertiary, and landowners sectors to follow the classification of most local studies

**Similar figures for women are not available because the local sample estimates do not provide such information

Table 17 Literacy rates by gender and location, *Antenati* sample, 15–30 cohort

Macroarea	Summary		Gender		Urban–rural	
	Average literacy	N	Male	Female	Urban	Rural
North	25.2	978	41.3	11.9	43.7	20.9
Centre	13.0	223	22.0	5.7	37.5	4.8
South	12.5	639	20.4	4.8	17.6	11.3
Italy	19.3	1840	31.3	8.8	34.1	15.7
N		1840	860	980	348	1482

*This balanced sample only considers spouses aged 15 to 30 [15–29 included] in around 1815

Table 18 Literacy rates by gender and location, *Antenati* sample, 30–40 cohort

Macroarea	summary		Gender		Urban–rural	
	Average literacy	N	Male	Female	Urban	Rural
North	21.9	194	27.7	8.47	37.0	17.3
Centre	18.9	49	19.4	17.65	33.3	14.6
South	13.4	129	21.2	4.9	20.0	11.8
Italy	18.6	376	24.7	8.0	31.3	15.0
N		376	239	137	83	293

*This balanced sample only considers spouses aged 30 to 40 [30–39 included] in around 1815, excluding second marriages

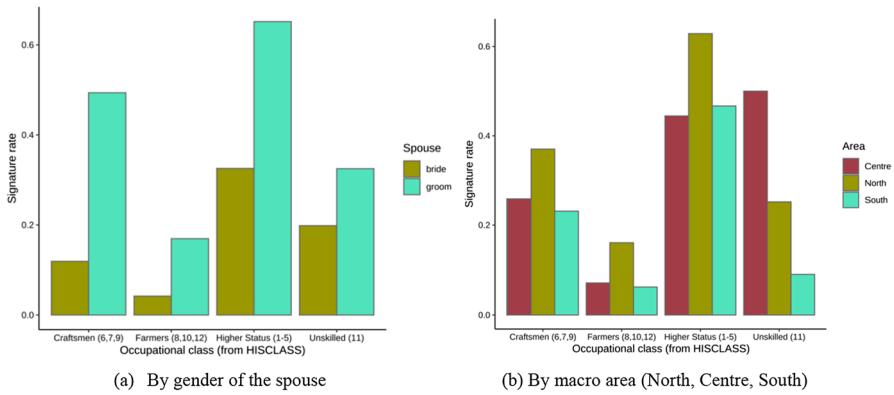


Fig. 13 Histogram showing the percentage of spouses able to sign by broad socio-occupational class

Fig. 14 Number of collected marriage certificates by socio-occupational class and macroarea

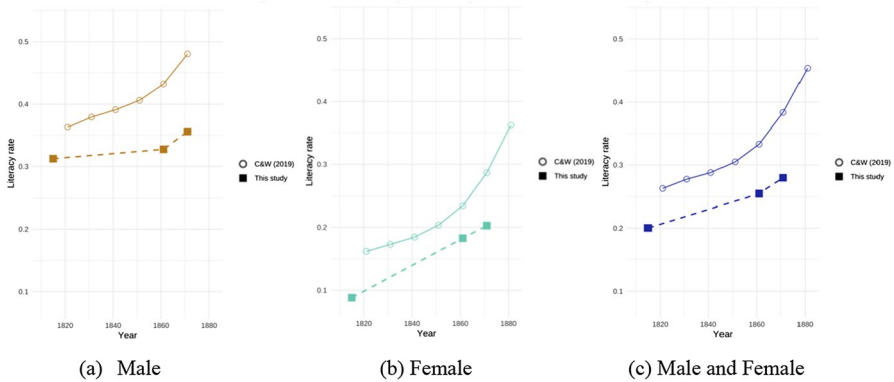
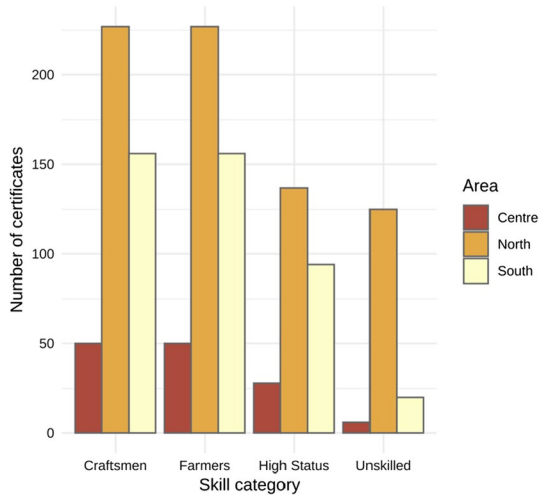


