Journal of Economic Growth
https://doi.org/10.1007/s10887-025-09263-5

Keeping up with the Joneses? The rise of modern universities and local economic development in Italy

Elena Cottini[1] · **Paolo Ghinetti**[2] · **Simone Moriconi**[3]

Received: 6 June 2022 / Accepted: 11 August 2025

Abstract

This paper uses own built historical register tracing the evolution of Italian higher education from the High Middle Ages to 2010 to analyse the economic effects of the rise of modern universities. The registry data prompts the creation of a panel dataset measuring higher education supply across 110 Italian provinces from 1870 to 2010, which corresponds to the transition from elite to mass higher education in Italy. Our econometric approach combines a fixed effects estimator with Bartik-style instruments that exploit the historical cross-province distribution of higher education institutions in Italy. We find significant economic returns of universities, particularly in synergy with industrialisation, contributing to Italy's transition to sustained growth. Our estimates predict that a province exposed to an average supply (equivalent to 7 universities, approximately) experiences an annual 2% increase in real gross value added per capita compared to a province with a zero supply. We show that about two thirds of these returns arise locally, while the remaining third originates from externalities. We argue that such high local concentration of economic returns contribute to fostering inequalities between small peripheral areas and large university hubs.

Keywords Productivity · Higher education supply · Neighbourhood effects · Historical data · Unified growth theory

JEL Classification I23 · I28 · N00 · R1

✉ Paolo Ghinetti
paolo.ghinetti@uniupo.it

Elena Cottini
elena.cottini@unicatt.it

Simone Moriconi
s.moriconi@ieseg.fr

[1] Department of Economics and Finance, Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123 Milano, MI, Italy

[2] Department of Economics and Business Studies, Università del Piemonte Orientale, Via Perrone 18, Novara (NO) Vercelli, Italy

[3] IESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France

Published online: 12 January 2026

1 Introduction

The emergence of modern higher education systems was closely linked to the transformations of the Industrial Revolution, playing a crucial role in fostering economic growth through human capital development. By the late 19th century, rapid industrial expansion and productivity gains created an increasing demand for skilled workers, essential for sustaining technological progress (Galor, 2011). In response, higher education institutions evolved from elite academic centers into more accessible institutions designed to equip a broader workforce with necessary skills (Galor and Moav, 2006). This transition was pivotal in shaping long-term economic trajectories, as investments in human capital not only supported industrialisation but also laid the foundation for sustained growth. The expansion of higher education thus represented a key structural shift, ensuring that labor markets could meet the evolving demands of a more complex and technologically advanced economy.

Despite the well-documented relationship between human capital accumulation and economic growth, the specific economic effects of higher education institutions—both locally and through regional spillovers—remain insufficiently understood. Research has shown that large stocks of human capital contributed significantly to GDP per capita growth in both developed and developing countries during the 20th century (Hanushek and Woessmann, 2015). However, because student and worker mobility is often geographically limited, these economic gains are likely concentrated in specific areas. This is the probable reason why OECD countries have strategically expanded their higher education systems with the goal of fostering regional economic development (OECD, 2007, 2014).

Yet, in this expansion process, regions may find themselves in a dynamic akin to the classic 'Keeping up with the Joneses' phenomenon, a term originally derived from the comic strip by Momand (1913). In this metaphor, rich regions with established universities play the role of the wealthier family, the Joneses, while poorer or less developed areas resemble the struggling McGinis family, striving to match their neighbours' educational and economic opportunities. The expansion of higher education may thus generate positive spillovers — through knowledge diffusion and productivity gains — but it may also create competitive pressures that exacerbate disparities, as underprivileged regions compete for limited resources such as students, faculty, and investment.

Understanding the full scope of these effects is essential for assessing the net economic benefits of higher education institutions and their cross-province distribution. If universities disproportionately benefit wealthier regions while limiting opportunities to keep up for others, they may inadvertently prevent regional economic convergence. From a policy perspective, quantifying the extent to which these effects are localised versus widespread is a necessary first step in evaluating whether the expansion of higher education may have contributed to the persistent regional differentials in income per capita since the end of WWII (Gennaioli et al., 2013).

This paper analyses the productivity effects of the rise of the modern supply of higher education in Italy, the country with the longest history of university education in the western world. Over Italian history, we register the expansion of university education at the level of the province. This is the oldest layer of governance in the Italian territory, historically responsible for the provision of several local public services and goods, and equivalent to the *départment* in France or the county in the US or UK (Federico et al., 2019). We evaluate the productivity effects of the rise of university institutions, by conducting a twofold empirical exercise. First, we estimate neighbourhood effects, i.e., the association between supply of higher education institutions in nearby provinces and the local supply. Second, we estimate economic returns from higher education institutions in the local province and its neighbourhood. We use estimated coefficients from these two empirical exercises to predict the actual distribution of economic returns in terms of productivity, and separate local economic returns from the productivity externalities of neighbouring provinces.

To analyse the long-term impact of higher education institutions, we constructed the History of Italian Universities (HIU) registry, a comprehensive dataset that traces the evolution of university institutions from the High Middle Ages to 2010. The HIU registry documents the historical development of higher education, from its origins as an elite-driven system to the democratisation of access in the 20th century (Schofer and Meyer, 2005). Using this registry, we built a panel dataset covering 110 Italian provinces between 1870 and 2010, measuring higher education supply through the total number of "faculties" in each province per year. In the Middle Ages, faculties functioned as the primary teaching units responsible for curriculum design and instruction in specific academic disciplines (de Ridder-Symoens, 1992).[1] This structure remained largely unchanged until the 2010 reform, which introduced a new organisational model for Italian universities. By focusing on faculties as the fundamental units of higher education supply, our dataset enables a detailed analysis of the role universities played in regional economic development over time. This historical perspective is crucial for understanding how institutional structures evolved (e.g., in terms of disciplines taught) and how they contributed to shaping Italy's educational and economic landscape.

To ensure a rigorous empirical analysis, we employ a fixed effects (FE) estimator as our primary econometric approach and complement it with an instrumental variable (IV) strategy to address endogeneity concerns. Our panel dataset spans a long time period and offers fine geographical disaggregation at the province level, making the FE estimator a natural benchmark for controlling for unobserved heterogeneity. However, to mitigate concerns related to reverse causality and time-varying sorting of faculties, we implement an IV approach based on Bartik-style instruments. Specifically, we construct instruments by interacting the historical cross-province distribution of faculties with the subsequent faculty growth process over Italian history. This approach leverages the fact that Italian provinces inherited their higher education institutions from pre-unitarian sovereign states, and we use historical data from the HIU registry to reconstruct initial faculty distributions in different preunitarian epochs, from the 13th century onwards. Our empirical framework assumes that these initial conditions had a persistent influence on the creation of new faculties, even after Italy's unification in the 19th century.[2] By incorporating these instruments, we account for potential omitted variables that could bias our estimates. Throughout the analysis, we extensively evaluate the validity of our Bartik instruments and conduct a series of robustness checks to ensure that our empirical strategy is not compromised by unobserved fac-

[1] In essence, a faculty is the core teaching unit, and it is meant to govern the supply of tertiary education in a given field of study (e.g., the Faculty of Economics, the Faculty of Law, etc.). The use of the term faculty should not be confused with its alternative use to denote the body of academic personnel, which is rather common among Anglo-Saxon universities. See Clark (2006) for details on historical traditions in Anglo-Saxon universities. We provide additional details on the Italian tertiary education system, also from an historical perspective, in Section 2.

[2] The political unification of the Italian territory into the sovereign state of Italy started in 1861 and was completed in 1870 with the capture of Roma. The newly born state was a monarchy (Kingdom of Italy) until 1946, when following an institutional referendum, a new republican constitution was adopted.

tors correlated with the historical distribution of faculties (Goldsmith-Pinkham et al., 2020; Adão et al., 2020).

Our results suggest there are sizeable and significant economic returns of higher education institutions, which are quite localised at the province level. We measure economic returns in terms of real Gross Value Added (GVA, from here on) per capita. Our estimates predict that a province exposed to the average supply of faculties in 2001 (28 faculties corresponding to approximately 7 universities either in the local province or the neighbourhood) presents, ten years later, a 20% higher real GVA per capita compared to a province with a zero supply. This is equivalent to an annual 2% increase in GVA per capita between 2001 and 2010 (i.e., an elasticity of GVA per capita to higher education supply equal to 0.2). The 64% of such returns (equivalent to an annual 1.3% increase of GVA per capita between 2001 and 2010) are due to local faculties, while faculties nearby explain the remaining 36% (equivalent to an annual 0.7% increase between 2001 and 2010). Finally, we show that the high geographical concentration of economic effects of universities contribute to inequalities between small peripheral provinces that enjoy very small annual returns (often below 1%) and large provinces where the supply of faculties explains annual increases up to 7% of local GVA per capita.

Our paper connects two strands of the literature that investigates the economic effects of higher education institutions. The first strand focuses on the pre-industrial era, i.e., an epoch where upper-tail education was reserved to a minority of elite interested in scientific progress (Mokyr, 2005). Most studies in this strand exploit cross-sectional variation in upper-tail knowledge and educational outcomes to analyse the role of human capital as a determinant of pre-industrial economic development (e.g., in terms of population growth, urban population density, and living standards) and the emergence of industrialisation (e.g., Squicciarini and Voigtländer, 2015; Becker et al., 2011). Up to our knowledge, there are only two studies in this literature that use indicators of supply of university institutions (rather than education outcomes) to measure the economic effects of human capital. Both studies look at Germany. Cantoni and Yuchtman (2014) exploit overtime variation associated with the foundation of 37 Medieval universities to study the role of distance from a university as a determinant of the establishment of markets after the papal schism in 1386. More recently, Dittmar and Meisenzahl (2022) investigate how distance from research universities affected innovation and industrial change in the 18th and 19th century.

The second strand of the literature analyses the effect of higher education on productivity and growth in postindustrial societies. The use of indicators of supply of higher education supply is more common among these studies. Andersson et al. (2004) use a 14-year panel dataset for Swedish municipalities to study the economic effects of higher education decentralisation on the output per worker. Valero and Van Reenen (2019) estimate economic returns of higher education institutions using the World Higher Education Database (WHED) a regional panel measuring supply of universities across 78 countries for the post-WWII period.

Our study provides new data and a new empirical methodology, which allows us making three major contributions to this literature. First, thanks to the HIU registry, we construct a novel cross-province panel dataset covering 150 years in one country, Italy, which is the western country with the longest history of higher education. Thus, we observe considerable variation in higher education, a persistent institution whose supply evolves very slowly over time. This makes it possible to identify the effects of higher education institutions from the onset of the industrial era to today, a period marked by two major transitions i.e. the transition from elite to mass education, and the transition from the Post-Malthusian to the Sustained regime of Economic Growth. This is the likely reason why we observe sizeable effects of higher education supply on GVA per capita, as well as small, yet noticeable effects of local supply also on local population and population density in the early 20th century, which we interpret as Post-Malthusian returns of upper tail human capital.

Second, compared to existing studies, with the HIU data we are able to measure higher education supply at the level of the faculty, which is the most disaggregated supply unit of higher education. This finest level of disaggregation allows investigating field-specific returns, distinguishing the two macro-areas of STEM (including faculties in Science, Technology, Engineering and Mathematics), and non-STEM education (including faculties in the fields of Humanities and Social Sciences). We find returns (particularly of STEM faculties) are larger in provinces characterised by higher levels of early industrialisation, which carries over to a gradual transition of sectoral employment shares away from agriculture towards industry and services. These results align with the complementarity hypothesis, which underscores the interplay between the development of human capital and technological progress during the shift from the Malthusian era to the era of Sustained Growth. This hypothesis is central to the Unified Growth Theory (Galor and Weil, 2000; Galor and Moav, 2004).

As a third important contribution, our fine-grained province-level data enable us to advance the measurement of geographical externalities. In particular, we find out there is a cost associated with the existence of faculties in a sufficiently close neighbourhood (e.g., in the same region, within a certain distance, etc.), which prevents opening a similar faculty in a province. Our results suggest a one standard deviation increase in higher education supply in each neighbour j (4.61 faculties) results in up to -2.95 faculties in province i over a ten years' period. We argue that this *displacement cost* emerges as long as there is some individual mobility within the neighbourhood (e.g. due to historically good railway infrastructures, or a high geographic accessibility of provinces), but high commuting costs outside that area so that each province insists on a limited portion of national demand for education. Discussion of this displacement cost is new in the literature. Andersson et al. (2004) and Valero and Van Reenen (2019) discuss only 'gross' positive externalities, thus neglect displacement forces from the neighbourhood. Our analysis suggests that this neglect leads to severely overestimate (by about 50%) the positive geographical externalities of higher education supply in productivity terms.

Finally, we also make a fourth contribution to the side of identification. Historical papers stress the role of cross-sectional variation coming from the regional distribution of initial conditions to identify the economic effect of upper-tail knowledge. Cantoni and Yuchtman (2014) show that the establishment of Germany's first universities in the Middle Ages had a causal impact on medieval Europe's Commercial Revolution. Becker et al. (2011) use the education level observed before the start of industrialisation (i.e., in 1816) as an instrument to identify the effect of education on industrialisation in Prussia. Andersson et al. (2004) and Valero and Van Reenen (2019) use longitudinal data and estimate fixed effects models at the sub-national level, but they lack suitable instruments, and focus on the post-WWII period. As in this literature, we use instruments whose validity depends on the exogeneity of the distribution of initial conditions. However, our Bartik instrument combines cross-sectional variation with time variation coming from faculty creation during Italian history. This allows us to incorporate IVs into an empirical specification that also includes fixed effects both at the province and region-by-year level. This provides new insights on the causal identification of mechanisms that are well-established in the context of Unified Growth Theory (Galor, 2011).

The paper is organised as follows. In the next section, we describe the institutional and conceptual framework employed. In Section 3, we present the data. The empirical strategy is presented in Section 4, while main results of neighbourhood and productivity effects are in Section 5. Section 6 discusses the potential channels driving neighbourhood and productivity effects of higher education supply. Section 7 presents the welfare analysis. Section 8 concludes.[3]

2 Academic institutions and economic development in Italy (1200-2010)

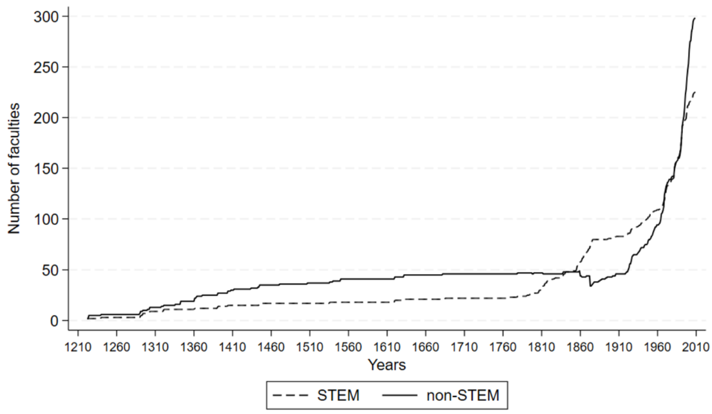
In this section, we explore how the higher education supply in Italian provinces evolved along Italian history. We will map higher education supply across economic development phases, as outlined by Unified Growth Theory (Galor, 2011). We will also address the importance of accounting for local effects on higher education provision. Many OECD countries treat higher education as a local market influenced by market forces rather than central government intervention (Catalano and Silvestri, 1999; Agasisti and Catalano, 2007). This rationale underpins considering local effects when analysing how higher education institutions contribute to economic development.

2.1 The rise of academic education from Malthusian stagnation to modern growth

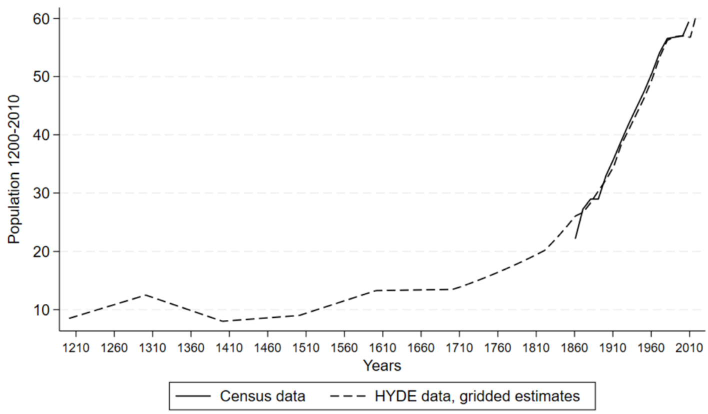
Figure 1 shows the evolution of higher education supply and population on the Italian territory between 1200 and 2010. It evidences some comovement between population and faculty creation, particularly starting from the second half of the 18th century. We can distinguish three broad phases in the development of higher education institutions (cfr. Panel (a)). Until roughly 1750, the group of faculties comprising Humanities, Law and Theology (we call them non-STEM faculties) outweighs the group comprising Mathematics, Natural Sciences, and Medicine (we denote them by STEM faculties), which conveyed the main clusters of scientific knowledge at that time (Curtis and de la Croix, 2023). During this epoch the growth of the number of faculties is balanced between these two groups, as academic knowledge was generalist by its own nature.[4] Relevant political actors in decisions to open an academic institution were the Holy Roman Empire (as in the case of faculties founded in Pavia in the 14th century), the Kingdom of Sicily (as in the case of the faculty of

[3] The paper also includes an Appendix that accompanies the main text (with equations, tables and figures numbered ‘A-’ followed by Arabic numerals), and an online Appendix (with tables and figures numbered ‘A-’ followed by Roman numerals).

[4] Medieval *Studia Generalia* always included faculties of Law, Medical Studies, Theology. In addition, there was a faculty of Liberal Arts, which included the Literary studies (*Trivium*, including grammar, rhetoric, dialectic), and Mathematics and Natural Sciences (*Quadrivium*, including arithmetic, geometry, music, astronomy). To foster comparability with the modern university system that arises in the 18th and 19th centuries, we treat Literary studies and Mathematics and Natural Sciences as two independent faculties belonging to the non-STEM and STEM group, respectively.



(a) Higher education supply 1200-2010



(b) Population 1200-2010

Notes: Panel (a) reports the number of preunitarian faculties in Italy between 1200 and 1861, distinguishing between STEM and non-STEM disciplines. The STEM group includes Agricultural Studies, Architecture, Chemistry, Pharmacy and Medical Studies, Geology and Biology, Engineering and Scientific Studies. The non-STEM group includes Social Sciences, Humanities, and Theology. Social Sciences include Law and, in the modern era, Economics and Statistics, and Sociopolitical Studies. Humanities include Education, Linguistic Studies, Literature, and Psychology. Panel (b) reports the level of population measured using data from Italian census 1861-2010 (grey line, scale measured over the left-hand side y-axis) and the estimated population based on gridded data at NUTS3-province level from HYDE3.2 data (black line, scale measured on the right-hand side yaxis). Authors' calculations based upon HYDE3.2, Census and HIU data.

Fig. 1 Population and faculty creation during 1200-2010

Law in Napoli in the 13th century), or the Spanish empire (as Catania in the 15th century), as well as smaller Initial states such as House of Este in Ferrara, the Republic of Genova, the Commune of Siena starting from the 13th century (see more on this in Section 2.2 below).

The second phase in the rise of higher education starts around 1760, until the end of WWII. During this period there is an acceleration in the growth of STEM faculties that outweigh non-STEM ones. This epoch is marked by the extension of the Napoleonic model of higher education to Italian territories, which led to the creation of faculties in novel STEM fields. This aimed at training professionals, favour the development of technologies, and create better and more prepared public employees (Alvazzi del Frate, 2021; Kickert, 2007).[5] Relatedly, during this period there is the decline of faculties of Theology, which determines the drop in the number of non-STEM faculties between 1860 and 1873. Theology faculties lose appeal during the Enlightments, and are legally suppressed by the new Kingdom of Italy in 1873 (see online Appendix A3 for details). For this reason we decide not to consider them for the rest of the analysis.

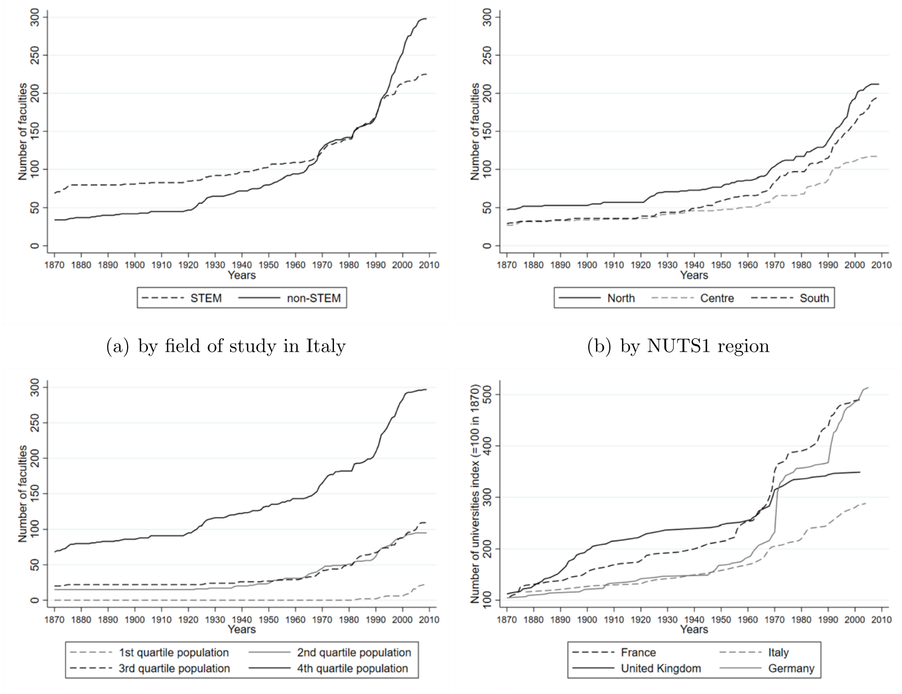
The third phase covers the Post-War period until 2010. This was the time in which the Italian university system, still internally organised to serve the ‘elite’, was asked to provide tertiary education for the ‘mass’. In online Appendix A2 we outline the major reforms that shaped the Italian university system during this period. This was marked by the liberalisation of the university access to all students with a 5 year diploma of secondary education, including those from technical schools, in 1969 (see also Eurydice, 2000). The consequent significant surge in the university student population exerted substantial pressure on the university system, which was compelled to expand the number of faculties. This expansion is concentrated into non-STEM disciplines (particularly Social Sciences), which end up outweighing STEM faculties again towards the end of the 20th century.