Fig. 15 Literacy rates by census source and gender. ‘This study’ includes elaborations from the *Antenati* sample for 1815, and the literacy rates derived unitary censuses of 1861 and 1871 for the years 1861 and 1871

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11698-023-00271-2>.

Acknowledgements I would like to thank Alessandro Nuvolari and Michelangelo Vasta for the insightful discussions and encouragements. This paper benefitted from the comments of Carlo Ciccarelli, Gabriele Cappelli, Martina Occelli, Giovanni Vecchi (discussants), Lorenzo Del Panta, and Leigh Shaw-Taylor, as well as from the helpful suggestions of David Chilosì, Giovanni Federico, Giuliana Freschi, Jacob Weisdorf, Michele Postigliola, Mauro Rota, Jochen Streb, Leonardo Ridolfi, and of the participants at the Spring 2020 Sant'Anna Quarantine Seminar, Fall 2020 Sant'Anna Workshop, 2021 Economic History Society Conference, 2021 Lost Highways Seminar (LHSem), and the 3rd Sapienza Economic History Workshop. I am also grateful to the staff of the libraries of Scuola Normale Superiore, Scuola Sant'Anna, to the libraries of Economics and of History and Philosophy of the Università di Pisa and of the state archives of Naples and Palermo for their support.

Funding Open access funding provided by Scuola Superiore Sant'Anna within the CRUI-CARE Agreement.

Data availability The data, methods, and replication codes underlying this article are available for further studies and replication that can be accessed via the following link: <https://deposit.icpsr.umich.edu/deposit/claimResource?tenant=openicpsr&claimId=45579>.

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

- A'Hearn B, Vecchi G (2017) Education. In: Vecchi G (ed) *Measuring the Wellbeing of Italians*. Oxford University Press, Oxford
- A'Hearn B (1998) Institutions, externalities, and economic growth in southern Italy: evidence from the cotton textile industry, 1861–1914. *Econ Hist Rev* 51:734–762. <https://doi.org/10.1111/1468-0289.00112>
- A'Hearn B, Delfino A, Nuvolari A (2022) Rethinking Age Heaping: a cautionary tale from nineteenth-century Italy. *Econ Hist Rev* 75:111–137. <https://doi.org/10.1111/ehr.13087>
- Andersson J, Berger T (2019) Elites and the expansion of education in nineteenth-century Sweden. *Econ Hist Rev* 72:897–924. <https://doi.org/10.1111/ehr.12738>
- Aymard M (1978) *La transizione dal feudalesimo al capitalismo*. Einaudi, Milano
- Álvarez B, Palencia FR (2018) Human capital and earnings in eighteenth-century castile. *Explor Econ Hist* 67:105–133. <https://doi.org/10.1016/j.eeh.2017.10.005>
- Astoul G (1992) L'alphabétisation en Haut-Languedoc au XVIIIe siècle: les aléas de la méthode Maggiolo dans le diocèse de Montauban. *Annales Du Midi* 104(198):175–194
- Banti AM (1996) *Storia della borghesia Italiana: L'età liberale* (Vol. 1). Donzelli Editore, Rome
- Basile R, Ciccarelli C (2018) The Location of the Italian Manufacturing Industry, 1871–1911: a sectoral analysis. *J Econ Geogr* 18:627–661. <https://doi.org/10.1093/jeg/lbx033>
- Basile R, Ciccarelli C, Groote P (2022) The legacy of literacy: evidence from Italian regions. *Reg Stud* 56(5):794–807. <https://doi.org/10.1080/00343404.2021.1926960>
- Benigno F (2014) *L'età moderna: dalla scoperta dell'America alla Restaurazione*. Laterza, Rome-Bari
- Bozzano M, Cappelli G, Vasta M (2023) Whither education? The long shadow of pre-unification school systems into Italy's Liberal Age (1861–1911). *J Econ Hist* (forthcoming)

- Brambilla E (1991) Alfabetismo e società nelle Marche in Età Napoleonica. In: Bartoli Langeli A, Toscani X (eds) *Istruzione, alfabetismo, scrittura. Saggi di storia dell'alfabetizzazione in Italia (sec. XV-XIX)*. Franco Angeli, Rome
- Braudel F (1949) *La Méditerranée et le monde méditerranéen à l'époque de Philippe II*. A. Colin, Paris
- Cappelli G (2017) Was Putnam Wrong? The determinants of social capital in Italy around 1900. *Rivista Di Storia Economica* 33:277–314. <https://doi.org/10.1410/89359>
- Cappelli G, Quiroga Valle G (2020) Literacy and Schooling in Italy and Spain (1860–1921): a comparative analysis. *Rivista Di Storia Economica* 36:87–123. <https://doi.org/10.1410/99052>
- Cappelli G, Quiroga Valle G (2021) Female teachers and the rise of primary education in Italy and Spain, 1861–1921: evidence from a new dataset. *Econ Hist Rev* 74:754–783. <https://doi.org/10.1111/ehr.13068>
- Cappelli G, Vasta M (2020a) A “Silent Revolution”: school reforms and Italy’s educational gender gap in the Liberal Age (1861–1921). *Cliometrica* 15:1–27. <https://doi.org/10.1007/s11698-020-00201-6>
- Cappelli G, Vasta M (2020b) Can school centralization foster human capital accumulation? A quasi-experiment from early twentieth-century Italy. *Econ Hist Rev* 73:159–184. <https://doi.org/10.1111/ehr.12877>
- Chen Y, Naidu S, Yu T, Yuchtman N (2015) Intergenerational mobility and institutional change in 20th century China. *Explor Econ Hist* 58:44–73. <https://doi.org/10.1016/j.eeh.2015.07.001>
- Ciccarelli C, Weisdorf J (2019) Pioneering into the Past: Regional literacy developments in Italy before Italy. *Eur Rev Econ Hist* 23:329–364. <https://doi.org/10.1093/ereh/hey014>
- Cinnirella F, Hornung E (2016) Landownership concentration and the expansion of education. *J Dev Econ* 121:135–152. <https://doi.org/10.1016/j.jdeveco.2016.03.001>
- Cipolla CM (1952) The decline of Italy: The case of a fully matured economy. *Economic History Review* 5(2):178–187
- Cipolla CM (1969) *Literacy and development in the west*. Penguin, London
- Cipolla CM (1989) *Le tre rivoluzioni e altri saggi di storia economica e sociale*. Il Mulino, Bologna
- Ciccarelli C, Fachin S (2017) Regional growth with spatial dependence: a case study on early Italian industrialization. *Papers Reg Sci* 96:675–695. <https://doi.org/10.1111/pirs.12217>
- Chilosi D, Ciccarelli C (2022) Evolving gaps: occupational structure in southern and northern Italy, 1400–1861. *Econ Hist Rev* 75:1349–1378. <https://doi.org/10.1111/ehr.13159>
- Chilosi D (2007) Old wine in new bottles: civic nation-building and ethnic nationalism in schooling in Piedmont, ca. 1700–1861. *Nations Natl* 13:417–436. <https://doi.org/10.1111/j.1469-8129.2007.00294.x>
- Clark G (2005) The condition of the working class in England, 1209–2004. *J Polit Econ* 113:1307–1340. <https://doi.org/10.1086/498123>
- Cvrček T, Zajicek M (2019) The making of a liberal education: Political economy of the Austrian school reform, 1865–1880. *Expl. Econ. Hist.* 73:101251. <https://doi.org/10.1016/j.eeh.2018.07.003>
- Daniele V, Malanima P (2011) *Il Divario Nord-sud in Italia, 1861–2011*. Rubbettino Editore, Catanzaro
- Dincecco M, Federico G (2021) Napoleon in Italy: A legacy of institutional reform? In: O’Brian K (ed) *The crucible of revolutionary and Napoleonic warfare and European transitions to modern economic growth*. Brill, Leiden
- Dincecco M, Federico G, Vindigni A (2011) Warfare, taxation, and political change: evidence from the Italian Risorgimento. *J Econ Hist* 71:887–914. <https://doi.org/10.1017/S002205071100218X>
- Federico G, Nuvolari A, Ridolfi L, Vasta M (2019a) The race between the snail and the tortoise: skill premium and early industrialization in Italy (1861–1913). *Cliometrica* 15:1–42. <https://doi.org/10.1007/s11698-019-00200-2>
- Federico G, Nuvolari A, Vasta M (2019b) The origins of the Italian regional divide: evidence from real wages, 1861–1913. *J Econ Hist* 79:63–98. <https://doi.org/10.1017/S0022050718000712>
- Felice E (2013) *Perché il Sud è rimasto indietro*. Il Mulino, Bologna
- Felice E (2014) Il Mezzogiorno fra storia e pubblicistica. Una replica a Daniele e Malanima. *Rivista Di Storia Economica* 30:197–242
- Felice E (2019) Regional income inequality in Italy in the long run (1871–2001). Patterns and determinants. In: Roses J, Wolf N (eds) *The Economic development of Europe’s regions. A quantitative history since 1900*. Routledge, London
- Felice E, Vasta M (2015) Passive modernization? The new human development index and its components in Italy’s regions (1871–2007). *Eur Rev Econ Hist* 19:44–66. <https://doi.org/10.1093/ereh/heu018>