Panel (b) displays population levels between 1200 and 2010.[6] The figure shows a sharp rise in population, alongside the growth in the number of STEM faculties, starting roughly from the second half of the 18th century. Such co-growth of population and higher education supply suggests it is important to account for Malthusian forces possibly correlated with the pre-unitarian supply of faculties (particularly STEM ones) in the 19th century. In Sections 4 and 5 we discuss extensively these concerns, and how we tackle them in the empirical analysis.

Figure 2 provides deeper insights into the evolution of higher education supply throughout the period 1870-2010. Panel (a) allows to better appreciate the surge in higher education provision, encompassing particularly non-STEM faculties in the second half of the 20th century. Panels (b) and (c) indicate that the total count of faculties increases across all

[5] New STEM faculties were established either as independent schools specialising in applied or technical fields—such as Architecture and Engineering within the Natural Sciences faculties—or as technical institutions operating outside the formal university system, including Pharmacy and Veterinary Studies. These new applied schools were unequivocally equivalent to university faculties, conferring a diploma of the same status as a university degree. Possession of this diploma was essential for employment in the Italian public sector or entry into field-specific highly skilled professions. Most of these specialised schools were established in the main capital cities of each Pre-unitarian State (for further details, refer to online Appendix A1).

[6] The longest time series (grey line in Panel (b)) is based upon estimated gridded data, so it may suffer from some lack of precision. However, comparison with the actual data from population census for the sample period (black line in Panel (b)) reveals an almost complete overlap between the two series. This reassures us over the reliability of the estimated gridded data for the pre-sample period.



(a) by field of study in Italy

(b) by NUTS1 region

(c) by province population quartile

(d) no. of universities in EU4

Notes: Panel (a) reports the number of faculties in Italy by field of study. The STEM group includes Agricultural Studies, Architecture, Chemistry, Pharmacy and Medical Studies, Geology and Biology, Engineering and Scientific Studies. The non-STEM group includes Social Sciences and Humanities. Compared to panel (a) of Fig. 1, here we exclude Theology faculties, which were suppressed by the Law No. 1251, January 26, 1873. Panel (b) reports the number of faculties in Italy in the Northern, Central and Southern regions. Panel (c) shows the evolution in the total number of faculties in Italy between 1870 and 2010 in each quartile of the provincial population in 2010. Panel (d) presents the number of universities in EU4 (France, Germany, the UK and Italy). This is an index equal to 100 in the first year (1870). Authors' calculations based on HIU data and WHED data.

Fig. 2 Evolution of the number of faculties in Italy from 1870-2010 and EU comparison

NUTS1 regions and within all quartiles of province population.[7] Finally, Panel (d) demonstrates that Italy's evolution of higher education supply from the mid-19th century shares a similar trajectory with that of France, Germany, and the UK. All four nations exhibit an upsurge in university numbers, particularly noticeable from the 1960s onwards. While Italy's profile is comparatively more gradual in contrast to the other three countries,[8] this

[7] Panel (c) highlights a more pronounced expansion of faculties in provinces belonging to the top quartile of population when contrasted with provinces in lower quartiles, a trend discernible even from the unification period. Subsequently, Fig. 1 elucidates these divergent trends that surfaced during the Renaissance era. Our empirical analysis thoroughly addresses and discusses the empirical concerns arising from this pre-existing trend.

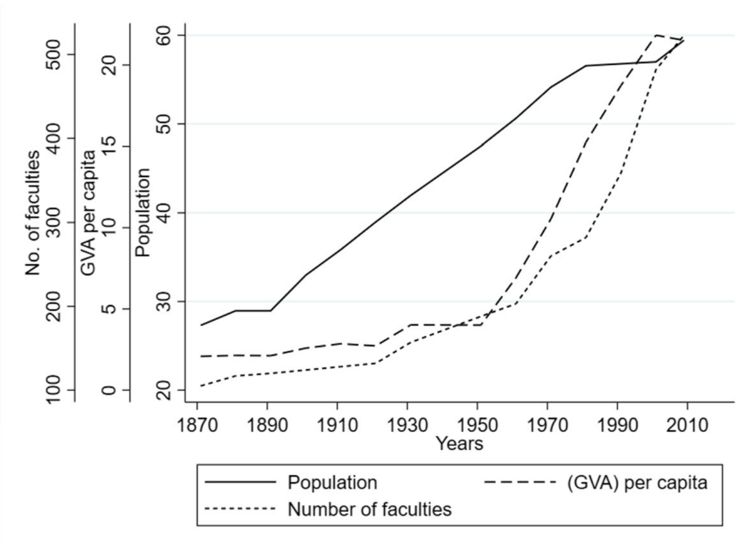
[8] Various factors could underpin this disparity, such as consistently fewer high school graduates in comparison to nations like Germany, particularly within the past 40 years.

pattern embodies the shift from an elite to a mass education system, mirroring developments that unfolded across Europe and the US during the same era (Eurydice, 2000; Smith, 2010).

Figure 3 plots total number of faculties, GVA per capita, and the overall population in Italy from 1860 to 2010. It shows three important points, which match the predictions from Unified Growth Theory (Galor, 2011). Firstly, the gradual rise in GVA per capita at the beginning of the sample period indicates that Italy had already moved away from the Malthusian era of economic stagnation, on average. Secondly, a noticeable strong link emerges between GVA per capita and the number of faculties (with a correlation coefficient of 0.97, which is statistically significant at the 1% level). This suggests a robust connection between human capital formation and economic growth, through the shift from the Post-Malthusian period to the Modern Growth Regime. This shift seems to have begun after WWI and was completed after WWII. Lastly, the influence of Malthusian forces continued to matter until the second half of the 20th century, when the growth of the population began to slow down.

2.2 A province-level provision of higher education

Our empirical analysis uses the province as geographical unit of observation. This is the natural benchmark for measuring human capital accumulation and analysing industrialisation, production, and wage trends (see, e.g., Squicciarini and Voigtländer, 2015; Ciccone and Peri 2006, Ciccone and Peri 2011; and Bratti and Leombruni, 2014). At the time of its political unification, Italy had 69 provinces. 59 became part of the sovereign state of Italy in 1861. Additional 8 provinces joined in 1866 (namely Belluno, Padova, Rovigo, Treviso, Udine, Venezia, Verona, and Vicenza), followed by Mantova in 1868. The final addition was Roma in 1870, marking the completion of Italian unification. The higher education system of the Italian Kingdom emerged as a confluence of substantially diverse education systems established by various states within the Italian territory since the Middle Ages. For this



Notes: Authors' calculations based on HIU and Census data.

Fig. 3 Higher education supply, GVA per capita and population 1870-2010 reason, the initial provision of higher education institutions across provinces was highly heterogeneous (Brizzi and Romano, 2007). As we discuss in Section 4 below, this heterogeneity constitutes the primary source of variation that we leverage for empirical identification.

Table 1 summarises information on the initial set of higher education institutions across several relevant dimensions. In Column [1], we list the 21 provinces with at least one academic institution at the time they joined Italy. In Column [2], we detail the year of foundation in the province, while in Column [3] we list the precise academic institution, whether a *studium*, a modern university, or a single faculty unit (e.g., an academy or a specialised school). This initial supply was geographically dispersed across the northern, central, and southern regions (see also Fig. A-I in the online Appendix). In general, higher education had been established during the Middle Ages, resulting in about one-third of Italian provinces hosting active faculties by the 16th century. Most Medieval institutions would take the form of a *Studium*, or a *Studium Generale*. The former would have local validity, providing higher education aiming at training local administrators in the community or region. The latter had official recognition from the Pope or the Emperor, and would offer advanced education into Literary Studies, Law, Medical Studies, Mathematics and Natural Sciences, and Theology. Modern academic institutions emerged in Cagliari, Macerata, Pesaro-Urbino, Sassari, Milano, Palermo, and Messina, established under the influence of the Enlightenment. Aligned with the Napoleonic model of higher education, these institutions emphasised scientific inquiry and secular education.

Column [4] reports the states where the first institution was founded. These are termed ‘Initial States’. The overwhelming majority of Initial States were Duchies, Counties, Principalities, Houses, Republics, or Communes, and covered the geographic expanse of just one province. In Column [5], we indicate the last state to which each province belonged prior to the unification of Italy (referred to as the ‘Pre-unitarian State’; see Fig. A-II for geographical boundaries). The comparison between a province’s Initial and the Pre-unitarian State sheds light on the evolving political dynamics occurring between the 11th and 19th centuries. Smaller Initial States would forfeit their political autonomy (for example, the House of Este in Ferrara was incorporated into the Papal States, and the Republic of Genoa was absorbed by the Kingdom of Sardinia). Some larger Initial States would gradually dissolve and cease to exist (such as the Holy Roman Empire), or undergo significant political metamorphoses due to shifts in ruling dynasties (as observed in the Kingdom of Sicily, Kingdom of Sardinia, and Papal States).[9] This transformation had far-reaching implications for the configuration of higher education at the province level. As an illustration, the Spanish Empire established new educational institutions to train local administrators in the occupied provinces (for instance, Cagliari and Catania). It would export its own academic model and enforce distinct accreditation regulations, which local institutions were obliged to adhere to in order to attain academic recognition (Roberts et al., 1996).

Column [6] provides the total count of faculties present in each province at the time of their integration into Italy, in the four academic disciplines that were retained by the Italian higher education system after the exclusion of Theology (these are Humanities, Mathemat-

[9] The Kingdom of Sicily transitioned into the Kingdom of Naples under the Angevin dynasty and later became the Kingdom of Two Sicilies under the Bourbons. The Kingdom of Sardinia initially aligned with the Council of Aragon, subsequently becoming part of the Spanish Empire in 1402. In 1720, it joined the Duchy of Savoy until Italian unification. The Papal States modified their territorial dominion over time, reaching their peak extent in the 18th century, encompassing much of central Italy as well as the Legations of Ravenna, Ferrara, and Bologna.

Table 1 Pre-unitarian higher education supply

[1] Pre-unitarian Province	[2] Foundation Year	[3] First Institution	[4] Initial State	[5] Pre-unitarian State	[6] Pre-unitarian Faculties	[7] Ranking Casati Law
Bologna	1088	Alma Mater Studiorum	Holy Roman Empire	Papal States	6	A
Cagliari	1620	Universitas Studiorum Caralitana	Kingdom of Sardinia (Spanish Empire)	Kingdom of Sardinia	4	B
Catania	1445	Studium Generale	Kingdom of Sicily (Spanish Empire)	Kingdom of Two Sicilies	4	B
Ferrara	1391	Studium Generale	House of Este	Papal States	4	Private
Genova	1481	Collegium Doctores	Republic of Genova	Kingdom of Sardinia	5	B
Macerata	1540	Studium Generale	Papal States	Papal States	3	C
Messina	1838	Università di Messina	Kingdom of Two Sicilies	Kingdom of Two Sicilies	3	B
Milano	1791	School of Veterinary Medicine	Duchy of Milano (Habsburg Empire)	Lombardy-Venetia	2	A,A
Modena	1224	Studium	Commune of Modena	Duchy of Modena and Reggio	4	B
Napoli	1224	Studium Generale	Kingdom of Sicily	Kingdom of Two Sicilies	6	A
Palermo	1806	Università di Palermo	Kingdom of Sicily (Spanish Empire)	Kingdom of Two Sicilies	4	A
Padova	1222	Studium Patavinum	Commune of Padova	Lombardy-Venetia	4	A
Perugia	1308	Studium Generale	Commune of Perugia	Papal States	2	Private
Pisa	1343	Studium Generale	Republic of Pisa	Gran Duchy of Tuscany	6	A
Parma	962	Studium	County of Parma	Duchy of Parma and Piacenza	5	B
Pesaro Urbino	1506	Collegium Doctores	Papal States	Papal States	2	Private
Pavia	1361	Studium Generale	House of Visconti	Kingdom of Sardinia	5	A
Roma	1303	Studium Urbis	Papal States	Papal States	5	A
Siena	1240	Studium	Republic of Siena	Gran Duchy of Tuscany	2	B

Table 1 (continued)

[1] Pre-unitarian Province	[2] Foundation Year	[3] First Institution	[4] Initial State	[5] Pre-unitarian State	[6] Pre-unitarian Faculties	[7] Ranking Casati Law
Sassari	1765	Università di Sassari	Kingdom of Sardinia	Kingdom of Sardinia	3	C
Torino	1404	Studium Generale	Principality of Savoy-Acaia	Kingdom of Sardinia	8	A,B

Notes: Column [1] reports the Pre-unitarian province.Column [2] reports the year of the first faculty in the province. Column[2] reports the first Higher education institution present in the territory. Columns [4] and [5] report the Initial and Pre-unitarian State to which the province belonged. In Column [6] we report the number of faculties observed in the province at the moment of entry into the Italian Kingdom. Column [7] reports the quality assessment of higher education institution(s) in the province according to the Casati Law. This evaluation refers to public institutions only. In Torino, the Casati Law assigns an A score to the University of Torino and a B score to the Polytechnic. In Milano, it assigns an A score to both the Scientific Academy and the Polytechnic. In Milano, the first higher education institution was the autonomous School of Veterinary Studies, while the University of Milan was established in 1924 only. The source is HIU Data

ics and Natural Sciences, Law, and Medicine; refer to de Ridder-Symoens, 1992). Province-level provision varies from a minimum of two faculties (in Milano, Perugia, Pesaro-Urbino) to a maximum of eight faculties (in Torino). Finally, Column [7] of Table 1 provides the initial evaluation of these higher education institutions based on the Casati Law (Law 3725/1859). This legislation established the criteria for accrediting existing institutions within the new Italian university system. In nine provinces, higher education supply was awarded an 'A' ranking, signifying compliance with the highest quality standards set by the Italian government for both teaching and research. In multiple provinces, pre-unitarian supply received a 'B' ranking. While not meeting the highest standards, these faculties served as valuable second-tier regional institutions. Faculties associated with the Universities of Sassari and Macerata fell below the minimum requirements and were classified as 'C'. The Casati Law also recognised private higher education provision in the provinces of Ferrara, Pesaro Urbino, and Perugia, where local authorities previously directly supported universities (Brizzi and Romano, 2007).[10]

The discussion above suggests that preunitarian universities and academic institutions developed independently across multiple regions and sovereign states on the Italian territory, rather than being centrally planned or governed by a single authority. In modern times, these polycentric origins resulted in a highly decentralised system, resistant to centralisation efforts throughout the 20th century. Historically, the province has been the key territorial unit for higher education and the provision of public services. Petracchi (1962) defines provinces as "large associations of municipalities devoted to the protection of each municipality's rights and to the management of their collective moral and material interests". Among these interests, the local provision of higher education was particularly significant. Although provinces lacked the political authority to influence national reforms—due

[10] Initially, the Casati Law was applied to the new territories of the Kingdom of Sardinia. The subsequent Matteucci Law in 1862 extended its application to all territories that progressively became part of the Italian Kingdom (refer to Table A-I in the online Appendix). De la Croix et al. (2024) demonstrate that the Casati ranking strongly aligns with their index of university prominence, which employs a distinct metric based on estimations of human capital for the top 5 academic scholars who spent time in that university during the period 1740-1800.

to constitutional limitations on their legislative competencies and their fragmented representation across Italy—they played a crucial role in shaping the geographical distribution of faculties.[11] Province-level *consortia* initiated proposals for new faculties or universities, which were subsequently reviewed by the central government (Brizzi and Romano, 2007). Local politicians coordinated academic resources and funding from banks, chambers of commerce, and local investors, while the central government primarily functioned as a regulator, setting administrative requirements but remaining neutral regarding the spatial allocation of faculties.

Between 1870 and 2010, the expansion of Italian higher education followed this decentralised model. New faculties were established at the provincial level, responsible for designing courses and awarding degrees. In many cases, the establishment of a faculty preceded the founding of a university, highlighting the pivotal role faculties played as local providers of higher education. For example, the University of Milan was established in 1923; however, the Faculty of Veterinary Medicine traces its origins back to 1791, when it was founded as the Scuola Veterinaria Minore.[12] In 1863, the Polytechnic of Milano was founded as an advanced technical institution known as the Regio Istituto Tecnico Superiore. The creation of faculties often precedes the establishment of universities in modern times as well. For instance, the University of Piemonte Orientale was founded in 1998, evolving from seven faculties that had been established five years earlier, in 1993, as branches of the University of Turin within the provinces of Novara, Vercelli, and Alessandria (online Appendix A provides further details and examples). Figure 4 compares the total count of faculties retained by the Italian higher education system after the completion of Italian unification to the number of faculties in 2010. Panels (a) and (b) visually illustrate that the surge in faculty numbers encompasses the entire expanse of the Italian territory. Even in 1870, faculty distribution was predominantly centred in the capitals of pre-unitarian states. By 2010, two-thirds of the provinces hosted at least one faculty. Provinces devoid of faculties are the most recent additions, established in the post-WWII era or more recently. A handful of provinces featured 15 faculties or more (including Bari, Bologna, Milano, Napoli, and Roma).[13]

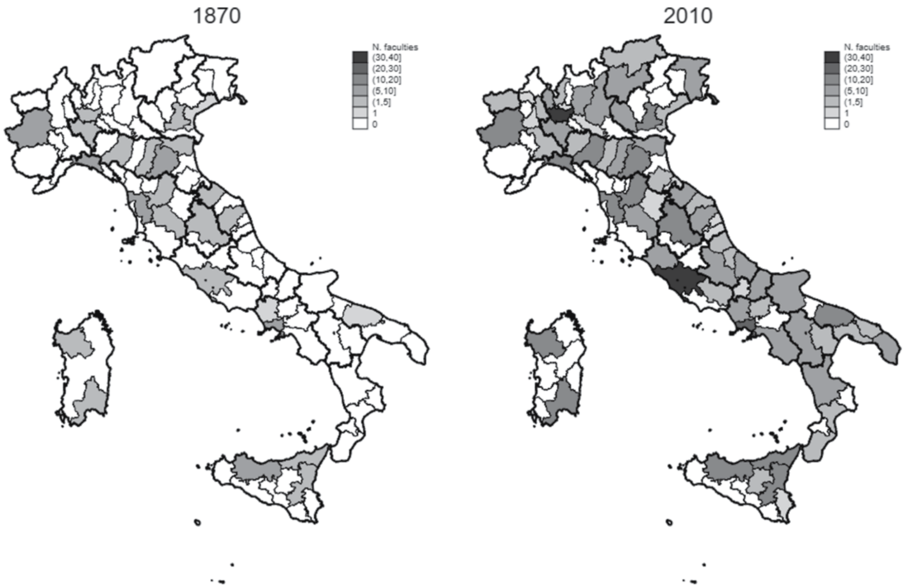
The preceding discussion highlights the importance of considering an empirical framework that allows to analyse the economic effects of a decentralised provision of higher education at the province level. We opt for a spatial economics model where each province independently determines the local offering of higher education based on local student demand.[14] This setup gives rise to neighbourhood effects within higher education supply, as multiple neighbouring provinces respond to the same local demand. This empirical setting is well suited to the Italian case, as the Italian public higher education system lacks price-based competition due to substantially uniform and low tuition fees, which prevents

[11] The Republican Constitutional Law omits specific tasks and legislative powers for provinces (see Article 114 and Fabrizzi 2008a, b, c). Their dispersion across the country also means each represents only a small fraction of the total population.

[12] It aimed to educate a specialised group of farriers capable of performing minor surgical interventions and treating common ailments in horses, cattle, and sheep.

[13] For further insights into temporal variations across different regions, consult Fig. A-III in the online Appendix.

[14] This approach is reminiscent of tax setting by autonomous jurisdictions (see, for example, Parchet, 2019), which is fully symmetric to traditional spatial competition models in the vein of Hotelling. In these models, it is the supplier that selects its position over space instead.



Notes: Authors' calculation based on HIU data using a map of Italian provinces and regions (whose borders are in bold). The borders of 1870 refer to the 69 provinces existing in 1870, those of 2010 to the 110 provinces of 2010. For visual convenience, we use the 2010 country borders also for the 1870 figure, which include territories that in 1870 were not already part of Italy (i.e., the provinces of Trento and Bolzano). The count does not consider faculties of Theology, as these were not retained by the Italian higher education system.