- Ferraresi A (1991) Alfabeti e analfabeti nell'Oltrepò pavese alla fine dell'Antico regime: Voghera, Stradella, Casteggio, Casatisma. In: Bartoli A, Toscani X (eds) *Istruzione, alfabetismo, scrittura. Saggi di storia dell'alfabetizzazione in Italia*. Franco Angeli, Rome
- Fleury M, Valmary P (1957) *Les progrès de l'instruction élémentaire de Louis XIV à Napoléon III, d'après l'enquête de Louis Maggiolo (1877–1879)*. Population 71–92
- Gargano A (2012) Numeri in dubbio. Scuola pubblica e scuola privata nell'Italia meridionale attraverso l'inedita inchiesta del 1861. In: Bianchi A (ed) *L'istruzione in Italia*. La Scuola, Brescia
- Genovesi G (2010) *Storia della scuola in Italia dal Settecento a oggi*. Laterza, Rome-Bari
- Griseri G (1973) *L'istruzione primaria in Piemonte: 1831–1856*. Deputazione subalpina di storia patria, Turin
- Houston RA (2014) *Literacy in early modern Europe*. Routledge, London
- Lupo M (2005) *Tra le provvide cure di Sua Maestà: stato e scuola nel Mezzogiorno tra Settecento e Ottocento*. Il Mulino, Bologna
- Lupo M, Gargano A, Marra A (2014) Collezione delle Leggi de' Decreti e di altri atti riguardante la Pubblica Istruzione promulgati nel già Reame di Napoli dall'anno 1806 in poi. ISMED - Consiglio Nazionale delle Ricerche (CNR), Naples
- Macry P (2002) *Ottocento. Famiglia, élites e patrimoni a Napoli*. Il Mulino, Bologna
- Malanima P (2005) Urbanisation and the Italian Economy during the Last Millennium. *Eur Rev Econ Hist* 9:97–122. <https://doi.org/10.1017/S1361491604001327>
- Malanima P (2016). Italian urban population 1300–1861 (The Database)
- Malanima P, Daniele V (2007) Il prodotto delle regioni e il Divario Nord-Sud in Italia (1861–2004). *Rivista di Politica Economica* 97:1–49
- Marchesini D (1992) Il Bisogno di scrivere: usi della scrittura nell'Italia moderna. Laterza, Rome-Bari
- Mariella V, Postigliola M, Rota M (2020) A reconstruction of the Italian provincial population. *Rivista Di Storia Economica* 36:1–39. <https://doi.org/10.1410/98355>
- Mariella V (2022) The agrarian origins of social capital. *J Econ Behav Organ* 193:543–568. <https://doi.org/10.1016/j.jebo.2021.11.029>
- Mastroluca S, Verrascina M (2012) L'evoluzione dei contenuti informativi del censimento della popolazione. *Annali di statistica* 12
- Meier zu Selhausen F, Van Leeuwen MH, Weisdorf JL (2018) Social mobility among Christian Africans: evidence from Anglican marriage registers in Uganda, 1895–2011. *Econ Hist Rev* 71(4):1291–1321. <https://doi.org/10.1111/ehr.12616>