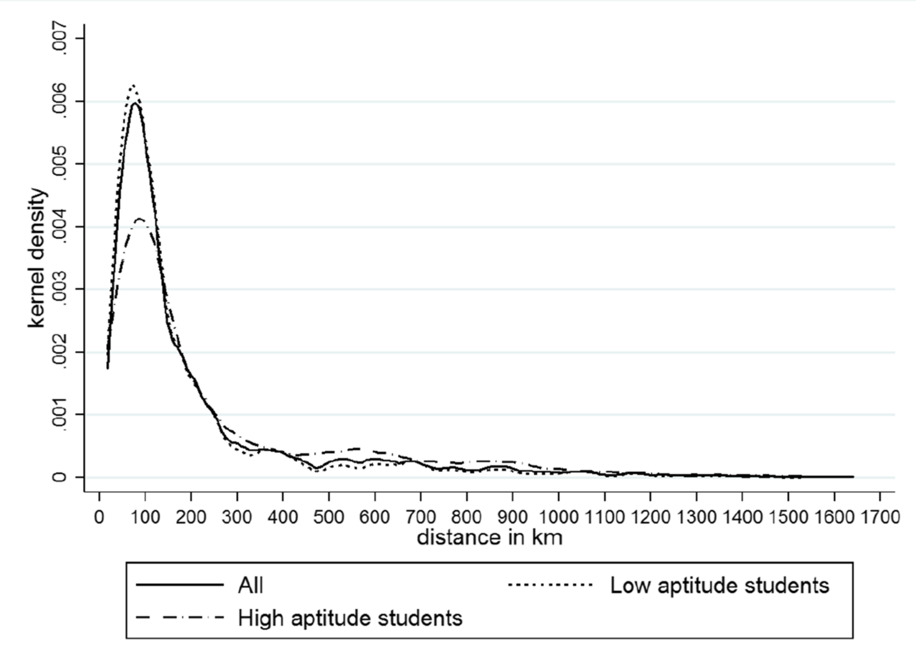
Fig. 4 Diffusion of faculties in Italian provinces

provinces from enticing students by reducing local university fees.[15] Consequently, neighbourhood effects may largely be associated with the opening of new faculties: an additional faculty in a province could attract more students, thereby potentially reducing enrolment in faculties of neighbouring provinces, as admissions are generally free.[16]

The conceptual framework just described is based on two important assumptions. The first assumption concerns student mobility. Neighbourhood effects become noticeable when there is a certain degree of individual mobility, albeit confined primarily within the local neighbourhood due to high commuting costs outside that area. This *local mobility assumption* suggests that each province can cater to a limited portion of the national demand. As a result, province-level supply aligns with this local demand and contributes to the local supply of skills rather than the national one. Figure 5 below indicates that a significant portion of Italian high school graduates opt for higher education institutions within a distance

[15] Throughout our study period, universities do not face a binding budget constraint: public universities are financed by the central government through lump-sum transfers, which increase with both student admissions and past university expenditures.

[16] These neighbourhood effects arise naturally due to historical competitive forces. However, we cannot dismiss the possibility of other complementary mechanisms. For instance, there might be yardstick competition among provinces aiming to expand their higher education supply to increase public transfers from the central government. Political coordination could also play a role, such as if two provinces with similar demand agree on the choice of location.



Notes: Authors' calculations based on data from the Italian survey of high school graduates, 2007 (ISTAT). We define students' aptitude using their final grades at lower secondary schools, which are set on a four-point scale. High aptitude students are those with the highest grades (4), while low aptitude students are those with the lowest grades (1 or 2).

Fig. 5 Students' mobility in faculty choice

of 150 km.[17] In Section 6 below, we validate that neighbourhood effects are concentrated within a comparable spatial reach (e.g., in terms of travel distance). We also show that these effects emerge in provinces located in geographically smooth areas with historically low out-migration rates, where the development of human capital becomes a critical aspect of the local labor supply. We interpret this as empirical evidence that supports the local mobility assumption. The second assumption is that there exists a certain degree of differentiation, such as in teaching content, among faculties in neighbouring provinces. If student preferences exhibit heterogeneity and are influenced by both faculty proximity and study content, selecting the nearest faculty might not be the optimal choice. Provinces may take advantage of this dimension of heterogeneity and differentiate their supply accordingly. Our mechanism analysis in Section 6 indeed indicates that neighbourhood effects are particularly strong within the same field of study.

When these two assumptions are met, the establishment of a new faculty can lead to welfare benefits for provinces. A province will introduce a new faculty if it maximises local welfare. When doing so, it will have to factor in the strategic responses of neighbouring provinces, which might foresee potential welfare reductions due to displacement of their own supply. In the empirical analysis, we extensively investigate the economic benefits

[17] Mobility is less common among students with lower study aptitudes, a trend observed in the US as well (refer to Hoxby, 2016).

from higher education supply, distinguishing between local returns and economic spillovers. In the welfare assessment detailed in Section 7, we will quantify the economic returns of higher education supply, net of neighbourhood effects at the province level.

3 Data and descriptive statistics: the history of italian universities (HIU) database

We have developed an original dataset on the History of Italian Universities (HIU), which compiles information on academic institutions across Italy from the year 1000 to 2010, when a major reform reorganised the institutional governance of the Italian higher education system. This dataset focuses exclusively on Italy, the Western country with the longest-standing tradition of higher education. Compared to existing datasets by Valero and Van Reenen (2019) and Dittmar and Meisenzahl (2022), HIU data provides more precise measures of higher education supply, distinguishing across several teaching fields. This finer level of disaggregation enables us to track the evolution of higher education provision within the two broad areas of STEM, and non-STEM faculties. We describe below with some detail the HIU Registry. We also explain how we use this registry to construct a province-level panel dataset of higher education institutions supply covering the period 1870-2010 of Italian history, as well as a full set of initial conditions for the Italian university system which we use for identification purposes.

HIU Registry The registry contains detailed and complete information on institutions providing higher education in Italy, disaggregated at the level of the faculty teaching unit, responsible for all Bachelor level (BSc) courses within a given disciplinary area. The dataset covers all years starting from around year 1000 (The first *Studia* in our data dates back to 962 in Parma, and 1088 in Bologna) up to 2010, when the Law 30/12/2010, n. 240 (the so-called ‘Gelmini reform’, after the name of the Public Education Minister in office at that time), eliminated faculty institutions from the governance of public universities.

The primary datasource for the construction of the register data was Brizzi and Romano (2007), which contains detailed information on the history of italian universities starting from their foundation onwards. We integrated this information using several sources on the history of specific universities (e.g., Fois, 1991 and Mattone, 2010 on University of Sassari; Pepe, 2007 on University of Ferrara; Mele, 2006 on the Politechnic of Torino.) as well as works with a specific focus on the evolution of specific faculties in Italy (e.g., Bidolli and Soldani, 2001 on Agricultural Studies; Silvestri, 2006 on engineering), or in specific universities (e.g., Cabassi and Liuzzo, 2001 on veterinary medicine studies in the University of Parma).

We also heavily relied on *Gazzetta Ufficiale della Repubblica Italiana*, a weekly publication that collects every public act taken by the government since 1861. This is the only official source of knowledge regarding the laws approved (and in force) in Italy. This is also the most authoritative tool for the dissemination, information and officialisation of legislative texts, public and private acts that must come to the knowledge of the whole community. We double-checked all information against that provided by open-source archives, i.e., Wikipedia, universities’ and faculties’ websites. Since faculties no longer exist after 2010, their websites are not readily available on the web today. We retrieved them using the Wayback Machine (*https://web.archive.org/*), a digital archive of the worldwide web created by the internet archive. We directly contacted the administrative representatives of faculties for which we could not find information from external sources. Finally, we validated the data against the current university list provided by the Italian Ministry for University, Education and Research (MIUR). In this way, we also verified that over the decades, there are very few faculty closures: besides Theology faculties that have been closed at the onset of Italian unification, almost all faculties that were created in Italian history are still present in 2010.[18]The final HIU registry includes 554 faculties and (in) 72 universities delivering a diploma equivalent to a BSc degree, which are registered on the Italian territory at some point between 1000 and 2010. The registry includes the name of each university, each faculty, its current address and 15 faculty-field identifiers. Data include the year when the faculty appears as a legally recognised provider of higher education, the type of governance (private or public), and quality assessments according to the Casati ranking and its following updates (see online Appendix B1 for details).

HIU Province-level panel dataset From the registry, we create a panel dataset that tracks the count of university faculties at the province level during Italian history. We exclude from the count all faculties which were officially suppressed after Italian unification (i.e. 18 faculties of Theology in 1873 and the two scientific faculties in Urbino in 2006), and 11 faculties of sport and physical education. We also exclude the years 1861-1869, as during this period the Italian unification process was still ongoing. We are left with 523 faculties (within 72 universities), which are active on the Italian territory between 1870 and 2010.

Thus, the province-level panel dataset includes yearly observations for 110 provinces spanning from 1870 to 2010. The primary variable of interest is the faculty count in province i in year t, presented both as a total and categorised by macro-area of science (STEM and non-STEM; cfr. the note of Fig. 1 for the definition of the two groups). Additionally, we include details about the total count of universities, private universities, and universities categorised as A-, B-, or C-level in province i during time t. The panel is unbalanced as there were only 69 provinces, which were inherited by pre-unitarian states at the moment of the Italian unification. During Italian history 41 new provinces appeared, as a result of the latest territory acquisitions, and autonomy acquired by some circondaries.[19] In the main analysis, we adopt a definition of pre-unitarian provinces, which privileges their role as administrative units, and allows their territory and borders to change overtime, alongside the creation of new provinces during Italian history. In online Appendix D, we confirm our

[18] The only exceptions are two scientific faculties (Environmental Sciences and Mathematical Sciences) at University of Urbino, which were closed in 2000. Moreover, four faculties of Engineering from the Polytechnic of Milano were closed and reopened with a different denomination and location in 2000 (see online Appendix B1 for details).

[19] The majority of new provinces were established during the late 1920s (Bolzano, Brindisi, Caserta, Enna, Frosinone, Gorizia, Matera, Nuoro, Pescara, Pistoia, Ragusa, Rieti, Savona, La Spezia, Terni, Trento, Trieste, Varese, Vercelli, and Viterbo), the early 1990s (Biella, Crotone, Lecco, Lodi, Prato, Rimini, Verbano Usio Cassola, and Vibo Valentia), and the early 2000s (Ogliastra, Olbia-Tempio, Monza, Medio-Campidano, Fermo, Carbonia-Iglesias, and Barletta-Andria-Trani). Two provinces in the 1930s (Asti, Latina,) and in the 1970s (Isernia and Oristano), one province in the late 1960s (Pordenone).

empirical analysis with different definitions of pre-unitarian provinces that consider alternative assumptions regarding their territory and borders in 1870 (cfr. Table A-VII).

The province-level panel data is the basis for the two primary samples used in our estimation processes. We create our first sample by merging the province-level panel data with information on productivity indicators like GVA per capita, and GVA per worker. These measures are adjusted for inflation and expressed in constant 2010 prices. We also incorporate indicators related to participation and employment shares within sectors such as industry, agriculture, and services. These data points are sourced from Italian censuses conducted at ten-years intervals, encompassing the period from 1870 to 2010.[20] After successfully merging the HIU province-level panel data with census data, our resulting dataset comprises a total of 1275 observations covering 110 provinces. These observations are recorded every ten years, spanning the timeline from 1870 to 2010.

For our second sample, we match each province *i* with the higher education supply of all its neighbouring provinces *j*. To establish these neighbouring relationships, we employ a matrix that documents contiguity connections among the 110 provinces. We also compute alternate matrices grounded in linear distance, travel distance, and travel time. Comprehensive details are available in the subsequent sections and online Appendix B. Allowing pairwise relationships between provinces *i* and *j* increases the sample size by approximately four times (as each province *i* has roughly four neighbours, on average). Moreover, to align with the matched HIU-census sample, we adhere to the same ten-year structure for this paired dataset. As a result, our second working sample encompasses 5536 province pairs denoted as '*ij*', maintaining the contiguity relationship over the period spanning 1870 to 2010.

Table 2 presents summary statistics for the main variables in both samples. Panel A describes the local higher education supply from the matched HIU-census dataset at the province level. On average, each province has a supply of over 2.6 faculties, mostly in STEM. Each province has approximately 0.5 universities, of which 12% are private, and around 77% are classified as of level A. Panel B turns to pairwise relationships, and reports the higher education supply of each individual neighbour *j* of province *i*. Our main objective is to model the local neighbourhood effects between the 110 provinces in our sample with the utmost precision. To accomplish this, we carefully consider the changes in provincial borders that occurred during Italian history, particularly due to the gradual creation of 41 new provinces. This approach allows us to capture evolving neighbourhood relationships (i.e., the sample of *j*s for each *i*) over the 1870-2010 period.[21] It is important to notice that new provinces on the Italian territory are born without an own supply of higher education.[22] All in all, a representative neighbour, on average, has approximately 2.95 faculties; other figures are similar to those displayed in Panel A. Finally, Panel C reports province-level

[20] An exception is the 9th wave of the Italian census, which, due to WWII, was carried out in 1936 instead of 1941. This interval deviates from the typical ten-year pattern, occurring five years after the prior wave and fifteen years before the subsequent one. Refer to Section B4 in the online Appendix for further information.

[21] Let us consider the province of Treviso as an illustrative example. In 1866, when it became part of the Kingdom of Italy, we identify the contiguous provinces of Udine, Belluno, Padova, Venezia, and Vicenza as neighbours of Treviso. However, starting from 1968, the new province of Pordenone is established on the former territory of the province of Udine, situated at the border with Treviso. As a result, from 1968 onwards, Pordenone replaces Udine in the set of neighbours associated with Treviso

[22] The only exception is the province of Monza-Brianza, that is established in 2004 with one faculty of Medicine that used to belong to the province of Milano before.

Table 2 Summary Statistics

[1] Variable	[2] Mean	[3] Std. Dev.	[4] Min.	[5] Max.	[6] Obs.
Panel A: Higher education supply in provincei					
no. of faculties in *i*	2.619	4.319	0	38	1275
no. of STEM faculties in *i*	1.387	2.16	0	12	1275
no. of non-STEM faculties in *i*	1.232	2.375	0	26	1275
no. of universities in *i*	0.472	0.789	0	7	1275
no. of private universities in *i*	0.069	0.336	0	4	1275
no. of A-level universities in *i*	0.366	0.602	0	3	1275
no. of faculties in $-i$ (neighbourhood of *i*)	12.718	10.475	0	68	1275
no. of universities in $-i$	2.213	1.687	0	11	1275
Panel B: Higher education supply of neighbouring provincej					
no. of faculties in *j*	2.952	4.61	0	38	5536
no. of STEM faculties in *j*	1.567	2.268	0	12	5536
no. of non-STEM faculties in *j*	1.385	2.576	0	26	5536
no. of universities in *j*	0.513	0.837	0	7	5536
no. of private universities in *j*	0.082	0.386	0	4	5536
no. of A-level universities in *j*	0.397	0.617	0	3	5536
Panel C: Economic and demographic characteristics of provincei					
Gross Value Added per capita, real terms (in 000s)	9.095	8.525	1.402	91.100	1275
Gross Value Added per worker, real terms (in 000s)	21.589	19.795	2.918	206.397	1275
Population (in 000,000s)	0.517	0.469	0.057	4.018	1275
Population density (100 inhabitants per km^2)	2.171	2.882	0.244	25.949	1275
Participation rate	44.491	7.586	25.478	98.571	1275
% of employed in industry	28.686	11.271	6.562	72.864	1275
% of employed in agriculture	38.752	25.926	0.646	85.686	1275
% of employed in services	32.562	19.74	6.643	84.243	1275
Industry-to-agriculture employment ratio	2.89	5.595	0.077	57.81	1275
Service-to-agriculture employment ratio	4.363	9.618	0.078	121.492	1275

Notes: Authors' calculation based on HIU data, ISTAT (Italian National Institute of Statistics), and Istituto Tagliacarne. Panels A and C consist of province-level data extracted from 15 waves of Italian Censuses conducted between 1870 and 2010. Panel B includes province pair information for the same time span. The Italian Census is typically conducted once every ten years, except for the 1936 census, which took place 5 years after the previous one held in 1931 and 15 years before the subsequent one in 1951. Gross Value Added (per capita or per worker) in real terms is expressed in euros at 2010 constant prices

historical information on local economy and demography. On average, in our sample, each province presents a yearly real GVA per capita equal to 9000 euros (constant 2010 prices) and a population of 517 thousand individuals. Approximately 45% of the population is active in the labor market, with 28% working in the industry sector, 38% in the agricultural sector, and 32% in the service sector.[23]

The Initial Conditions of Italian Supply of Higher Education As we discuss in the next section, we will exploit cross-provincial variation in the distribution of faculties observed at

[23] It should be noted that in historical census data, the definition of active population is very blurred as unemployed individuals were not counted until 1961.

different moments before the Italian unification to address concerns regarding endogenous sorting of faculties across Italian provinces during the post-unification period 1870-2010. The 1000-1861 coverage of HIU data allows us to observe higher education supply at different moments of preunitarian history of Italian territories. Table 8 in Appendix B describes the higher education supply at the province level moving from the end of Middle Ages (i.e. in 1400) throughout the Reinassance, the Age of Enlightenment, the Napoleonic Period, up to the Restoration and Italian Unification (1861). These faculties provide the initial conditions of the higher education supply in our working sample 1870-2010. On average, in province *i* in 1400 we observe 0.4 faculties, mostly non-STEM. Supply of faculties more than doubles during the following four centuries, particularly in STEM fields and during the period 1800-1861. In 1861 we observe on average 1.07 faculties in province *i*, of which 0.68 are STEM and 0.39 are non-STEM.

4 Empirical strategy

In this section, we outline our empirical approach. We first describe the model used to estimate neighbourhood effects on local higher education supply. Then, we estimate province-level returns for both local and neighbouring higher education supply in terms of real GVA per capita. These coefficients form the basis for the welfare analysis outlined in Section 7.

Neighbourhood Effects of Higher Education Supply We estimate neighbourhood effects by specifying the higher education supply F in the *i*-th Italian province at time *t* as a function of the corresponding supply in their neighbouring provinces *j*s. Our baseline specification is a linear spatial competition model based on geographical contiguity, where neighbours are all provinces *j* that share a border with *i*. We also use alternative models based on linear distance, travel distance, and travel time.[24] All models feature heterogeneous neighbourhood relationships, which are immune to Manski (1993)'s reflection issues.[25]

We estimate a pairwise model, which is well suited to account for the omitted determinants of province *i*'s supply (e.g., related to geographical factors) that are spatially correlated with those of its *j* neighbours (see Parchet, 2019):

$$F_{ijt} = a + bF_{jt-10} + c_{ij} + d_{r(i)jt} + X_{ijt-10}f + e_{ijt}, \tag{1}$$

where F_{ijt} and F_{jt} are the higher education supply of province *i*, and each of its neighbouring provinces *j*, respectively.

Equation 1 can be estimated for all contiguous *ij* pairs in our dataset. Each pair of provinces is included twice in the equation, once on the left side and once on the right side. As noted earlier, we account for shifts in provincial borders that occurred throughout Italian

[24] This is done in Table 5, and in Table A-XII in online Appendix D.

[25] The reflection problem arises when interactions occur in a fixed 'reference group'. Neighbourhood relationships (e.g., based on distance, contiguity, and travel time) imply that each Italian province interacts in a different way with all other provinces, so reference groups differ for different provinces. By definition, these groups only partially overlap because the sets of neighbours of two provinces do not perfectly coincide. See De Giorgi et al. (2010) for a detailed discussion in the context of social interactions.

history. This approach acknowledges that neighbourhood relationships (i.e., the set of *j*s for each *i*) evolve over the period from 1870 to 2010.[26] In the equation, c_{ij} represents the fixed effect of the pair, capturing time-invariant omitted factors for each *ij* couple. $d_{r(i)jt}$ denotes region-by-year fixed effects (where *r*(*i*) refers to the region to which province *i* belongs). X_{ijt} is a vector of time-varying covariates from province *i* and its neighbour *j*. In the baseline specification, this includes the total number of universities in each province and for each neighbour. In this way we isolate the neighbourhood relationships that pertain to the faculty-level provision (e.g., teaching units, curricula etc) from any 'higher scale' aggregate interactions that concerns university openings. In sensitivity analyses, we also consider additional potentially relevant covariates such as population growth, the share of the active population in agriculture, controls for higher education quality in the province, and geographical accessibility.[27]

The heterogeneity that is left in Eq. 1 after the inclusion of the province pair fixed effects and region-by-year fixed effects is the variability over time across province pairs within the same region. Our main coefficient of interest is *b*, which captures the average effect of each neighbour *j* supply on the local supply of province *i*. The inclusion of the number of universities in vector X_{ijt} ensures that the estimated effects are driven solely by changes in the number of faculties.[28] Identification of this coefficient is achieved through a comparison of different time-varying patterns in the number of faculties across provinces. Finally, as discussed in Section 2, several actors are involved in the decision to open a new faculty, which implies a delay in the effect of neighbours on local supply. For this reason Eq. 1 models a ten-year lagged effect.