- Milanesi A (1991) La città e il borgo: un caso padano. In: Bartoli Langeli A, Toscani X (eds.) *Istruzione, alfabetismo, scrittura. Saggi di storia dell'alfabetizzazione in Italia*. Franco Angeli, Rome
- Milner B (2021). The impact of state-provided education: evidence from the 1870 education act. *Mimeo*
- Mitch DF (1986) The impact of subsidies to elementary schooling on enrolment rates in nineteenth-century England. *Econ Hist Rev* 39:371–391. <https://doi.org/10.1111/j.1468-0289.1986.tb00410.x>
- Montalbo A (2020) Industrial activities and primary schooling in early nineteenth-century France. *Cliometrica* 14:325–365. <https://doi.org/10.1007/s11698-019-00191-0>
- Nuvolari A, Vasta M (2017) The geography of innovation in Italy, 1861–1913: evidence from patents data. *Eur Rev Econ Hist* 21(3):326–356. <https://doi.org/10.1093/ereh/hex006>
- Pellizzari MR (1987) Alfabeto e fisco. Tra cultura scritta e oralità nel Regno di Napoli a metà Settecento. In: Pellizzari MR (ed) *Sulle vie della scrittura: alfabetizzazione, cultura scritta e istituzioni in età moderna*. Atti del convegno di Studi del centro studi "Antonio Genovesi" per la storia economica e sociale. Edizioni Scientifiche Italiane, Salerno, 99–152
- Pellizzari M (2000) *La penna e la zappa. Alfabetizzazione, cultura e generi di vita nel Mezzogiorno moderno*. Laveglia e Carlone editore, Salerno
- Piseri M (2002) *L'alfabeto delle riforme: Scuola e alfabetismo nel basso Cremonese da Maria Teresa all'Unità*. Vita e Pensiero, Milan
- Piseri M (2017) *La scuola primaria nel Regno italico: 1796–1814*. FrancoAngeli, Rome
- Petrucci A (1987) Prospettive di Ricerca e Problemi di Metodo per una Storia Quantitativa dell'Alfabetismo. In: Pellizzari MR (ed) *Sulle vie della scrittura: alfabetizzazione, cultura scritta e istituzioni in età moderna*. Atti del convegno di Studi del centro studi "Antonio Genovesi" per la storia economica e sociale. Edizioni Scientifiche Italiane, Salerno, 21–27
- Postigliola M, Rota M (2020) Institutions and literacy rates: the legacy of Napoleonic reforms in Italy. *Eur Rev Econ Hist* 25:757–779. <https://doi.org/10.1093/ereh/hea021>
- Postigliola M, Rota M (2021) Institutional shocks and social capital: evidence from the French domination in Italy. *J Mod Ital Stud* 1:1–24. <https://doi.org/10.1080/1354571X.2021.1912251>