Productivity effects of higher education supply We estimate the following model:

$$Y_{it} = \phi_1 F_{it-10} + \phi_2 F_{-it-10} + X_{it-10}\phi_3 + \varphi_i + \xi_{r(i),t} + \mu_{it}. \tag{2}$$

where Y_{it} is GVA per capita in province *i* at time *t*, in real terms.[29] F_{it-10} is the total number of faculties in province *i* at time $t-10$. $F_{-it-10} = \sum_j F_{ijt-10}$ is the total number of faculties in the neighbourhood of province *i* at time $t-10$. X_{it-10} is the same vector of province-level controls as in Eq. 1 above, including the total number of universities in prov-

[26] In online Appendix D, we explore alternative definitions of province borders that keep neighbourhood relationships constant over the time span (see Table A-VII).

[27] This information is presented in Table 11 in Appendix B. To evaluate higher education quality, we include variables like the count of private and A-level universities, along with interactions involving the notability index developed by De la Croix et al. (2024) with university counts in both *i* and *j*. However, we avoid including all these province-level controls in the baseline estimates due to the potential problem of 'bad controls' over the extensive 150-year period. Including bad controls can introduce bias to the estimates of the parameters of interest, even when the related regressors are fully exogenous (Angrist and Pischke, 2009).

[28] Notice that a separate empirical identification of the effect of universities from the effect of faculties requires that each university does not consist of a relatively constant number of faculties overtime. This is exactly the case of Italy, where the variability over time in the number of faculties within universities is substantial. Universities are less likely to react and quickly adjust to changes in the local environment compared to faculties.

[29] This definition is broadly equivalent to the Gross Domestic Product (GDP, hereafter) which is indeed the sum of the GVA created through the production of goods and services in the individual sectors of the economy. Conceptually, the main differences between GVA and GDP based definition of output regards financial intermediation services and product taxes, which are not counted into GVA (while they are in GDP).

ince i and its $-i$ neighbourhood.[30] We include also province FEs, φ_i, and region-by-year FEs, $\xi_{r(i),t}$. Finally, μ_{it} is the error term.

Our main parameters of interest from Eq. 2 are ϕ_1 and ϕ_2. These measure the returns to province *i* of faculties set up locally, and in the neighbourhood, respectively. As in Eq. 1, these estimates are conditional on the supply of universities. Our aim is to show that a non negligible part of productivity effects of higher education happens at a more disaggregated level (the supply of faculties), conditional on the supply of universities. This is different from Valero and Van Reenen (2019), whose analysis is at the university level instead.

For both Eqs. 1 and 2 we exploit time variation consistent with the 10 years' timeline imposed by census data collection. This implies that over the 150-year sample period, we exploit an unbalanced panel with $T = 15$ covering the period of 1870-2010. We will use the predicted returns ϕ_1 and ϕ_2, net of displacement cost b, to quantify the total contribution of higher education supply to the economic welfare of the local province, distinguishing between effects of local supply, and net externalities coming from the neighbourhood. We will return to these calculations in Section 7 below.

Estimation issues and identification The two models just discussed present endogeneity concerns. The most important one is omitted variables. In Eq. 1, unobservable time-varying determinants of one province's number of faculties (e.g., local economic conditions, shocks and demand for higher education) are likely spatially correlated such that $cov(e_{ijt}, F_{ijt}|c_{ij}, d_{r(i)jt}, X_{ijt}) \neq 0$. Even if these factors were region specific, they would not be captured by region-by-year fixed effects as long as not all neighbours belong to the same region as province *i*. Similarly, omitted factors in Eq. 2 may motivate both an increase in GVA per capita, and the opening of faculty such that $cov(\mu_{it}, F_{it-10}|\varphi_i, \xi_{r(i)t}, X_{it}) \neq 0$ and $cov(\mu_{it}, F_{-it-10}|\varphi_i, \xi_{r(i)t}, X_{it}) \neq 0$. Despite these are mitigated by the use of ten years' lagged effects, there are also concerns of reverse causality. In Eq. 1, the number of faculties of a neighbouring set of provinces itself depends on the number of faculties of province *i*. In Eq. 2 richer and/or more productive provinces may express higher (or lower) demand for higher education. Thus, the opening of a faculty may result from this demand.

To address endogeneity concerns, we implement a Bartik instrumental variables approach that exploits cross-provincial variation in the distribution of faculties observed at some initial time t_0, which denotes some precise moment before the Italian unification (see more below on the choice of t_0). These faculties provide the initial conditions of the Italian higher education supply. Call IC_{it_0} the total number of faculties operating in province *i* at time t_0, such that we can compute the province *i*'s share of the total supply in the Kingdom of Italy at that time, i.e.:

$$sh_{it_0} = \frac{IC_{it_0}}{\sum_i IC_{it_0}}. \tag{3}$$

We construct respective Bartik indicators B_{it} and B_{jt} as the counterfactual number of faculties in provinces *i* and *j* at time *t*, as we allocate to each province the number of faculties

[30] We perform the same set of robustness checks as for Eq. 1 and control for population growth, the share of the active population in agriculture, controls for higher education quality in the province, and geographical accessibility (see Table 11 in Appendix B).

observed in Italy at time t, according to its respective initial share at time t_0. We also construct a synthetic Bartik indicator B_{-it} for the neighbourhood of i:

$$\begin{aligned}
B_{it} &= sh_{it_0} * F_t, \quad where \quad F_t = \sum_i F_{it}; \\
B_{jt} &= sh_{jt_0} * F_{-it}, \quad where \quad F_{-it} = \sum_{j \neq i} F_{jt}; \\
B_{-it} &= \sum_{j \neq i} B_{jt} = \sum_{j \neq i} sh_{jt_0} F_{-it}.
\end{aligned} \tag{4}$$

By looking at Eq. 4, our IV approach consists of using ten-year lagged values of B_{jt} as an instrument for F_{jt-10} in Eq. 1; tenyear lagged values of B_{it} as an instrument for F_{it-10} in Eq. 2; ten-year lagged values of B_{-it} as an instrument for F_{-it-10} in Eq. 2. This approach features a research design whereby initial conditions measure the exposure of pre-unitarian provinces to the faculty growth process during Italian history (cfr. Fig. 2). This is an exogenous 'shift' to the higher education supply of each individual province. As discussed in Section 2.1 above, the faculty creation process over Italian history is driven by general higher education reforms, which are independent of previous provincial conditions and likely uncorrelated with past shocks. Following Goldsmith-Pinkham et al. (2020), we argue that the identification of Bartik instruments Eq. 4 comes from the cross-province variation of IC_{it_0} and IC_{jt_0}, through respective initial shares. Higher education institutions inherited from the past had a persistent impact on the process of faculty creation at the province level. In the main analysis, we measure initial conditions at $t_0 = 1860$, and $t_0 = 1700$. The former timing captures the annexion of pre-unitarian provinces to the Kingdom of Italy,[31] while the latter predates the Age of Enlightment and the industrial revolution in Italy, being predetermined to the sharp acceleration in the supply of faculties of the second half of the 18th century (cfr. Panel (a) of Fig. 1). In the robustness section, we also consider earlier timings, back to 1400, i.e. the end of a century of declining population (cfr. Panel (b) of Fig. 1).

As our IV approach builds upon the exogeneity of initial conditions, in Table 3, we present correlations between the initial distribution of the number of faculties and observable characteristics of province i before the start of our sample period.[32] In Columns [1] - [8], we check whether initial conditions correlate with time invariant characteristics at the onset of Italian unification. We consider linear distances (in km) from Bologna (i.e., the city with the oldest active university), Roma (the capital of the Italian Kingdom during our sample period), the respective capital of each pre-unitarian state (cfr. Fig. A-II in online Appendix A), GVA per capita, the share of workers in the agriculture and industry sectors, urbanisation rate, and total population density. In Columns [9] and [10], we consider population growth between 800AD and 1500, and between 1500 and time t_0, when we measure initial conditions. These correlations are meant to capture any long-term historical pre-trends during different epochs. We subject these correlations to an alternative clustering of standard errors. We report standard errors clustered at the NUTS2 level between (...). Standard errors

[31] We adopt the notation $t_0 = 1860$ for simplicity but this is not precise. Belluno, Padova, Rovigo, Treviso, Udine, Verona, Vicenza, Venezia, Mantova, and Roma became part of Italy between 1866 and 1870.

[32] Since our baseline specification uses the 69 pre-unitarian provinces to maintain a balanced sample throughout the period of analysis, we focus solely on these 69 provinces in the pre-trends of Table 3.

are clustered at the level of the pre-unitarian state the province belonged to between [...] and at the NUTS1 level between {...}.

With very few statistically significant correlations, evidence in Table 3 does not suggest historical factors that pose a systematic threat to our identifying assumptions. Panel A presents the battery of correlations of the initial conditions IC_js and IC_is, with the aforementioned set of preexisting characteristics of province i as we set $t_0 = 1860$. This first battery of tests suggests that IC_js are uncorrelated with linear distances and time invariant economic characteristics of province i (cfr. Columns [1] - [8]). A negative correlation emerges with the initial level of population density, which however becomes statistically not significant as we cluster standard errors at the NUTS1 level. Tests of pre-trends in Columns [9] and [10] evidence a negative spillover from population growth during the 1500 - 1860 period. The battery of correlations for the IC_is on time invariant characteristics suggests a positive significant correlation of initial conditions with the share of industry, and local population density (cfr. Columns [6] and [8]). Further to that, an inspection of changes into the pre-sample period reveals a statistically significant pre-trend relating to population growth between 1500 - 1860 (cfr. Column [11]), which becomes less evident when we split the sample by geographic area (see Table A-IV in the online Appendix). This last correlation suggests some potential co-growth between population, demand for human capital, and faculty creation starting after the end of Middle-Ages in line with predictions by the Unified Growth Theory (Galor, 2011). In Panel B, we replicate the battery of tests after excluding provinces in the top decile of the population distribution back in 1500. All correlations become non-significant, with the only exception of the negative spillover between population growth during 1500 - 1860 to initial conditions of neighbours, IC_js (cfr. Column [10]). Finally, in Panel C we perform the same battery of correlation tests as we set $t_0 = 1700$. These tests still showcase some correlation of IC_js with the level of population density in Column [8]; however (and more importantly), we no longer observe any significant population pre-trend in Columns [9] and [10]. This is reassuring against endogeneity concerns - e.g., reverse causality, going from economic development to demand for technical skills, and thus the distribution of faculties (particularly in STEM fields) measured in 1700.

In the empirical section, we present results for both $t_0 = 1860$ (with and without the top population decile in 1500), and $t_0 = 1700$. We also verify that our main results hold when we anticipate further initial conditions to 1600, 1500, and 1400, thus even before the start of the Renaissance. We present an extensive sensitivity analysis that ensures that our empirical results are not driven by other omitted factors that could have affected the initial distribution of faculties among pre-unitarian provinces and might be correlated with local observable characteristics. We address various forms of spatial auto-correlation. We present the results of additional robustness and placebo tests that serve to further validate our identifying assumptions. This thorough battery of robustness checks consistently yields similar qualitative and quantitative outcomes (see Section 5.1 and online Appendix D for detailed information). This reinforces our confidence that our estimates of b, ϕ_1, and ϕ_2 from Eqs. 1 and 2 are not being determined by unaccounted factors or outliers. This is a prerequisite for conducting the welfare analysis in Section 7.

While evidence presented in this section seems reassuring, it is worth noting that any remaining omitted factors deeply correlated with the underlying cross-provincial distribution of initial conditions (e.g., culture), would not invalidate our analysis. Rather, it would prompt an interpretation of F_i and F_{-i} as 'proximate' determinants of local welfare, albeit not exact ones.[33]

[33] We acknowledge Nico Voigtländer for highlighting this insightful point.

Table 3 Balancing and Pre-trend analysis for IC_j and IC_i

	Initial Levels (in 1861)								Pre-trends	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
	Dist. of i, Bologna	Dist. of i, Roma	Dist. of i, *Capital*	Per capita GVA of i	Share of agric in i	Share of industry in i	Urban. in i *(6000 inh.)*	Total pop. density	$\Delta_{800-1500}$ pop. in i	Δ_{1500-t_0} pop. in i
	Panel A: Initial conditions$t_0 = 1860$									
IC_j	–0.0010	–0.0007	–0.0001	0.3784	–0.0048	–0.0142	–0.6109	–0.0012	3.1237	–1.8473
	(0.0010)	(0.0014)	(0.0013)	(0.6836)	(0.0079)	(0.0167)	(0.9122)	(0.0003)***	(3.3786)	(0.4875)***
	[0.0006]	[0.0017]	[0.0013]	[0.7926]	[0.0076]	[0.0191]	[0.6184]	[0.0005]*	[5.1109]	[0.8250]*
	{0.0012}	{0.0010}	{0.0014}	{0.9197}	{0.0062}	{0.0151}	{0.7251}	{0.0006}	{3.4460}	{0.5053}**
Observations	286	286	286	286	286	286	239	286	286	286
IC_i	–0.0007	–0.0008	–0.0006	1.6973	–0.0429	0.0395	0.7859	0.0037	–10.5807	2.7827
	(0.0009)	(0.0016)	(0.0019)	(1.1806)	(0.0233)*	(0.0347)***	(1.2903)***	(0.0006)	(5.4222)	(1.7422)
	[0.0008]	[0.0018]	[0.0011]	[1.2095]	[0.0265]	[0.0322]	[0.5615]	[0.0007]***	[7.0213]	[0.8633]**
	{0.0011}	{0.0013}	{0.0022}	{1.2111}	{0.0211}	{0.0099}**	{1.1000)	{0.0007}***	{5.7625}	{1.1661}*
Observations	69	69	69	69	69	69	59	69	69	69
	Panel B: Initial conditions$t_0 = 1860$, *without provinces in the top population decile*									
IC_j	–0.0008	–0.0004	–0.0005	0.7907	–0.0118	–0.0054	–0.0311	–0.0016	4.2986	–1.5042
	(0.0009)	(0.0018)	(0.0015)	(0.7362)	(0.0073)	(0.0174)	(0.7529)	(0.0030)	(3.5391)	(0.4955)***
	[0.0005]	[0.0022]	[0.0014]	[0.8962]	[0.0104]	[0.0226]	[0.6943]	[0.0050]	[5.3433]	[1.0378]
	{0.0011}	{0.0015}	{0.0014}	{0.9632}	{0.0120}	{0.0165}	{0.3273}	{0.0044}	{3.8794}	{0.5163}**
Observations	254	254	254	254	254	254	212	254	254	254
IC_i	–0.0007	0.0002	0.0019	0.3845	–0.0268	0.0304	0.3791	0.0074	–13.2706	1.4551
	(0.0011)	(0.0013)	(0.0016)	(1.1736)	(0.0187)	(0.0331)	(0.5802)	(0.0056)	(5.3017)**	(1.6602)
	[0.0007]	[0.0016]	[0.0017]	[1.4404]	[0.0199]	[0.0257]	[0.4790]	[0.0055]	[6.9956]	[1.4389]
	{0.0014}	{0.0017}	{0.0017}	{1.4853}	{0.0141}	{0.0206}	{0.7787}	{0.0046}	{6.4548}	{1.7582}
Observations	61	61	61	61	61	61	52	61	61	61
	Panel C: Initial conditions$t_0 = 1700$									
IC_j	–0.0009	–0.0007	0.0006	0.3125	–0.0030	–0.0085	–0.4819	–0.0007	–0.3846	9.0179
	(0.0006)	(0.0008)	(0.0007)	(0.4263)	(0.0042)	(0.0081)	(0.5275)	(0.0002)***	(2.3629)	(13.7687)

Table 3 (continued)

	Initial Levels (in 1861)								Pre-trends	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
	Dist. of *i*, Bologna	Dist. of *i*, Roma	Dist. of *i*, *Capital*	Per capita GVA of *i*	Share of agric in *i*	Share of industry in *i*	Urban. in *i* *(6000 inh.)*	Total pop. density	$\Delta_{800-1500}$ pop. in *i*	Δ_{1500-t_0} pop. in *i*
	[0.0003]**	[0.0009]	[0.0009]	[0.5627]	[0.0038]	[0.0083]	[0.3585]	[0.0003]**	[3.0012]	[9.6263]
	{0.0008}	{0.0008}	{0.0007}	{0.5982}	{0.0043}	{0.0072}	{0.4958}	{0.0004}*	{1.7126}	{18.1201}
Observations	286	286	286	286	286	286	239	286	286	286
IC_i	–0.0007	–0.0008	0.0008	0.5583	–0.0130	–0.0002	0.1904	0.0009	–5.8161	11.5749
	(0.0006)	(0.0009)	(0.0012)	(0.7174)	(0.0106)	(0.0174)	(0.5142)	(0.0005)	(3.6941)	(24.4610)
	[0.0005]	[0.0010]	[0.0010]	[0.8530]	[0.0105]	[0.0138]	[0.2952]	[0.0005]	[4.9745]	[16.3630]
	{0.0007}	{0.0011}	{0.0011}	{0.9075}	{0.0091}	{0.0090}	{0.4456}	{0.0005}	{4.5848}	{27.9605}
Observations	69	69	69	69	69	69	59	69	69	69

Notes: The sample includes only the 69 pre-unitarian provinces. In Columns [1] - [8], we report correlations of initial conditions with time invariant characteristics of province *i*. The urbanisation rate is measured in 1861 for provinces that already belonged to the Italian Kingdom by then. All other variables are measured in 1870 for Belluno, Mantova, Padova, Roma, Rovigo, Treviso, Udine, Vicenza, and Verona and in 1861 for all other provinces. For column [3] refer to Figure A-II in the online Appendix A1 for identifying the capital cities of pre-unitarian regions.In Columns [9] and [10], we report correlations of initial conditions with growth rates of province i's population as reported in each column. Standard errors are clustered at the NUTS2 regional level between (). Standard errors are clustered at the NUTS1 regional level between []. Standard errors are clustered at the level of the pre-unitarian country between {}. Significance levels: $*$: 10% $**$: 5% $***$: 1%

Table 4 Neighbourhood and productivity effects in the higher education supply

	All provinces				Pre-unitarian provinces			
	[1] OLS FE	[2] IV FE	[3] IV FE	[4] IV FE	[5] OLS FE	[6] IV FE	[7] IV FE	[8] IV FE
		(t_0=1860)	(t_0=1860; drop top pop. decile)	(t_0=1700)		(t_0=1860)	(t_0=1860; drop top pop. decile)	(t_0=1700)
	Panel A: Neighbourhood effects							
Total no. faculties in *j*	-0.088**	-0.217***	-0.129***	-0.260***	-0.042*	-0.162**	-0.155**	-0.205**
	(0.042)	(0.074)	(0.048)	(0.094)	(0.023)	(0.066)	(0.065)	(0.094)
Total no. of universities in *i*	3.339***	3.288***	2.633***	3.271***	3.651***	3.595***	2.453***	3.575***
	(0.683)	(0.626)	(0.318)	(0.627)	(0.639)	(0.593)	(0.467)	(0.595)
Total no. of universities in *j*	0.023	0.018	0.011	0.017	0.015	0.005	0.003	0.001
	(0.022)	(0.024)	(0.013)	(0.026)	(0.025)	(0.025)	(0.015)	(0.028)
K-P rk Wald F-stat		54.996	46.711	40.894		50.558	39.140	29.650
K-P rk LM-stat		38.154	33.518	30.230		28.527	23.574	20.672
p-value KP		0.000	0.000	0.000		0.000	0.000	0.000
		F_{jt}	F_{jt}	F_{jt}		F_{jt}	F_{jt}	F_{jt}
B_{jt}		0.311***	0.280***	0.236***		0.305***	0.280***	0.203***
		(0.042)	(0.045)	(0.037)		(0.040)	(0.045)	(0.037)
Observations	5209	5209	4703	5209	3868	3868	3496	3868
	Panel B: Productivity							
Total no. faculties in *i*	0.287***	0.326***	0.327**	0.349***	0.241**	0.283**	0.333**	0.311**
	(0.071)	(0.102)	(0.134)	(0.109)	(0.115)	(0.136)	(0.150)	(0.150)

Table 4 (continued)

	All provinces				Pre-unitarian provinces			
	[1] OLS FE	[2] IV FE	[3] IV FE	[4] IV FE	[5] OLS FE	[6] IV FE	[7] IV FE	[8] IV FE
		(t_0=1860)	(t_0=1860; drop top pop. decile)	(t_0=1700)		(t_0=1860)	(t_0=1860; drop top pop. decile)	(t_0=1700)
Total no. faculties in $-i$	0.060*	0.108**	0.095*	0.127**	0.071*	0.116**	0.092*	0.130**
	(0.035)	(0.053)	(0.053)	(0.050)	(0.039)	(0.056)	(0.056)	(0.052)
Total no. of universities in i	-0.093	-0.245	-0.406	-0.329	0.123	-0.044	-0.190	-0.142
	(0.702)	(0.632)	(0.501)	(0.609)	(0.932)	(0.791)	(0.518)	(0.762)
Total no. of universities in $-i$	-0.101	-0.256	-0.458*	-0.319	-0.161	-0.307	-0.445*	-0.349
	(0.223)	(0.256)	(0.240)	(0.241)	(0.257)	(0.278)	(0.255)	(0.256)
K-P rk Wald F-stat		10.751	18.083	12.602		22.819	12.514	34.063
K-P rk LM-stat		18.551	22.132	14.900		22.315	20.085	14.136
p-value KP		0.000	0.000	0.000		0.000	0.000	0.000