- Poussou JP (1993) La méthode Maggiolo et la mesure de l'alphabétisation du Sud-Ouest de la France à la fin du XVIIIème siècle. *Annales Du Midi* 105:209–223
- Pruneri F (2011) L'istruzione in Sardegna 1720–1848. Il Mulino, Bologna
- Putnam RD, Leonardi R, Nanetti RY (1994) Making democracy work: civic traditions in modern Italy. Princeton University Press, Princeton
- Rediker M (1989) Between the devil and the deep blue sea: merchant seamen, pirates and the Anglo-American maritime world, 1700–1750. Cambridge University Press, Cambridge
- Rettaroli R (1990) Age at marriage in nineteenth-century Italy. *J Fam Hist* 15:409–425. <https://doi.org/10.1177/036319909001500123>
- Rossi N, Toniolo G, Vecchi G (2001) Is the Kuznets curve still alive? Evidence from Italian household budgets, 1881–1961. *J Econ Hist* 61:904–925. <https://doi.org/10.1017/S0022050701042024>
- Sargant WL (1867) On the progress of elementary education. *J Stat Soc Lond* 30:80–137
- Scirocco A (1987) Note sull'alfabetizzazione a Napoli nell'Ottocento attraverso i Registri Matrimoniali. In Pellizzari MR (ed.) *Sulle vie della scrittura: alfabetizzazione, cultura scritta e istituzioni in età moderna*. Atti del convegno di Studi del centro studi "Antonio Genovesi" per la storia economica e sociale, Edizioni Scientifiche Italiane, Salerno, 259–291
- Schofield RS (1968) The measurement of literacy in pre-industrial England. *Lit Trad Soc* 10:310–326
- Schofield RS (1973) Dimensions of Illiteracy, 1750–1850. *Explor Econ Hist* 10:437–454. [https://doi.org/10.1016/0014-4983\(73\)90026-0](https://doi.org/10.1016/0014-4983(73)90026-0)
- Shaw-Taylor L, Wrigley EA (2014) Occupational structure and population change. In: Floud R, Johnson R (eds) *The Cambridge economic history of modern Britain*, vol 1. Cambridge University Press, Cambridge, pp 53–88
- Som RK (2020) *Practical sampling techniques*. CRC Press, Boca Raton
- Tapia FJB, Martinez-Galarraga J (2018) Inequality and education in pre-industrial economies: evidence from Spain. *Explor Econ Hist* 69:81–101. <https://doi.org/10.1016/j.eeh.2017.12.003>
- Toscani X (1987) L'alfabetismo nelle campagne dei Dipartimenti del Mincio e del Mella e nelle Alte Valli del Serio e dell'Adda. In: Pellizzari MR (ed) *Sulle vie della scrittura: alfabetizzazione, cultura scritta e istituzioni in età moderna*. Atti del convegno di Studi del centro studi "Antonio Genovesi" per la storia economica e sociale. Edizioni Scientifiche Italiane, Salerno, 549–6093
- Terzi L (2001) Le scuole Normali a Napoli tra Sette e Ottocento. Documenti e ricerche sulla "pubblica uniforme educazione" in antico regime, L'Orientale Editrice, Naples
- Van Leeuwen MH, Maas I, Miles A (2002) *HISCO: historical international standard classification of occupations*. Leuven University Press, Leuven
- Van Leeuwen MH, Maas I (2011) *HISCLASS: a historical international social class scheme*. Leuven University Press, Leuven
- Vasta M (1999) Capitale umano e ricerca scientifica e tecnologica. In: Amatori F, Bigazzi D, Giannetti R, Segreto L (eds) *Storia d'Italia, Annali*. Einaudi, Turin, pp 1041–1124
- Vigo G (1971) *Istruzione e sviluppo economico in Italia nel secolo XIX*. ILTE, Turin
- Vigo G (2017) *Il vero sovrano dell'Italia: l'istruzione degli adulti nell'Italia dell'Ottocento*. Il Mulino, Bologna
- Zamagni V (1993) *The Economic History of Italy 1860–1990*. Clarendon Press, Cambridge
- Zazo A (1927) *L'istruzione pubblica e privata nel Napoletano (1767–1860)*. Il Solco, Città di Castello

Primary sources

- Antenati – Ministero per i Beni e le Attività Culturali e per il Turismo (2023) In: Buzzi AM, Vitali S, Corbellini R, Corbo A, Ferrari D, Gagliardoni L, Garis E, Maestrini F, Salemme F, Vedovotto N, and Nicoletti R (eds) Antenati. <http://dl.antenati.san.beniculturali.it/gallery> Accessed 24 Sept 2022
- ASI – Annuario Statistico Italiano (1853) *Annuario Economico-Statistico dell'Italia* (Vol. 1027). Turin, Ferrero e Franco.
- dell'Istruzione M (1890) *Delle condizioni della istruzione elementare in Italia e del suo progresso dal 1861 in poi*. Stabilimento Tipografico Sinimberghi, Rome
- MAIC – Ministero dell'Agricoltura, dell'Industria e del Commercio (1866a) *Censimento generale della popolazione del 31 dicembre 1861 Vol. II*. Tipografia Letteraria, Florence
- MAIC – Ministero dell'Agricoltura, dell'Industria e del Commercio (1866b) *Censimento generale della popolazione del 31 dicembre 1861 Vol. III*. Tipografia Letteraria Florence

- MAIC – Ministero dell’Agricoltura, dell’Industria e del Commercio (1875) Censimento generale della popolazione del 31 dicembre 1871 Vol. II. Tipografia Cenniniana, Rome
- MAIC – Ministero dell’Agricoltura, dell’Industria e del Commercio (1883) Censimento della popolazione del Regno d’Italia al 31 dicembre 1881. Vol. III. Tipografia Bodoniana, Rome
- MAIC – Ministero dell’Agricoltura, dell’Industria e del Commercio, Direzione Generale di Statistica (1900) Studi e proposte per l’esecuzione del IV censimento generale della popolazione del Regno. Tipografia Nazionale di G. Bertero, Rome

Secondary sources

- Bianchi A (2019a) L’istruzione in Italia tra Sette e Ottocento. Dal Regno di Sardegna alla Sicilia Borbonica: istituzioni scolastiche e prospettive educative. La Scuola, Brescia
- Bianchi A (2019b) La scuola primaria nell’alto Piemonte-Valle d’Aosta Napoleonici. I dipartimenti della Dora e della Sesia. In: Angelo B (ed) L’istruzione in Italia tra Sette e Ottocento. Dal Regno di Sardegna alla Sicilia Borbonica: istituzioni scolastiche e prospettive educative. La Scuola, Brescia
- Blok A (1969) South Italian Agro-towns. *Comp Stud Soc Hist* 11:121–135. <https://doi.org/10.1017/S0010417500005235>
- Curtis D (2013) Is there an “Agro-town” Model for Southern Italy? Exploring the diverse roots and development of the agro-town structure through a comparative case study in Apulia. *Contin Chang* 28:377. <https://doi.org/10.1017/S0268416013000362>
- King R, Strachan A (1978) Sicilian Agro-Towns (Sizilianische Agro-Städte). *Erdkunde*, 110–123
- Marzeddu R (2019) Fonti per la storia dell’istruzione in Piemonte tra Sette e Ottocento. Spunti e itinerari di ricerca. In: Angelo B (ed.) L’istruzione in Italia tra Sette e Ottocento. Dal Regno di Sardegna alla Sicilia Borbonica: istituzioni scolastiche e prospettive educative. I-Studi; II-Carte Storiche. La Scuola, Brescia, 57–75
- Salvo C, Zichichi L (2003) La Sicilia dei signori: il potere nelle città demaniali. Sellerio, Palermo
- Toscani X (1993) Scuole e alfabetismo nello Stato di Milano da Carlo Borromeo alla Rivoluzione. La Scuola, Brescia

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