Table 4 (continued)

	All provinces							Pre-unitarian provinces						
	[1] OLS FE	[2] IV FE		[3] IV FE		[4] IV FE		[5] OLS FE	[6] IV FE		[7] IV FE		[8] IV FE	
		(t_0=1860)		(t_0=1860; drop top pop. decile)		(t_0=1700)			(t_0=1860)		(t_0=1860; drop top pop. decile)		(t_0=1700)	
		F_{it}	F_{-it}	F_{it}	F_{-it}	F_{it}	F_{-it}		F_{it}	F_{-it}	F_{it}	F_{-it}	F_{it}	F_{-it}
B_{it}		0.301***	0.045	0.268***	0.085	0.261***	0.052		0.247***	0.093	0.261***	0.085	0.208***	0.095
		(0.060)	(0.091)	(0.040)	(0.087)	(0.050)	(0.079)		(0.037)	(0.080)	(0.041)	(0.092)	(0.025)	(0.067)
B_{-it}		-0.002	0.336***	-0.005	0.329***	-0.012	0.293***		0.012	0.325***	-0.003	0.333**	-0.002	0.289****
		(0.022)	(0.053)	(0.017)	(0.053)	(0.019)	(0.046)		(0.020)	(0.054)	(0.018)	(0.057)	(0.018)	(0.047)
Observations	1231	1231		1112		1231		1025	1025		906		1025	

Notes:. Regressions in Column [1] are based upon the entire set of provinces, including those established during 1870-2010 (these are 110 provinces). Regressions in Column [2] exploit the sample of pre-unitarian provinces only for the entire period (these are 69 provinces). In each column, we present three sets of IV FE estimates (which are performed using 2SLS estimation) with Bartik instruments. The first set computed the instruments using initial conditions at time$t_0 = 1860$; the second set uses initial conditions at time$t_0 = 1700$. The third set considers$t_0 = 1860$, but excludes provinces that belonged to the top decile of the Italian total population in 1500 (Bari, Caserta, Como, Lecce, Milano, Napoli, Roma, Torino). This implies the last set of IV estimates is based on a sample of Regressions of 102 provinces in Column [1], and 61 provinces in Column [2]. In all regressions we take 10 years lags of regressor and controls. Specifications in Panel A include province pair fixed effects. Specifications in Panel B include province fixed effects. All specifications include region-by-year fixed effects. They also include the total number of universities in province i, and its neighbours (at the level of j province in Panel A, and in the neighbourhood$-i$in Panel B) as provincial controls. In IV estimates, the total number of faculties in province $i, j, -i$is instrumented by the respective Bartik instrumentB_{it},B_{jt},B_{-it}, as they are defined in Eq. 4 above. Standard errors clustered at the level of the local province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%

5 Empirical results

Table 4 showcases the main empirical findings. Panel A presents neighbourhood effects derived from Eq. 1, while Panel B displays the estimates of productivity effects from Eq. 2. In Columns [1] - [4], we presents results for all provinces in our sample (110 in the baseline), while Columns [5] - [8] concentrate solely on pre-unitarian provinces (69 in the baseline). We present OLS FE estimates (Columns [1], [5]), and three different IV FE estimates. In the first IV estimation (Columns [2], [6]), we use Bartik instrument which considers $t_0 = 1860$. In the second one (Columns [3], [7]), we still use $t_0 = 1860$, and drop provinces in the top decile of the population distribution. In the third (Columns [4], [8]), we assume $t_0 = 1700$ instead. In all estimations, we cluster standard errors at the level of the local province.[34]

First stage results are satisfactory in both panels. The first stage coefficients of each Bartik instrument are strongly correlated with their corresponding indicator for the total number of faculties. F-statistics signal that instruments have strong predictive power.[35] In Panel A, the coefficient of the total number of faculties in a neighbouring province is always negative and significant at the 1% or 5% level for both OLS FE and IV FE estimates, irrespective of the specification. IV FE coefficients are up to five times larger than the corresponding OLS FE coefficient. This observation suggests that unobservable and time-varying factors contribute to increasing the number of faculties in both province *i* and its neighbours. These factors diminish the OLS estimates of neighbourhood effects. As discussed above, a plausible explanation for this trend lies in technical advancements under the Post-Malthusian regime, which generate a local demand for human capital. This may induce an expansion of higher education supply in both province *i* and province *j*, consequently moderating the neighbourhood effects estimated by Eq. 1 using the OLS FE model.[36]

Taken at face value, IV FE estimates of Panel A suggest that having one faculty more in province *j* is associated with a decrease by $0.16 - 0.26$ local faculties in province *i*, on average. We can better gauge the magnitude of this effect considering that each Italian province has four neighbours, on average. IV-FE estimates in Column [8] suggest that a one standard deviation increase in higher education supply in each neighbour *j* (4.61 faculties) results in up to $-3.69 (= -0.20 * 4.61 * 4)$ faculties in province *i* over a ten years' period. This is a decrease in local supply 40% larger than the sample mean, and four fifths of the standard deviation of local supply of province *i* (cfr. Table 2).

Turning to productivity effects, in Panel B, the estimated coefficients of the total number of faculties in province *i* and its neighbourhood are positive and statistically significant. Compared to Panel A, the higher similarity in the size of coefficients between OLS FE and IV FE suggests that province and region-by-year fixed effects already account for many

[34] Appendix B showcases results obtained through clustering standard errors at the regional or kingdom level, along with outcomes of Eq. 1 featuring two-way clustered standard errors at the *ij* level (refer to Table 10). These methodologies enable us to address potential concerns regarding correlation and heteroskedasticity within the data, thereby reinforcing the robustness of our findings.

[35] It should be acknowledged that in dyadic regressions (such as those presented in Panel A) having multiple similar entries may result in inflated *t*-stat and F-stat values compared to having only one entry for each province-time combination. We thank an anonymous referee for this remark. In online Appendix E we provide a detailed discussion and argue that a reasonable estimate of the correct F-stat is lower than the F-stat computed in Table 4 Panel A, but still safely above the standard threshold level of 10. See Table A-XV.

[36] Other potential factors that could counteract these effects include administrative agreements, political arrangements, or the establishment of forty-one new provinces throughout Italian history, which might mitigate competition among existing *ij* pairs. We explore some of these hypotheses in the sensitivity analysis.

time-varying unobserved factors that influence both productivity and higher education supply at the provincial level. From these coefficients, we quantify GVA returns of higher education supply, thus the opportunity cost of neighbourhood effects in productivity terms. IV FE estimates suggest that, on average, one more faculty in the local province is worth about 280 €−350€ per capita after ten years, while one more faculty in the neighbourhood is worth 90€−130€ per capita to the local province (after ten years). More precisely, IV FE estimates in, e.g., Column [8] imply that one standard deviation increase in local supply (4.3 faculties) and neighbours' supply (10.4 faculties) create in ten years per capita value to the local province up to 2,689€= $(311 * 4.3 + 130 * 10.4)$. This is equivalent to about one third of a standard deviation of GVA per capita. Notice that these are effects determined by a large scale intervention that increases the relevant supply of higher education by 14.7 faculties in total (summing up local and neighbours' supply). Such an intervention cannot be plausibly put in place in one specific year. The decennial structure of our productivity data indeed suggests that our effects should be interpreted over a time horizon of ten years.

In what follows, we select as our preferred specification the one in Table 4 Column [8]. By doing so, we focus on the 69 pre-unitarian provinces only. This is meaningful from an identification perspective, as initial conditions cannot be defined for the 41 post-unitarian ones. Also the choice to measure initial conditions at $t_0 = 1700$ can be made without loss of generality: estimates in Column [8] are fully confirmed both in Column [6] as we choose $t_0 = 1860$ and in Column [7], as we drop provinces in the top population decile. However, the use of $t_0 = 1700$ is less exposed to pre-trend concerns than $t_0 = 1860$ (see Columns [9] and [10] of Table 3), where we evidenced some co-movement of initial conditions with population growth.[37]

In Appendix A we use the IV FE estimated coefficients in Column [8] to evaluate the productivity effect in terms of GVA per capita elasticity. We show that a balanced 10% increase in the supply of faculties in province i and its neighbourhood evaluated at sample averages produces an annual increase of 0.27% in per capita GVA over a time horizon of ten years. We also discuss to what extent this elasticity based on faculties openings is comparable with elasticity estimates obtained from university openings (notably by Valero and Van Reenen, 2019). In the welfare analysis we find a slightly smaller elasticity (0.20 instead of 0.27), when we evaluate effects at 2001 supply (see Section 7 for details). Finally, in online Appendix C we compare our estimates to the long-run effect of higher education supply that emerge from conditional convergence models (Barro, 1991, 2012; Gennaioli et al., 2013).

5.1 Sensitivity analysis

In this section we discuss robustness checks presented in Appendix B.

In Table 9 we carry out a first set of robustness checks to verify how our results change for alternative timing of initial conditions. We consider $t_0 = 1800, 1600, 1500, 1400$. By doing so, we compute instruments (see Eq. 4) using initial conditions that go back to Middle Ages, at the end of a century of declining population (cfr. Panel (b) of Fig. 1). While the point estimate for the local productivity effect of faculties is not precisely estimated when

[37] Also results in the sensitivity and mechanisms analyses presented in Sections 5.1 and 6 for $t_0 = 1700$ are confirmed in terms of both significance and direction as we use $t_0 = 1860$ with or without the top decile of population. These estimates are available upon request from the authors.

$t_0 = 1500, 1400$, our results are confirmed. Notice that going back in time reduces by up 50% the number of faculties considered to compute the initial conditions.

In Table 10 we carry out a first set of robustness checks to account for spatial autocorrelation, which may potentially bias our estimates. In Columns [1] and [2] we implement Conley's and Driscoll and Kraay (1998)'s corrections. In Column [3], we correct standard errors by two-way clustering, i.e. both at the local and neighbouring province. In Column [4] we cluster at the level of the pair *ij*, i.e. we allow for serial correlation within the same pair of observation. In Column [5] we explicitly account for the dyadic structure of the data used to estimate Eq. 1, and apply the correction proposed by Fafchamps and Gubert (2007); Aronow et al. (2015) and more recently by Carlson et al. (2024) to account for errors being correlated between all pairs that have in common either province *i* or *j* (or both).[38] Results from this first set of robustness shows that, if anything, standard errors are lower than in baseline estimates, so that the coefficients of interest are estimated more precisely. In Columns [6] and [7], the clustering of standard errors at the NUTS2 region, and the pre-unitarian state level (respectively) confirms our findings. Next, in Table 11 we account for potentially omitted factors, both for the local province and the neighbours, which may affect both faculty creation and GVA per capita. In Columns [1] and [2] we consider population shocks, or shocks on the overall level of development. We show that our results go through as we control for population growth of the province and its neighbourhood, and the share of employment in the agricultural sector in province and its neighbourhood, respectively.[39] After that, we consider unobserved quality of higher education supply at the province level. Thus, in Column [3] we control for the number of private universities and elite universities according to the Casati ranking (cfr. Table 1, Column [6]), while in Column [4] we include the interaction between the number of universities and the academic notability of university supply in the pre-unitarian province by De la Croix et al. (2024).[40] In Column [5], we consider altitude as a measure of ruggedness. So we include province-level interactions of a dummy equal to one if the local province has an average altitude above 350 metres (above sea level) and time fixed effects. Our results go through also in this robustness. In Column [6], we run our baseline specification without controls for the number of universities in the local province and its neighbours. Results are confirmed again, however loss in precision of estimates in Panel A as well as the drop of the F-stat in Panel B confirm it is important to control for this dimension of higher education supply in baseline estimates.

A battery of additional robustness and placebo exercises that provide further validations of our identifying assumptions is reported in Section D of the online Appendix. In line with suggestions by Goldsmith-Pinkham et al. (2020), we replicate the entire analysis using alternative estimators that combine the moment conditions in potentially more efficient ways, and check the sensitivity of our results to the adoption of alternative empirical strategies

[38] Robustness of Columns [3] to [5] can be implemented only for estimates of Eq. 1 as the unit of observation is the province pair *ij*. We cannot cluster two-ways standard errors computed when estimating Eq. 2, as the unit of observation for these estimate is the local province *i*.

[39] In the next section we show that sectoral employment shares are to be viewed as alternative productivity outcomes in the Post-Malthusian transition during the time span of our analysis. As such, these variables may be viewed as 'bad controls' in our setting. This is why we do not include them in the main analysis.

[40] The quality assessments provided by the Casati ranking is time-varying, due to updates made during Italian history (see online Appendix B for details). Conversely, the time variation in the academic notability index stops at the end of the pre-unitarian period. So we take its last value (which refers to the 1740-1800 period) for the interaction term.

(see Table A-VI). In these cases, the results remain qualitatively unchanged in terms of both significance and direction. For neighbourhood effects, we confirm our results as we consider alternative and time invariant definitions of territory and borders of pre-unitarian provinces (see Table A-VII); we verify that our results are not driven by specific sub-periods or regional outliers (see Table A-VIII); we apply the Adão et al. (2020) (AKM) correction to account for arbitrary correlation in the error term across provinces with similar initial conditions (see Table A-IX).

In Fig. D we display coefficients from a complementary exercise, where we replicate estimates of Eq. 1 100 times, using random draws of 30 provinces out of the 49 in our sample with a initial supply of faculties equal to zero, in order to have a more balanced number of provinces with and without initial faculties. With these reduced samples, in some cases estimates are less precise than with the full sample. However, on average, size and statistical significance of the estimated coefficients from these replications are comparable to the baseline estimates. This suggests that our main results are not driven by unobserved province-level factors, possibly correlated with an initial zero supply of faculties.

Finally, in Table A-XI we check the sensitivity of productivity effects to an alternative definition of initial shares. In each column, the placebo Bartik instrument is constructed multiplying the total number of faculties in any given year in province i (or in provinces $-i$) with initial shares (see Eq. 3) computed upon initial characteristics of pre-unitarian provinces, other than its initial number of faculties.[41] The power of first stages drop tremendously, to worryingly low levels, which makes second stage estimates of Eq. 2 meaningless. Indeed, the pattern in the estimated coefficients does not seem anyway close to the effects estimated in Panel B, Column [6] of Table 4, and are in general not significant. This suggests that, indeed, only the interaction with shares of initial faculties is what matters for our results.

6 Mechanism analysis

The aim of this section is twofold: firstly, to validate empirically the conceptual framework discussed in Section 2.2 for the emergence of neighbourhood effects; and secondly, to discover evidence that aligns with the insights from Unified Growth Theory regarding the role of human capital formation in driving economic development (Galor, 2011). We discuss these channels in the next two sub-sections.

6.1 Neighbourhood effects: a spatial competition channel?

An unanswered question is whether neighbourhood effects can be understood within the conceptual framework of spatial competition described in Section 2.2, or if alternative explanations could be at play. To shed light on this question, Table 5 offers insights into potential mechanisms driving neighbourhood effects. We explore varying effects between

[41] These characteristics are total GVA, GVA per capita, GVA per worker, participation rate, population, population density, share of active population in agriculture, services and industry. Since for $t_0 = 1700$ we do not have information for all of these characteristics, the placebo initial conditions IC_{it_0} we use for this exercise are computed for $t_0 = 1860$.

STEM and non-STEM fields; analyse historical, regional, and political factors; delve into mobility costs; and investigate the spatial reach of neighbourhood effects.

STEM vs non-STEM Faculties We start with the field of study, and analyse separately neighbourhood effects triggered by faculties that operate in STEM and non-STEM fields of study.

Table 5 Mechanisms for neighbourhood effects

	Panel A: across study field		*Panel B: within study field*	
	[1] non-STEM (outcome var.) on STEM	[2] STEM (outcome var.) on non-STEM	[3] STEM on STEM	[4] non-STEM on non-STEM
STEM faculties in *j*	-0.330*		-0.224**	
	(0.195)		(0.092)	
no-STEM faculties in *j*		-0.047		-0.201*
		(0.041)		(0.107)
Observations	3868	3868	3868	3868
K-P rk Wald F-stat	53.173	27.456	53.173	27.456
	Panel C: pre-WWI vs. post-WWI period		*Panel D: Intra vs. Inter-regional effects*	
	[5] pre-WWI	[6] post-WWI	[7] different region	[8] same region
no. of faculties in *j*	-0.081**	-0.216**	-0.062	-0.235**
	(0.036)	(0.105)	(0.114)	(0.116)
Observations	1370	2498	1640	2228
K-P rk Wald F-stat	48.105	23.850	3.475	127.548
	Panel E: Political color before 1884		*Panel F: Outmigration*	
	[9] same political color	[10] different political color	[11] low outmigration	[12] high outmigration
no. of faculties in *j*	-0.510***	-0.263*	-0.323***	-0.003
	(0.192)	(0.137)	(0.124)	(0.093)
Observations	1118	1469	1959	1909
K-P rk Wald F-stat	24.084	19.580	39.453	8.165
	Panel G: Spatial reach (linear distance)			
	[13] within 90Km	[14] btw 90Km and 180Km	[15] btw 180Km and 270Km	[16] btw 270 and 360Km
no. of faculties in *j*	-0.199**	-0.036	0.007	-0.058*
	(0.081)	(0.036)	(0.029)	(0.034)
Observations	4801	9504	10646	8378
K-P rk Wald F-stat	30.651	71.448	101.425	58.877

Notes: In Panels A and B, we use the STEM and non-STEM classification described in Fig. 2. In Panel C, the pre-WWI period includes years 1870-1911 while post-WWI period starts in 1921. In Panel D, intra-regional effects are estimated between provinces in the same NUTS2 region, while inter-regional effects are between provinces belonging to different NUTS2 regions. In Panel E, we define the political colour of a province based upon the party (historical left or historical right) of the 50%+1 of the deputies from this province elected in the parliament of the Italian Kingdom between legislatures VIII (1861-65) and XIII (1876-80). In Panel F we split the sample between provinces characterised by an out-migration rate in 1881 below or above the median (2.07%). The out-migration rate considers only international emigration to a European country, or transatlantic emigration. In Panel G, we use linear distances between province centroids extracted by the Google Maps. Details on sources and construction of these indicators are in online Appendix B. Standard errors are clustered at the province level. Significance levels:$*$: 10%$**$: 5%$***$: 1%

Panel A examines cross-field neighbourhood effects, while Panel B focuses on within-field neighbourhood effects. The results demonstrate that non-STEM faculties are influenced by both STEM and no-STEM fields (see Columns [1] and [4], respectively). In contrast, STEM faculties are only affected by neighbouring faculties within the same field. These findings align with the perspective discussed in Section 2.2, suggesting that diversifying own supply from supply of provinces nearby may protect the local province from neighbourhood effects. More precisely, STEM faculties may still experience neighbourhood effects from neighbouring STEM faculties. However they are not subject to displacement forces arising from non-STEM faculties, which have proliferated in more recent periods (see Fig. 2(a)).

Historical and regional factors Next, we investigate the role of historical and regional factors. In Panel C, we show that neighbourhood effects exist both before and after WWI, being more than twice as large in the second spell (i.e., 1921 - 2010). As argued above, this period sees the transition from the Post-Malthusian to the Sustained Growth Regime, and meanwhile the transition from elite to mass higher education. In Panel D we account for the fact that provinces i and j may or may not belong to the same administrative region. Thus, we analyse separately neighbourhood effects between provinces that belong to different regions (cfr. Column [7]), and between provinces that belong to the same region (cfr. Column [8]). Caution is needed in interpreting neighbourhood effects across regions in Column [7] due to the low power of the first stage; neighbourhood effects appear large and precisely estimated within the same region, instead. The negative, non significant coefficient in Column [7] suggests that displacement forces between provinces in different regions are weaker, even though these provinces share a border. This result is consistent with the local mobility assumption presented in Section 2.2, as it is well known that Italian university students are much less mobile between different regions than within the same region (DeAngelis et al., 2017).[42]

History of political coordination Results in Panel D may also emphasise intra-regional cooperation as a trigger of neighbourhood effects: from a public governance perspective, coordinating on the supply of higher education is easier for provinces belonging to the same region than for provinces belonging to different regions. The next heterogeneity exercise provides more insight into the political cooperation versus spatial competition interpretation of neighbourhood effects. In Panel E, we allow for heterogeneous effects between provinces that historically shared the same political colour (Column [9]), and provinces historically characterised by a different political orientation (Column [10]).[43] A concentration of neigh-

[42] For example, in 2007 only 10% of 1st-year BSc students enrolled in a university outside their region of residence. More than 80% of these students chose to study in their home region. Among them, 28% moved to another province within the same region, while over half of the students stayed in their province of residence. This trend was even more pronounced in the regions of the ‘Mezzogiorno,’ where only 5% of students ventured to a province in a different region (cfr. DeAngelis et al., 2017; Table 1).

[43] We define the historical political orientation of the province based on the party (either historical left or historical right) of the majority of local Deputies elected in the Lower House of the Italian Parliament at national elections taking place during 1861-1880. We focus on Deputies, as they were elected through majority rule, while Senators and other key province-level political roles were not directly elected but appointed. To ensure sufficient data, we consider multiple national elections. However, we stop data collection after the 13th legislature (1876-1880) because, from the 1880 elections onwards, salience of deputies’ political party affiliations declines due to the phenomenon commonly known as ‘Transformism’. See online Appendix B3 for details over the construction of this variable.

bourhood effects within historically politically-aligned pairs of provinces would suggest strong evidence for pure political coordination channel (i.e., provinces with the same history of political representation may coordinate more easily). Estimates in Panel E do not seem to support a pure political coordination explanation, as neighbourhood effects are sizeable and significant in both columns. Nonetheless, the coefficient in Column [9] is larger than coefficient in Column [10], indicating the potential presence of political channels that complement spatial competition mechanisms rather than offering an alternative explanation.

Mobility costs In Section 2.2 we discussed mobility costs as the main trigger to spatial competition mechanisms. We argued neighbourhood effects arise as long as there is some mobility between *ij* province couples, and this is somewhat limited to the neighbourhood itself (e.g., due to high costs faced by individuals to migrate to a non-neighbouring province, another distant region or abroad). In Panel F, we consider costs of migrating outside the neighbourhood. We proxy them by historical province level emigration rates abroad (to another European country or the US) between 1861 and 1884 (Gray et al., 2019). We distinguish provinces characterised by an out-migration rate below and above the median (equal to a yearly out-migration between 1861 and 1884 of about 2 x 1000 inhabitants). Neighbourhood effects are concentrated in low emigration provinces (cfr. Column [11]), while there is no evidence in high emigration provinces (cfr. Column [12]). This provides further support to the local mobility assumption, indicating that neighbourhood effects emerge in areas with a stable incumbent population that consistently demands education to foster local skills and human capital.

In Table 12 of Appendix B, we consider two additional heterogeneity exercises. First we use altitude as a measure of geographic isolation, and we estimate neighbourhood effects separately for relatively isolated provinces (with an average altitude above the median, 350 meters above sea level) and more accessible provinces (with an altitude below the median). Then, we use the presence of transportation infrastructure in the pre-sample period as a historical proxy for highly accessible provinces. We examine the effects separately for provinces that lacked any railway endowment and provinces that already had some railway endowment (i.e., at least one railway connection) by 1861. Reported coefficients are not always precisely estimated, but results suggest that neighbourhood effects are concentrated among smoother, more geographically accessible provinces, characterised by an initial endowment of railways. This evidence is consistent with the hypothesis that accessibility increases the salience of neighbourhood effects, arguably through a higher local mobility.

Spatial reach of neighbourhood effects The higher education supply of province *i* may be influenced by supply in province *j* if the two provinces are positioned close enough, regardless of whether they share a border or not. In Panel G of Table 5, we propose a complementary definition of a neighbourhood, based on the linear distance d_{ij} between province capitals, regardless of their contiguity. In practice, we compute *i*-to-*j* distances from Google Maps applications and assume province *j* is a neighbour of province *i* as long as $d_{ij} < D$, where D is a supposed distance bandwidth for neighbourhood effects (see Parchet, 2019). In Columns [13] - [16], we consider *i*-to-*j* linear distances $D = 90, 180, 270, 360$km. The results reveal that, on average, neighbourhood effects occur within a 90-kilometer linear distance. In Table A-XII of online Appendix D, we replicate the analysis based on travel-based bandwidths measured as of today. We show that, on average, neighbourhood effects operate within a bandwidth of 120 kilometres of travel, and a bandwidth of 80 minutes of travel time. This is the average 'spatial reach' of neighbourhood effects, measured in linear distance, travel distance, or travel time terms.[44] Beyond these spatial reaches, estimated coefficients are much smaller and non significant. This does not exclude that neighbourhood effects may be at work, e.g., at a small scale and within specific sets of *ij* couples, even though they are very distant.[45]

6.2 Productivity effects: industrialisation and human capital formation

In this section, we analyse the productivity effects of higher education institutions through the lens of the Unified Growth Theory. We point out effects on the multiple components of GVA per capita; we scrutinise the impact of STEM supply, and its interaction with early industrialisation, and explore sectoral composition changes. This will help us to highlight the role of higher education through the Post-Malthusian era, and the transition to Sustained Economic Growth in Italy.

Break down of the productivity effect In Table 6, we present the results of Eq. 2 on a set of outcomes that allow to disaggregate the effect of higher education supply on GVA per capita. For comparability purposes in Column [1] we report baseline estimates (cfr. Table 4, Panel B, Column [2]). In Columns [2] and [3], we show a positive effect of local higher education supply on both population and population density, respectively. This describes a yearly elasticity of population to higher education supply equal to 0.03, which is equivalent to the 4.3% of the spike in population growth Western European countries experienced between 1870 and 1913, on average (equal to 0.7; cfr. Figure 2.9 (b) from Galor, 2011).[46] Figure 10 in Appendix B, shows these returns start to decline in the immediate aftermath of WWI, then disappear after WWII. This timing seems consistent with the demographic transition in Italy (cfr. e.g. Galor, 2011 Figure 2.17). It is also in line with the transition of the Italian higher education from elite knowledge to modern tertiary education. Overall, we can interpret estimates in Table 6 (Columns [2], [3]), and Fig. 10 as small yet noticeable Post-Malthusian returns of upper tail human capital. In Table A-XIV of the online Appendix we use the pre-1500 value of the index of historical potential crop yield by Galor and Özak (2016) to investigate the trajectory of provinces potentially at different stages of development. These estimates suggest a larger, more precise effect of higher education supply on productivity measures (GVA per capita and per worker) as we include in the sample provinces in the top quartile of the distribution of the caloric suitability index. Conversely, higher education supply displays a larger and more precisely estimated effect on population and

[44] Table A-XIII in the online Appendix presents results with complementary neighbourhood definitions, incorporating factors such as distance and contiguity, as well as spatial restrictions imposed by Italian geography.

[45] For example, Table A-X in the online Appendix reports the distribution of negative and statistically significant coefficients as we re-estimate neighbourhood effects considering for each province in the sample, 100 random draws of five fictitious neighbours across provinces that are located over 500 km away. If we repeat the exercise 100 times, (100 repetitions of 100 extractions each), we observe that, on average, only about the 57% of extractions involves some evidence of neighbourhood effects.

[46] We compute a population elasticity of 0.31(=0.062*2.619/0.517) over a decennial time horizon, or a 0.03 on a yearly basis, on average. This positive effect confirms population would be a bad control in baseline productivity estimates. The same holds true for sectoral employment shares (see below).

Table 6 Compositional effects of higher education supply on productivity

	[1] GVA per capita	[2] Population	[3] Pop. density	[4] Participation rate	[5] GVA per worker
Total no. faculties in i	0.311**	0.062***	0.310**	0.417	0.551*
	(0.150)	(0.019)	(0.147)	(0.261)	(0.295)
Total no. faculties in $-i$	0.130**	-0.017**	-0.044	0.073	0.289***
	(0.052)	(0.007)	(0.029)	(0.140)	(0.090)
Observations	1025	1025	1025	1025	1025
K-P rk Wald F-stat	34.063	34.063	34.063	34.063	34.063

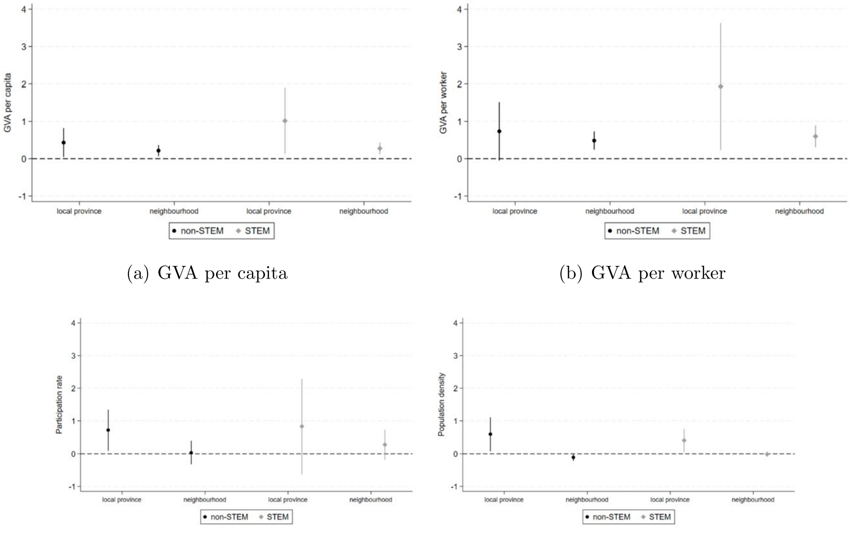
Note: Column [1] reports IV FE results of the baseline specification (see Table 4, Panel B, Column [2]) for comparison purposes. In Columns [2]-[5] we report results using related measures of province-level productivity. In Column [2] we use total population. In Column [3], we use population density measured as total population per Km^2. In Column [4] we use participation rate in the labor market (total active population/total population). In Column [5] we consider GVA per worker (total GVA of the province/total active population). All specification include the usual set of controls (i.e. the total number of universities in province i, and its neighbourhood $-i$), province and region-by-year fixed effects. Standard errors are clustered at the province level. Significance levels: $*$: 10% $**$: 5% $***$: 1%

population density, as we include in the sample provinces in the bottom quartile of the index distribution (thus more likely to be on a Malthusian stage).

Neighbourhood faculties show a negative coefficient on local population in Column [2], suggesting they may counterbalance population concentration within the province; however, this impact is not statistically significant in terms of population density in Column [3]. In Columns [4] and [5], we consider participation rate of workers in the labor market and GVA per worker, respectively the extensive and intensive margins of productivity. Our findings indicate that local supply of higher education contributes to both margins of productivity, although coefficients are not estimated too precisely. Faculties in the neighbourhood create positive externalities only on the intensive margin.

(STEM) faculties, industrialisation, and economic transition So far we implicitly assumed that all faculties provide an equal boost to local productivity. However, it is realistic assuming heterogeneous effects among institutions operating in different fields: for example the presence of an engineering faculty (in the province and/or in the neighbourhood) may have more sizeable productivity effects than a humanities faculty. In Fig. 6, we present separate regression plots for the estimated coefficients of STEM and non-STEM faculties on productivity measures reported in Table 6. Evidence in this figure suggests that productivity effects of STEM education are more localised compared to non-STEM faculties. In fact, local STEM faculties have approximately double the economic impact compared to local non-STEM faculties on GVA measures. Conversely the productivity externality does not seem to differ significantly between STEM and non-STEM faculties, as well as effects on participation rate and population density.

One possible explanation for evidence in Fig. 6 is a stronger interaction of STEM faculties with the local economy, particularly during early phases of industrialisation. STEM human capital could provide a crucial local endowment to reach further stages of local development. To explore this hypothesis, we investigate whether the productivity effect of higher education supply varies based on the early stage of industrialisation in the province.



(a) GVA per capita (b) GVA per worker

(c) Participation rate (d) Population density

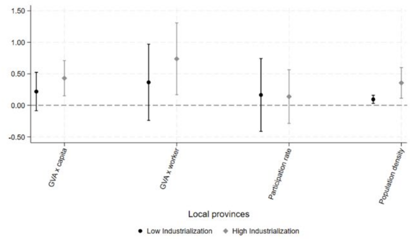
Notes: authors' calculations on HIU and Italian Census data. The graphs plot coefficients of IV estimates of Eq. 2 for the number of faculties in the STEM and non-STEM fields in the local province and its neighbourhood (using respective STEM and no-STEM version of Bartik instruments shown in Eq. 4) on the dependent variables reported in Panels (a)-(d)). All the regressions include the usual set of controls (i.e. the total number of universities in province i, and its neighbourhood $-i$), region-by-year and province fixed effect. Details of all regressions are available to the authors upon request. Standard errors are clustered at the province level. The shadowed area represent the 90% interval of confidence.

Fig. 6 Productivity effects of higher education supply by higher education sector (STEM vs. non-STEM)

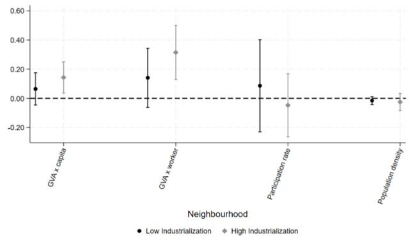
Thus, we categorise provinces depending on whether they display a 'low' early level of industrialisation (below the sample median of the index by Ciccarelli and Fenoltea, 2013), or a 'high' level of early industrialisation (above the sample median).[47] In Fig. 7, we show the estimated coefficients for all faculties and STEM faculties; Panels (a) and (c) report coefficients for faculties in the local province; Panels (b) and (d) report coefficients for faculties in the neighbourhood. Results suggest a complementarity between early industrialisation and higher education supply, particularly of STEM faculties. Productivity effects are concentrated in provinces with high early industrialisation. Furthermore, productivity returns are magnified as we dive into the supply of STEM faculties (cfr. y-axes in Panels (b) and (d) have very different scales).

If our estimates are capturing the contribution of higher education supply to the structural transformation of the Italian production system through the transition from the Post-Malthusian to the Modern Growth Regime, we should observe meaningful effects of higher education supply on sectoral employment shares in industry and services relative to agriculture. Figure 8 plots the sectoral composition effects of total supply, and STEM supply respectively. Panel (a) reports estimated coefficients on the industry-to-agriculture employ-

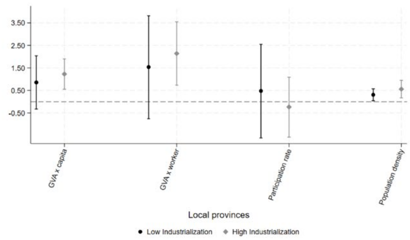
[47] The relative index of early industrialisation by Ciccarelli and Fenoltea (2013) measures the province's average share of the national industrial value added, relative to the share of the male population at the beginning of our sample period (1881-1911).



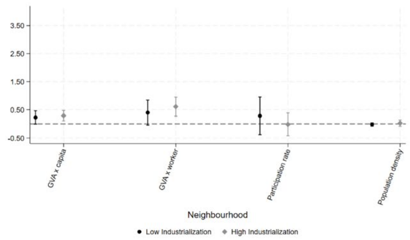
(a) all faculties, local province



(b) all faculties, neighbourhood



(c) STEM faculties, local province



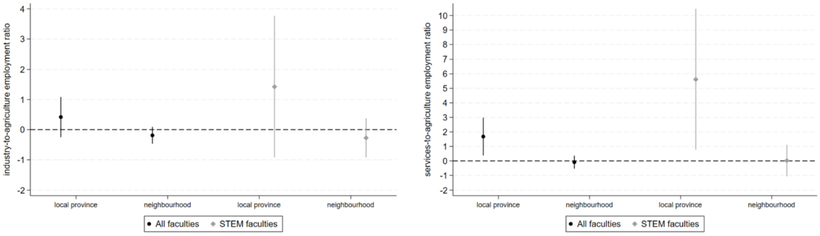
(d) STEM faculties, neighbourhood

Notes: authors' calculations on HIU and Italian Census data. The graphs plot coefficients of IV estimates of Eq. 2 for all faculties (Panel (a) for local faculties, and (b) for faculties in the neighbourhood), and STEM faculties (Panels (c) for local faculties, and (d) for faculties in the neighbourhood), using respective versions of Bartik instruments shown in Eq. 4, on the dependent variables reported on the x axis. All the regressions include the usual set of controls (i.e. the total number of universities in province *i*, and its neighbourhood−*i*), region-by-year and province fixed effect. Details of all regressions are available to the authors upon request. Standard errors are clustered at the province level.

Fig. 7 Productivity effects of higher education supply by initial level of industrialisation

ment ratio; Panel (b) focuses on the services-to-agriculture employment ratio. Results are in line with expectations. Local supply of faculties is associated with increasing employment ratios, featuring a gradual province-level transition from agriculture to industry and services. Gradients of these sectoral transition are larger for STEM faculties, and more precisely estimated for transition to services than for transitions to industry: one more STEM faculty in the local province explains the 25% of a standard deviation of the industry-to-agriculture ratio (very noisily estimated), and the 58% of a standard deviation of the service-to-agriculture ratio over the sample period.[48] Interestingly, we find no positive externalities on sectoral ratios; if any, there is a negative externality on the local province's industry-to-agriculture intensity ratio, which seems in line with the negative externality on local population in Table 6 above: higher education supply in neighbouring provinces not only avoids the concentration of population in province *i*, but may also prevent concentration of industrial activities.

[48] By dividing each estimated coefficient by the standard deviation of the respective variable we obtain $25\%(= 1.43/5.595 * 100)$ for the industry-to-agriculture ratio and $58\%(= 5.6/9.62 * 100)$ for the services-to-agriculture ratio.



(a) industry-to-agriculture ratio (b) services-to-agriculture ratio

Notes: authors' calculations on HIU and Italian Census data. The graphs plot coefficients of IV estimates of Eq. 2 for all faculties, and STEM faculties, using respective versions of Bartik instruments shown in Eq. 4, on the dependent variables reported in Panels (a)-(c). All the regressions include the usual set of controls (i.e. the total number of universities in province *i*, and its neighbourhood $-i$), region-by-year and province fixed effect. Details of all regressions are available to the authors upon request. Standard errors are clustered at the province level.

Fig. 8 Higher education supply and sectoral employment ratios

7 Welfare analysis

We are now ready to quantify the total contribution of higher education supply to the welfare of the local province, measured in per capita GVA terms. To do this, we have to take into full consideration how productivity effects of faculties in the neighbourhood complement the productivity effect of local faculties. Our analysis thus far has highlighted two opposite externalities. The first is measured by ϕ_2 in Eq. 2. This reflects a positive effect of the higher education supply of neighbouring provinces on local GVA per capita, which is well known in the literature. The second is shown by $b < 0$ in Eq. 1. This features a negative *displacement externality*, which the literature has never pointed out.

To measure the 'net' externality, we perform a simulation exercise using the value of coefficients from Eqs. 1 and 2. In practice, we compare the *actual returns* from higher education supply with the *counterfactual returns* in the absence of neighbourhood effects. For simplicity, we take as a reference the last year of our sample, i.e., $t = 2010$ (from now on we omit subscript *t*). The actual returns in province *i* relate to the number of faculties in province *i* and its neighbourhood, i.e.:

$$\widetilde{Y}_i = \widehat{\phi}_1 F_i + \widehat{\phi}_2 F_{-i}, \tag{5}$$

where F_i, F_{-i} are the observed number of faculties in province *i* and its neighbourhood ten years before, i.e. in 2001 (which is mandated by the time structure of our data), and $\widehat{\phi}_1$ and $\widehat{\phi}_2$ are the estimated productivity returns. If there were no faculties in the neighbourhood, province *i* would neither benefit from positive GVA externalities nor suffer from negative displacement effects from those faculties. Counterfactual returns in this scenario are as follows:

$$\widetilde{\widetilde{Y}}_i = \widehat{\phi}_1 \widehat{F}_i = \widehat{\phi}_1 (F_i - \widehat{b} \sum_{j \neq i} F_{ij}) = \widehat{\phi}_1 (F_i - \widehat{b} F_{-i}), \tag{6}$$

where $\widehat{b} < 0$ is the neighbourhood effect predicted from Eq. 1, and $\widehat{F}_i$ measures the counterfactual number of faculties in province *i* in the absence of neighbourhood effects. This is higher than observed supply, as relaxing neighbourhood pressures from contiguous provinces eliminates disincentives for local provision. Net externality Ext_i is the difference between Eqs. 5 and 6:

$$Ext_i = \widetilde{Y}_i - \widetilde{\widetilde{Y}}_i = (\widehat{\phi_2} + \widehat{\phi_1}\widehat{b})F_{-i}. \tag{7}$$

This is the sum between the positive value of direct welfare externalities $\widehat{\phi}_2 F_{-i}$ and the negative value of displacement externality $\widehat{\phi_1}\widehat{b}F_{-i}$.

Table 7 reports summary statistics of returns predicted by Eqs. 5 - 7 using $\widehat{b}$ and $\widehat{\phi}$s from our preferred specifications.[49] Based on the predicted values $\widehat{\phi_1}F_i = 1.401$ and $\widehat{\phi_2}F_{-i} = 2.993$ in Table 7, the local supply of faculties seems to contribute roughly half as much as the neighbouring supply of faculties. More importantly, once we take into account the opportunity cost of cross-provincial displacement forces $\widehat{\phi_1}\widehat{b}F_{-i}$, the net productivity contribution of local faculties ($\widetilde{\widetilde{Y}} = 2.815$) is more than one and half times larger than the corresponding contribution of faculties in the neighbourhood, on average ($Ext = 1.58$).

Some back-of-envelope calculations may be useful to assess magnitudes. Predictions in Table 7 suggest that average local supply alone in 2001 (equal to 5 faculties) explains about 64% of total productivity returns in the average province in 2010. These gains (compared to a benchmark situation with zero local faculties in 2001) are equivalent to the 13% of average (province level) GVA per capita in 2010. Net externalities (from an average supply of 23 faculties in the neighbourhood in 2001) explain the remaining 36% of total returns in 2010. This is equivalent to about 7% of average GVA per capita in 2010 (relative to having

Table 7 Simulation exercise: summary of welfare and productivity effects

[1] Variable	[2] Mean	[3] Std. Dev.	[4] Min.	[5] Max.	[6] Obs.
$\widehat{\phi_1}F_i$	1.401	2.001	0	11.82	107
$\widehat{\phi_2}F_{-i}$	2.993	1.832	0	8.586	107
$\widehat{\phi_1}\widehat{b}F_{-i}$	-1.414	0.904	-4.211	0	107
$\widetilde{Y}_i = \widehat{\phi_1}F_i + \widehat{\phi_2}F_{-i}$	4.394	2.523	0	13.771	107
$\widetilde{\widetilde{Y}}_i = \widehat{\phi_1}(F_i - \widehat{b}F_{-i})$	2.815	2.092	0	12.777	107
$Ext_i = (\widehat{\phi_2} + \widehat{\phi_1}\widehat{b})F_{-i}$	1.58	1.029	0	4.821	107
Annual effects (in % of $\overline{Y}$)					
$\widetilde{Y}_i/\overline{Y}$	2.092	1.214	0	7.080	107
$\widetilde{\widetilde{Y}}_i/\overline{Y}$	1.335	0.971	0	6.083	107
$Ext_i/\overline{Y}$	0.756	0.502	0	2.208	107
$\overline{Y}$ (mean in 2010)	21.518	5.198	12.413	32.571	107

Notes: Predicted effects from Eqs. (5)-(7) in the text. Values are computed using 2010 data. The annual effects are obtained dividing each ratio by 10 (since the effects computed in the upper panel refer to a decade) and multiplying by 100 (to express them in % terms)

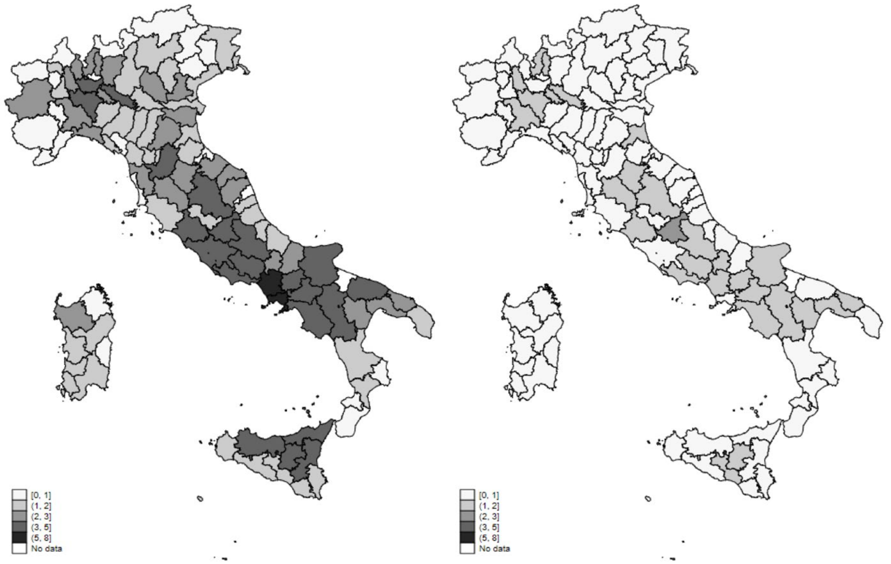
[49] These are IV FE estimates in Table 4, Panel A, Column [8] for Eq. 1 and Table 4, Panel B, Column [8] for Eq. 2.

zero faculties in the neighbourhood in 2001).[50] As these productivity gains accrue over a ten years period, their annual counterparts (reported in the lower rows of Table 7) are one tenth of the total effect, and are equal to respectively the 1.3% and 0.7% of average GVA per capita in 2010. This is a smaller total effect than the one we computed in Appendix A (0.20 vs. 0.27 in elasticity terms), where we evaluated $\widehat{\phi}_1$ and $\widehat{\phi}_2$ at sample averages.

In the Introduction we had stated that quantifying benefits and costs is necessary to understand the role of higher education supply for the persistent differentials in income per capita that have prevented regional economic convergence since the end of WWII (Gennaioli et al., 2013). With the help of Fig. 9 we move an important step in this direction. In Panel (a) we draw the cross-province distribution of the predicted total annual returns in $\%$ of local GVA per capita in 2010. The figure shows that economic returns tend to be negligible (less than 1% of local GVA per capita; transparent grey provinces) or small (between 1% and 2%, light grey areas) in peripheral provinces without own supply, that only benefit from faculties of a few small neighbours. Small provinces with own local faculties, or benefitting from externalities from major metropolitan areas enjoy returns of up to annual 3% of GVA per capita (grey provinces). Annual returns in large provinces such as Milano and Roma that benefit from local university hubs are between 3% and 5% of local GVA per capita (dark grey provinces). In the case of Napoli, annual returns explain over 7% of local GVA per capita (darkest grey province). Panel (b) shows the geographical distribution of the net annual externalities, also in % terms of GVA per capita. Everywhere in the Italian territory, externalities are non-negative, suggesting that predicted economic returns are larger than displacement costs in all provinces. However, externalities are small. In most border provinces (particularly in the northeast), net annual externalities account for less than 1% of local GVA per capita (transparent grey provinces). Around most urban areas, they explain between 1% and 2% of local GVA per capita (light grey areas).

This evidence already suggests that annual total returns in cities that benefit from a local supply, are likely larger than the corresponding positive externalities on neigbouring provinces. We can have a more precise picture, by looking at the complete distribution of annual returns in Table 13. Consider a province characterised by a small local supply (4 faculties) such as Bergamo. Based upon our estimates, this supply produces local annual returns equivalent to 1.35% of local GVA per capita. This is larger than the net externalities accruing to the neighbouring province of Sondrio (equal to 0.47% of local GVA per capita), which does not have a local supply. In practice, faculties in Bergamo contribute to creating a differential between Bergamo and Sondrio, two provinces whose levels of GVA per capita in 2010 are quite similar (28.5 for Bergamo and 28.2 in Sondrio). Evidence is more striking when we consider a larger university centre such as Roma. Based on our estimates, its large local supply (38 faculties) generates local annual returns equal to 4.31% of local GVA per capita. This is larger than annual total net externalities enjoyed by neighbouring provinces such as Latina (1.68% of local GVA per capita), or Rieti (2.21% of local GVA per capita, grey province in Fig. 9, Panel (b)), that do not have an own supply. In practice, these 38 faculties tend to widen the productivity gap observed in 2010 between Roma (GVA per capita

[50] Shares of local supply and supply in the neighbourhood over total returns are computed as $(\underset{\sim}{\widetilde{Y}} * 100/\widetilde{Y}) = (2.815 * 100/4.394) = 64\%$ and $(Ext * 100/\widetilde{Y}) = (1.58 * 100/4.394) = 36\%$. Productivity gains are quantified as follows: $(\widetilde{\widetilde{Y}}_i * 100/\overline{Y} = 2.815 * 100/21.518) \approx 13\%$, and $(Ext_i * 100/\overline{Y} = 1.58 * 100/21.518) \approx 7\%$.



(a) total returns, $\widetilde{Y}_i/Y_i$ (b) net externalities, Ext_i/Y_i

Notes: Values are expressed in % of local real GVA of province i (gross terms). The Yearly effects are obtained dividing each ratio by 10 (to obtain yearly effects from effects that refer to a decade) and multiplying by 100 (to express them as percentages). All values are computed using 2010 levels. Information on GVA is not available for the provinces of Barletta, Fermo, Monza-Brianza. Detailed values by province are shown in Table 13.

Fig. 9 Geographical distribution of predicted returns (annual effects, in %)

equal to 29.6) and its neighbours (19.8 in Rieti, and 20.1 in Latina). Overall, these results suggest that, if any, higher education supply contributes to widen differentials in GVA per capita between Italian provinces rather than fostering economic convergence.

8 Conclusions

This study leveraged a unique historical dataset to examine the long-term economic benefits of higher education supply at the province level in Italy. We constructed a comprehensive dataset tracing the establishment of universities from around the year 1000, allowing us to assess their impact on productivity over time. Our empirical analysis spanned 150 years (1870–2010), covering the period of unified Italy and capturing the transformation of higher education from an elite system to a more widespread and accessible one. This extensive time-frame enabled us to analyse the evolving role of universities from the early industrial era to the present. By integrating historical and economic perspectives, our study provided insights into how higher education contributed to economic development across different growth phases. Specifically, we examined the transition from the Post-Malthusian phase, characterised by gradual productivity improvements, to the Sustained Growth epoch, where human capital became a key driver of long-term economic expansion.

To identify the productivity effects of higher education supply, we exploited historical variations in the initial distribution of university faculties across different preunitarian periods. Our estimates indicated that the elasticity of productivity—measured as per capita Gross Value Added (GVA)—to higher education supply is approximately 0.2. By providing a more precise measurement of net externalities compared to previous studies, our study revealed that economic returns from university institutions are predominantly local. Specifically, we computed that 64% of annual returns are generated by higher education institutions located within the same province, while only 36% can be attributed to spillover effects from neighbouring provinces. We argued that such localised nature of university-driven economic benefits may contribute to persistent disparities in GVA per capita between wealthier provinces, that host large university hubs (the Joneses, following the metaphor we have cited in the title) and more peripheral ones, that mostly benefit from education externalities. From a policy perspective, this highlights the importance of aligning higher education supply with regional development needs. A more balanced spatial distribution of universities should aim at mitigating (or at least not exacerbating further) regional inequalities, when externalities that naturally spread economic gains beyond local boundaries are not strong enough.

Finally, a note of caution is needed. Our findings are particularly relevant for countries like Italy, where higher education is highly regulated, tuition fees are centrally set, and universities primarily compete locally. In contrast, factors like decentralised tuition (e.g., the UK) or a strong private sector (e.g., France) may alter these dynamics. Future research should examine how higher education supply shapes regional disparities across OECD countries, strengthening the argument that its expansion may have sustained rather than reduced economic inequalities since WWII (see Gennaioli et al., 2013).

Appendix A Evaluating the productivity effect of faculties

Let us define the elasticities $f_{i,i}$ and $f_{i,-i}$ of GVA per capita in province *i* to the supply of faculties in the same province *i*, and its neighbourhood $-i$, and evaluated at sample means. Recall that Eq. 2 considers a lag of ten years to measure the returns of faculty supply. By utilising the coefficients reported in Table 4 (IV FE estimates in Panel B, Column [8]), and evaluating the elasticities at the sample means reported in Table 2, we obtain the following results:

$$\begin{aligned} f_{i,i} =& \frac{dY_{i,t}}{dF_{i,t-10}} * \frac{\overline{F}_i}{\overline{Y}_i} = 0.311 * 2.6/9.09 = 0.09, \\ f_{i,-i} =& \frac{dY_{i,t}}{dF_{-i,t-10}} * \frac{\overline{F}_{-i}}{\overline{Y}_i} = 0.130 * 12.7/9.09 = 0.18. \end{aligned} \tag{A1}$$

These elasticities imply that a balanced 10% increase in the supply of faculties in province i and its neighbourhood produce a $0.9 + 1.8 = 2.7\%$ increase in local GVA per capita that develops over a time horizon of 10 years. This is equivalent to a yearly increase of 0.27%, on average. In Section 7 we find a smaller total elasticity (equal to 0.20 instead of 0.27), as we evaluate marginal effects at 2001 supply instead of the sample averages.

It is useful to understand under what conditions $f_{i,i}$, $f_{i,-i}$ are comparable to corresponding elasticities of GVA per capita to supply of universities $u_{i,i}$ (and $u_{i,-i}$):

$$u_{i,i} = \frac{dY_{i,t}}{dU_{i,t-1}} * \frac{\overline{U_i}}{\overline{Y}_i} \tag{A2}$$

We can decompose the first term in Eq. A-2 as

$$\frac{dY_{i,t}}{dU_{i,t-1}} = \frac{dY_{i,t}}{dF_{i,t-1}} * \frac{dF_{i,t-1}}{dU_{i,t-1}} \tag{A3}$$

which implies that the productivity effect of a marginal increase in the number of universities is equivalent to the productivity effect of opening up one faculty multiplied by the marginal increase in the number of faculties associated with opening up a university. A similar condition exists for $\frac{dY_{i,t}}{dU_{-i,t-1}}$. Let us assume that the second term is constant, and equal to the ratio between faculties and universities at their sample means i.e.

$$\frac{dF_{i,t-1}}{dU_{i,t-1}} = \frac{\overline{F}_i}{\overline{U}_i} \qquad and \qquad \frac{dF_{-i,t-1}}{dU_{-i,t-1}} = \frac{\overline{F}_{-i}}{\overline{U}_{-i}} \tag{A4}$$

Plug in Eq. A-4 into Eq. A-2 and obtain:

$$u_{i,i} = \left(\frac{dY_{i,t}}{dF_{i,t-1}} * \frac{\overline{F}_i}{\overline{U}_i} \right) * \frac{\overline{U_i}}{\overline{Y}_i} = f_{i,i} \tag{A5}$$

Similary it holds $u_{i,-i} = f_{i,-i}$. Thus, from Eq. A-4, GVA elasticities to faculty supply or university supply are equivalent as long as as opening up a new university does not have a differential productivity effect through the number of faculties opened.

Appendix B Additional figures and tables

Table 8 Summary Statistics on Initial Conditions in province *i*

[1] Variable	[2] Mean	[3] Std. Dev.	[4] Min.	[5] Max.	[6] Obs.
no. of faculties in *i* in 1400	0.4	0.973	0	4	1275
no. of faculties in *i* in 1500	0.482	1.07	0	4	1275
no. of faculties in *i* in 1600	0.529	1.123	0	4	1275
no. of faculties in *i* in 1700	0.6	1.171	0	4	1275
no. of faculties in *i* in 1800	0.682	1.295	0	5	1275
no. of faculties in *i* in 1860	1.071	1.952	0	8	1275
no. of STEM faculties in *i* 1400	0.165	0.482	0	2	1275
no. of STEM faculties in *i* 1500	0.2	0.527	0	2	1275
no. of STEM faculties in *i* 1600	0.212	0.555	0	2	1275
no. of STEM faculties in *i* 1700	0.259	0.597	0	2	1275
no. of STEM faculties in *i* 1800	0.318	0.673	0	3	1275
no. of STEM faculties in *i* in 1860	0.682	1.304	0	6	1275
no. of non-STEM faculties in *i* in 1400	0.235	0.567	0	2	1275
no. of non-STEM faculties in *i* in 1500	0.282	0.625	0	2	1275
no. of non-STEM faculties in *i* in 1600	0.318	0.655	0	2	1275
no. of non-STEM faculties in *i* in 1700	0.341	0.661	0	2	1275
no. of non-STEM faculties in *i* in 1800	0.365	0.701	0	2	1275
no. of non-STEM faculties in *i* in 1861	0.388	0.705	0	2	1275

Notes: Authors' calculation based on HIU data

Table 9 Sensitivity Analysis (1): alternative t_0s

	[1] $t_0 = 1800$	[2] $t_0 = 1600$	[3] $t_0 = 1500$	[4] $t_0 = 1400$
	Panel A: Neighbourhood effects			
Total no. faculties in *j*	-0.193**	-0.198**	-0.188**	-0.193*
	(0.079)	(0.092)	(0.096)	(0.100)
Observations	3868	3868	3868	3868
K-P rk Wald F-stat	53.197	31.220	31.772	25.095
K-P rk LM-stat	29.768	21.041	20.360	16.814
p-value KP	0.000	0.000	0.000	0.000
	Panel B: Productivity			
Total no. faculties in *i*	0.309**	0.272*	0.258	0.299
	(0.142)	(0.152)	(0.166)	(0.205)
Total no. faculties in $-i$	0.117**	0.119**	0.128**	0.122**
	(0.050)	(0.050)	(0.054)	(0.057)
Observations	1025	1025	1025	1025
K-P rk Wald F-stat	34.319	31.981	28.660	23.011
K-P rk LM-stat	16.781	13.953	11.883	9.484
p-value KP	0.000	0.000	0.001	0.002

Notes: We take as preferred specifications IV FE one used to produce results reported in Table 4, Column [8] (Panel A for Neighbourhood effects; Panel B for productivity effects). In Columns [1] to [4] we report IV FE results as we compute initial conditions at t_0s reported in each column. Significance levels: $*$: 10% $**$: 5% $***$: 1%

Table 10 Sensitivity Analysis (2): spatial correlation and clustering

	[1] Conley correction	[2] DKRAAY correction	[3] Two-way clustering	[4] Pair-level clustering	[5] Dyadic clustering	[6] Region-level clustering	[7] Kingdom clustering
	Panel A: Neighbourhood effects						
Total no. faculties in j	-0.205***	-0.205***	-0.205***	-0.205***	-0.205***	-0.205***	-0.205***
	(0.035)	(0.016)	(0.073)	(0.036)	(0.063)	(0.064)	(0.044)
Observations	3868	3868	3868	3868	3868	3868	3868
K-P rk Wald F-stat	32.730	37.909	15.574	45.208	–	16.807	12.981
	Panel B: Productivity						
Total no. faculties in i	0.311***	0.311***	–	–	–	0.311**	0.311***
	(0.111)	(0.117)				(0.137)	(0.076)
Total no. faculties in $-i$	0.130**	0.130***	–	–	–	0.130**	0.130***
	(0.062)	(0.042)				(0.055)	(0.026)
Observations	1025	1025	–	–	–	1025	1025
K-P rk Wald F-stat	42.496	18.310	–	–	–	22.718	43.774

Notes: We take as preferred specifications IV FE one used to produce results reported in Table 4, Column [8] (Panel A for Neighbourhood effects; Panel B for productivity effects). In Columns [1] and [2] we implement (Driscoll and Kraay, 1998)'s corrections, respectively to account for spatial auto-correlation. In Column [3] we compute two-way clustering following the methodology by Gu (2019). In Column [4] we cluster the standard errors at the level of the pair. In Column [5] we report DCRSE (dyadic clustered robust standard errors), using the procedure proposed by Aronow et al. (2015). For Columns from [3] to [5] the clustering procedure is applicable only to the pairwise model. For Column [5], the dcr routine in Stata does not provide K-P statistics on instruments validity. In Column [6] we cluster standard errors at the regional level. In Column [7] we cluster standard errors at the level of the pre-unitarian state. Significance levels: $*$: 10% $**$: 5% $***$: 1%

Table 11 Sensitivity Analysis (3): varying the set of control variables

	[1] Pop Growth controls	[2] Control for %s in agriculture	[3] Elite & Private Uni. Controls	[4] (Notability)*(no. Univ.) Controls	[5] (Ruggedness)*(Year FE)	[6] Drop all controls
	Panel A: Neighbourhood effects					
Total no. faculties in j	-0.207**	-0.223**	-0.212**	-0.205**	-0.232**	-0.197
	(0.094)	(0.099)	(0.091)	(0.100)	(0.098)	(0.122)
Observations	3868	3868	3868	3859	3868	3868
K-P rk Wald F-stat	29.716	29.178	29.969	31.905	28.768	30.304
	Panel B: Productivity effects					
Total no. faculties in i	0.310**	0.392**	0.310**	0.418**	0.307**	0.342**
	(0.152)	(0.192)	(0.142)	(0.189)	(0.149)	(0.159)
Total no. faculties in $-i$	0.129**	0.133**	0.111**	0.120**	0.160***	0.136**
	(0.054)	(0.062)	(0.049)	(0.054)	(0.054)	(0.063)
Observations	1025	1025	1025	1025	1025	1025
K-P rk Wald F-stat	33.961	31.848	35.296	27.243	33.421	9.329

Notes: We take as preferred specifications IV FE one used to produce results reported in Table 4, Column [8] (Panel A for Neighbourhood effects; Panel B for productivity effects). In Column [1] we control for the rate of growth of population in province i, and its neighbourhood. In Column [2] we include in the set of controls the shares of active people in the agricultural sector. In Column [3], we include in the set of controls indicators for the number of elite universities and the number of private universities, both for the local province and its neighbourhood. In Column [4], we include in the set of controls the interaction between the number of universities in each province and the corresponding indicator of academic notability (noto7) by De la Croix et al. (2024). In Column [5], we include interactions between time dummies and a dummy for local province having an altitude higher than 350 metres, on average. In Column [6], we drop controls for the total number of universities (in the province and the neighbourhood). Coefficients of all additional control variables used in this table are available to the authors upon request. In Columns [1]-[6], standard errors clustered at the level of the local province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%

Table 12 Mechanism analysis: altitude and transportation

	Panel A: Railway connections in 1860		*Panel B: Low vs. high altitude (sea level)*	
	[1] No connection	[2] Some connection	[3] Altitude below 350m	[4] Altitude above 350m
no. of faculties in *j*	-0.160	-0.223*	-0.335**	-0.074
	(0.109)	(0.136)	(0.167)	(0.085)
Observations	1871	1997	2046	1822
K-P rk Wald F-stat	16.980	12.336	9.402	15.674

Notes: In Panel A, we split the sample between provinces with or without railway connections in 1860. In Panel B we split the sample between provinces characterised by a standard deviation in the altitude above and below the sample median (350 metres above sea level)

Table 13 Welfare analysis of higher education supply

				Yearly effects (in %)			
Province	F_{-i}	F_i	$\widehat{F_i}$	$\widetilde{Y_i}/Y_i$	$\widetilde{\widetilde{Y}}_i/Y_i$	Ext_i/Y_i	Y_i
Agrigento	11.00	0.00	2.26	1.04	0.51	0.53	13.76
Alessandria	37.00	3.00	10.59	2.37	1.36	1.01	24.24
Ancona	29.00	5.00	10.95	2.10	1.34	0.76	25.36
Arezzo	45.00	1.00	10.23	2.53	1.31	1.23	24.34
Ascoli Piceno	25.00	1.00	6.13	1.74	0.93	0.81	20.46
Asti	17.00	0.00	3.49	0.95	0.46	0.48	23.39
Avellino	44.00	0.00	9.02	3.77	1.85	1.92	15.17
Bari	10.00	16.00	17.03	4.09	3.45	0.64	15.36
Belluno	18.00	0.00	3.69	0.89	0.43	0.45	26.40
Benevento	40.00	3.00	11.20	4.22	2.40	1.82	14.54
Bergamo	41.00	4.00	12.41	2.31	1.35	0.95	28.53
Biella	17.00	0.00	3.49	0.92	0.45	0.47	24.14
Bologna	26.00	15.00	20.33	2.70	2.12	0.58	29.79
Bolzano/Bozen	6.00	2.00	3.23	0.46	0.33	0.13	30.60
Brescia	18.00	6.00	9.69	1.57	1.13	0.45	26.79
Brindisi	25.00	0.00	5.13	2.32	1.14	1.18	14.04
Cagliari	0.00	10.00	10.00	1.60	1.60	0.00	19.50
Caltanissetta	22.00	0.00	4.51	1.99	0.98	1.02	14.35
Campobasso	22.00	3.00	7.51	2.08	1.28	0.80	18.27
Carbonia-Iglesia	10.00	0.00	2.05	1.01	0.49	0.51	12.94
Caserta	34.00	9.00	15.97	5.39	3.71	1.68	13.41
Catania	12.00	11.00	13.46	3.37	2.84	0.54	14.77
Catanzaro	9.00	3.00	4.85	1.24	0.89	0.35	16.95
Chieti	13.00	4.00	6.67	1.60	1.13	0.47	18.31
Como	40.00	2.00	3.23	2.40	0.41	1.99	24.23
Cosenza	7.00	6.00	7.44	1.81	1.51	0.30	15.31
Cremona	57.00	1.00	12.69	3.10	1.58	1.52	24.93
Crotone	9.00	0.00	1.85	0.94	0.46	0.48	12.41

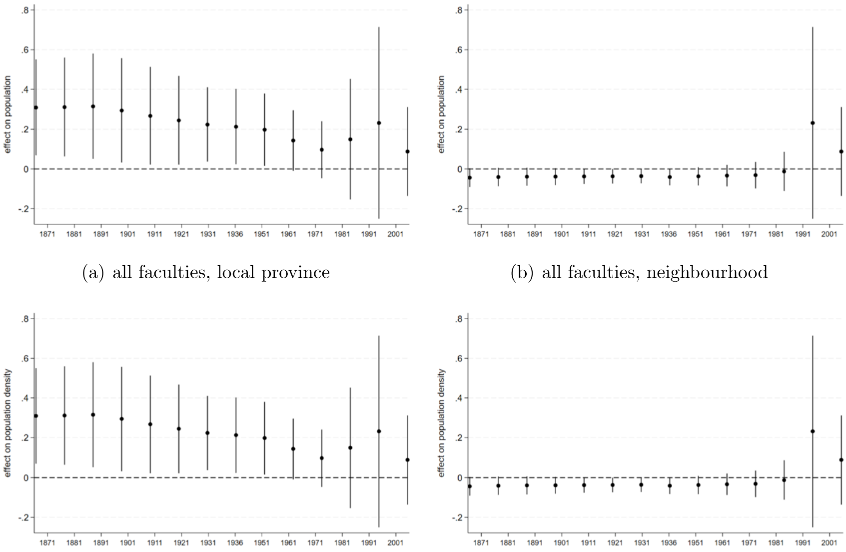
Table 13 (continued)

Province	F_{-i}	F_i	$\widehat{F}_i$	Yearly effects (in %) $\widetilde{Y}_i/Y_i$	$\widetilde{\widetilde{Y}}_i/Y_i$	Ext_i/Y_i	Y_i
Cuneo	14.00	0.00	2.87	0.67	0.33	0.34	27.04
Enna	33.00	0.00	6.77	3.11	1.53	1.59	13.79
Ferrara	22.00	8.00	12.51	2.30	1.67	0.63	23.26
Firenze	40.00	11.00	19.20	3.12	2.16	0.96	27.62
Foggia	26.00	5.00	7.05	3.56	1.58	1.98	13.85
Forli-Cesena	13.00	5.00	7.67	1.17	0.86	0.31	27.76
Frosinone	54.00	4.00	15.08	3.89	2.21	1.68	21.26
Genova	16.00	11.00	14.28	2.21	1.79	0.43	24.86
Gorizia	22.00	0.00	4.51	1.24	0.61	0.63	23.04
Grosseto	24.00	0.00	4.92	1.33	0.65	0.68	23.42
Imperia	0.00	0.00	0.00	0.00	0.00	0.00	22.92
Isernia	26.00	1.00	6.33	2.17	1.16	1.01	17.05
L'Aquila	55.00	6.00	17.28	4.89	2.92	1.98	18.43
La Spezia	21.00	0.00	4.31	1.20	0.59	0.61	22.86
Latina	51.00	0.00	10.46	3.30	1.62	1.68	20.09
Lecce	1.00	8.00	8.21	1.78	1.73	0.05	14.72
Lecco	40.00	0.00	1.23	2.05	0.15	1.90	25.35
Livorno	11.00	0.00	2.26	0.61	0.30	0.31	23.42
Lodi	46.00	0.00	9.44	2.56	1.26	1.31	23.36
Lucca	32.00	0.00	6.56	1.60	0.78	0.82	26.03
Macerata	17.00	9.00	12.49	2.27	1.76	0.51	22.05
Mantova	42.00	0.00	8.61	1.94	0.95	0.99	28.20
Massa-Carrara	13.00	0.00	2.67	0.81	0.40	0.41	20.80
Matera	27.00	0.00	5.54	2.11	1.04	1.08	16.62
Medio Campidano	10.00	0.00	2.05	1.01	0.49	0.51	12.93
Messina	22.00	11.00	15.51	4.09	3.14	0.95	15.38
Milano	24.00	34.00	38.51	4.21	3.68	0.53	32.57
Modena	26.00	7.00	12.33	1.93	1.33	0.60	28.74
Napoli	21.00	23.00	27.31	7.08	6.08	1.00	13.96
Novara	49.00	3.00	13.05	2.93	1.63	1.30	24.95
Nuoro	21.00	0.00	4.31	1.59	0.78	0.81	17.18
Ogliastra	10.00	0.00	2.05	0.82	0.40	0.42	15.80
Olbia-Tempio	11.00	0.00	2.26	0.71	0.35	0.36	20.04
Oristano	21.00	0.00	4.31	1.76	0.86	0.90	15.54
Padova	12.00	13.00	15.46	2.16	1.85	0.31	25.96
Palermo	11.00	11.00	13.26	3.16	2.68	0.47	15.36
Parma	18.00	10.00	13.69	1.98	1.55	0.43	27.53
Pavia	44.00	8.00	17.02	3.47	2.24	1.23	23.66
Perugia	33.00	11.00	17.77	3.60	2.58	1.02	21.43
Pesaro e Urbino	18.00	9.00	12.69	2.18	1.67	0.51	23.60
Pescara	15.00	3.00	6.08	1.55	1.01	0.53	18.66
Piacenza	33.00	3.00	9.77	1.95	1.13	0.81	26.85
Pisa	19.00	11.00	14.90	2.37	1.86	0.51	24.85

Table 13 (continued)

				Yearly effects (in %)			
Province	F_{-i}	F_i	$\widehat{F_i}$	$\widetilde{Y}_i/Y_i$	$\widetilde{\widetilde{Y}}_i/Y_i$	Ext_i/Y_i	Y_i
Pistoia	33.00	0.00	6.77	1.89	0.93	0.96	22.68
Pordenone	15.00	0.00	3.08	0.79	0.39	0.40	24.58
Potenza	36.00	4.00	11.38	3.48	2.08	1.40	17.05
Prato	26.00	0.00	5.33	1.35	0.66	0.69	24.98
Ragusa	12.00	0.00	2.46	1.04	0.51	0.53	14.97
Ravenna	39.00	0.00	8.00	2.00	0.98	1.02	25.40
Reggio Calabria	3.00	3.00	3.62	0.92	0.78	0.14	14.38
Reggio Emilia	17.00	3.00	6.49	1.22	0.78	0.44	25.79
Rieti	66.00	0.00	13.54	4.33	2.12	2.21	19.82
Rimini	15.00	1.00	4.08	0.83	0.47	0.37	27.11
Roma	15.00	38.00	41.08	4.65	4.31	0.34	29.64
Rovigo	33.00	0.00	6.77	1.70	0.83	0.87	25.26
Salerno	27.00	9.00	14.54	3.93	2.81	1.11	16.07
Sassari	0.00	11.00	11.00	2.06	2.06	0.00	16.63
Savona	14.00	0.00	2.87	0.77	0.38	0.39	23.77
Siena	39.00	8.00	16.00	2.98	1.96	1.02	25.40
Siracusa	11.00	1.00	3.26	1.12	0.65	0.47	15.59
Sondrio	20.00	0.00	4.10	0.92	0.45	0.47	28.23
Taranto	24.00	1.00	5.92	2.29	1.23	1.06	14.97
Teramo	10.00	5.00	7.05	1.58	1.21	0.37	18.08
Terni	24.00	0.00	4.92	1.54	0.76	0.79	20.26
Torino	6.00	14.00	15.23	2.10	1.94	0.16	24.40
Trapani	11.00	0.00	2.26	1.02	0.50	0.52	14.02
Trento	15.00	6.00	9.08	1.40	1.04	0.36	27.26
Treviso	18.00	0.00	3.69	0.93	0.46	0.48	25.07
Trieste	0.00	12.00	12.00	1.31	1.31	0.00	28.48
Udine	5.00	10.00	11.03	1.49	1.35	0.13	25.31
Aosta	15.00	2.00	5.08	0.98	0.60	0.38	26.32
Varese	39.00	6.00	14.00	2.67	1.68	1.00	25.97
Venezia	23.00	5.00	9.72	1.71	1.14	0.57	26.54
Verbano-Cusio-Ossola	10.00	0.00	2.05	0.65	0.32	0.33	20.03
Vercelli	30.00	1.00	7.15	1.66	0.88	0.78	25.39
Verona	25.00	7.00	12.13	2.04	1.42	0.62	26.65
Vibo Valentia	6.00	0.00	1.23	0.57	0.28	0.29	13.59
Vicenza	26.00	0.00	5.33	1.25	0.61	0.64	27.01
Viterbo	46.00	5.00	14.44	3.65	2.18	1.48	20.64
Total	23.01	4.50	9.05	2.09	1.34	0.76	21.52

Notes: All values refer to 2010 levels. Information on Y_i is not available for provinces of Barletta, Fermo and Monza-Brianza. The Yearly effects are obtained dividing each ratio by 10 (since the effects computed in the upper panel refer to a decade) and multiplying by 100 (to express them as percentages)



(a) all faculties, local province

(b) all faculties, neighbourhood

(c) all faculties, local province

(d) all faculties, neighbourhood

Notes: authors' calculations on HIU and Italian Census data. The graphs plot the coefficients and 90% confidence intervals of the total no of faculties in the local province and its neighbourhood on dependent variables reported on the y axis for the sub-sample starting on the year reported on the x-axis (i.e. dropping from the sample all previous years). All the regressions include the usual set of controls (i.e. the total number of universities in province i, and its neighbourhood $-i$), region-by-year and province fixed effect. Details of all regressions are available to the authors upon request. Standard errors are clustered at the province level).

Fig. 10 Effects of higher education supply on population and population density

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10887-025-09263-5.

Acknowledgements We appreciate the constructive comments and suggestions provided by three anonymous referees and the Editor. We are grateful to David De la Croix for generously sharing his data on academic notability. We also thank Anna Valero, Thomas Baudin, Giovanni Peri, and Nico Voigtländer for their valuable suggestions. We are especially indebted to Dirk Foremny and Marine de Talance for their insightful discussions. Finally, we acknowledge the Istituto Tagliacarne for supplying data on provincial income. This paper was presented at the 2017 LEER Workshop in Leuven; 2017 IWAAE Workshop in Catanzaro; 2017 APET Conference in Paris; 2017 AIEL Conference in Cosenza; 2018 ESCOS International Conference at University of Napoli; 2018 JMA Workshop in Bordeaux; 2018 Annual Conference of the French Economic Association in Paris; 2018 IZA/SOLE Transatlantic Meeting of Labor Economists in Munich; 2018 CESifo Public Sector Economics Area Conference in Munich; 2018 EEA Conference in Cologne; 2018 CREA Workshop on Culture and Comparative Development; and 2018 Second Winter Symposium in Economics and Finance. It was also presented at seminars at Catholic University of Milan and at invited seminars at the School of Economics of the University of Barcelona, the Italian Agency of Research and University (ANVUR), CREA University of Luxembourg, and Lille Economics and Management (LEM). We wish to thank all colleagues who attended these presentations. We are particularly grateful to Pietro De Ponti, Marika Fasola, Ginevra Gallassi, Giuseppe Grasso, George Jacob, Purushottam Mohanty, Rodrigo Rocha and Roberta Terranova for providing assistance at various stages of our research. The usual disclaimer applies.

Author Contributions All authors contributed to each stage of the study. All authors read and approved the final manuscript.

Funding Open access funding provided by Università degli Studi del Piemonte Orientale Amedeo Avogrado within the CRUI-CARE Agreement. This paper is part of the project “Migration And Labor supplY wheN culturE matterS", financed by French National Research Agency (ANR-18-CE26-000, AAPG 2018). We acknowledge ANR for financial support. Finally, we thank the Department of Economics and Finance, Università Cattolica del Sacro Cuore for selecting this project for the Research Assistance Program. The authors have no relevant financial or non-financial interests to disclose.

References

Adão, R., Kolesár, M., Morales, E., & Zhang, X. (2020). REG_SS: Stata module to compute confidence intervals, standard errors, and p-values in a linear regression in which the regressor of interest has a shift-share structure. Boston College Department of Economics: Statistical Software Components.

Agasisti, T. and Catalano, G. (2007). I quasi-mercati nell’istruzione universitaria. Un modello interpretativo per un confronto Europeo. Bartezzagli E., Raffa M., Zollo G.(eds.), Produzione e trasferimento di conoscenze. *Edizioni Scientifiche Italiane*.

Alvazzi del Frate, P. (2021). Les universités napoléoniennes en Italie . *Cahiers Jean Moulin*, (7).

Andersson, R., Quigley, J. M., & Wilhelmson, M. (2004). University decentralization as regional policy: the swedish experiment. *Journal of Economic Geography, 4*(4), 371–388.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton University Press.

Aronow, P. M., Samii, C., & Assenova, V. A. (2015). Cluster–robust variance estimation for dyadic data. *Political Analysis, 23*(4), 564–577.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics, 106*(2), 407–443.

Barro, R. J. (2012). Convergence and modernization revisited. Working Paper 18295, National Bureau of Economic Research.

Becker, S. O., Hornung, E., & Woessmann, L. (2011). Education and catch-up in the industrial revolution. *American Economic Journal: Macroeconomics, 3*(3), 92–126.

Bidolli, A. and Soldani, S. (2001). L’istruzione agraria in Italia. Fonti per la storia della scuola, Ministero per i beni e le attivita culturali, Direzione Generale per gli Archivi 2001.

Bratti, M. and Leombruni, R. (2014). Local human capital externalities and wages at the firm level: Evidence from Italian manufacturing. *Economics of Education Review*, *41*(C), 161–175.

Brizzi, G.P. Del Negro, P. and Romano, A. (2007). Storia delle Università in Italia. CISUI.

Cabassi, E. and Liuzzo, G. (2001). L’insegnamento medico-veterinario a Parma. *Graphital Edizioni*.

Cantoni, D., & Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics, 129*(2), 823–887.

Carlson, J., Incerti, T., & Aronow, P. M. (2024). Dyadic clustering in international relations. *Political Analysis, 32*(2), 186–198.

Catalano, G. and Silvestri, P. (1999). Regolamentazione e competizione nel sistema universitario italiano: effetti e problemi del nuovo sistema di finanziamento. D. Fabbri e G. Fiorentini (eds.), *Regolamentazione dei servizi pubblici*. Carocci.

Ciccarelli, C., & Fenoltea, S. (2013). Through the magnifying glass: provincial aspects of industrial growth in post-unification italy. *The Economic History Review, 66*(1), 57–85.

Ciccone, A., & Peri, G. (2006). Identifying human-capital externalities: theory with applications. *Review of Economic Studies, 73*(2), 381–412.

Ciccone, A. and Peri, G. (2011). Schooling Supply and the Structure of Production: Evidence from US States 1950-1990. NBER Working Papers 17683, National Bureau of Economic Research, Inc.

Clark, W. (2006). *Academic Charisma and the Origins of the Research University*. Chicago and London: University of Chicago Press.

Curtis, M. and de la Croix, D. (2023). Seeds of knowledge: premodern scholarship, academic fields, and european growth. WorkingPaper 18321, CEPR Press.

Giorgi, G., Pellizzari, M., & Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics, 2*(2), 241–75.

Croix, D., Docquier, F., Fabre, A., & Stelter, R. (2024). The academic market and the rise of universities in medieval and early modern europe (1000–1800). *Journal of the European Economic Association, 22*(4), 1541–1589.

de Ridder-Symoens, H. (1992). History of the European university, Vol. 1: Universities in the Middle Ages. Cambridge University Press.

DeAngelis, I., Mariani, V., & Torrini, R. (2017). New evidence on interregional mobility of students in tertiary education: the case of italy. *Politica Economica, 1*, 73–96.

Dittmar, J. and Meisenzahl, R. R. (2022). The research university, invention and industry: evidence from German history. CEP Discussion Papers dp1856, Centre for Economic Performance, LSE.

Driscoll, J., & Kraay, A. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics, 80*(4), 549–560.

Eurydice (2000). *Two Decades of Reform in Higher Education in Europe: 1980 Onwards*. Eurydice European Unit: Eurydice studies.

Fabrizzi, F. (2008a). La provincia: storia istituzionale dell'ente locale più discusso. dalla riforma di crispi all'assemblea costituente. Technical Report 12/2008, federalismi.it.

Fabrizzi, F. (2008b). La provincia: storia istituzionale dell'ente locale più discusso. dall'assemblea costituente ad oggi. Technical Report 23/2008, federalismi.it.

Fabrizzi, F. (2008c). La provincia: storia istituzionale dell'ente locale più discusso. dall'unificazione amministrativa alla caduta della destra storica. Technical Report 7/2008, federalismi.it.

Fafchamps, M., & Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics, 83*(2), 326–350.

Federico, G., Nuvolari, A., & Vasta, M. (2019). The origins of the italian regional divide: evidence from real wages, 1861–1913. *The Journal of Economic History, 79*(1), 63–98.

Fois, G. (1991). L'Università di Sassari nell'Italia liberale. Università di Sassari.

Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.

Galor, O., & Moav, O. (2004). From physical to human capital accumulation: inequality and the process of development. *The Review of Economic Studies, 71*(4), 1001–1026.

Galor, O., & Moav, O. (2006). Das human-kapital: a theory of the demise of the class structure. *The Review of Economic Studies, 73*(1), 85–117.

Galor, O., & Özak, O. (2016). The agricultural origins of time preference. *American Economic Review, 106*(10), 3064–3103.

Galor, O., & Weil, D. N. (2000). Population, technology, and growth: from malthusian stagnation to the demographic transition and beyond. *The American Economic Review, 90*(4), 806–828.

Gennaioli, N., Porta, R. L., Silanes, F. L., & Shleifer, A. (2013). Human capital and regional development. *The Quarterly Journal of Economics, 128*(1), 105–164.

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: what, when, why, and how. *American Economic Review, 110*(8), 2586–2624.

Gray, R., Narciso, G., & Tortorici, G. (2019). Globalization, agricultural markets and mass migration: Italy, 1881–1912. *Explorations in Economic History, 74* .

Gu, A. (2019). vcemway: A one-stop solution for robust inference with multiway clustering. *Stata Journal*, *19*(4), 900–912(13).

Hanushek, E. A. and Woessmann, L. (2015). The Knowledge Capital of Nations: Education and the Economics of Growth. MIT Press.

Hoxby, C. (2016). The dramatic economics of the U.S. market for higher education. The 8th annual Martin Feldstein lecture, National Bureau of Economic Research.

Kickert, W. (2007). Public management reforms in countries with a napoleonic state model: France, Italy and Spain, pp. 26–51. Palgrave Macmillan UK, London.

Manski, C. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies, 60*(3), 531–542.

Mattone, A. (2010). Storia dell'Università di Sassari. *vol. 1*, in Storia dell'Università di Sassari. Ilisso.

Mele, C. (2006). Origini e formazione del Politecnico di Torino: dagli esordi alla sua formazione l'8 luglio 1906. In: Storia dell'Ingegneria, (Atti del Primo Convegno Nazionale) . Cuzzolin Editore.

Mokyr, J. (2005). Long-term economic growth and the history of technology. In P. Aghion & S. Durlauf (Eds.), *Handbook of Economic Growth, volume 1 , chapter 17,* pp. 1113–1180. Elsevier.

Momand, A. R. (1913). Keeping Up with the Joneses. New York World. Comic strip first published in 1913 in the New York World newspaper.

OECD (2007). *Higher Education and Regions: Globally Competitive*. OECD Paris: Locally Engaged. Final Report.

OECD (2014). The State of Higher Education 2014. A. Glass, (eds.). The OECD Higher Education Programme (IMHE). OECD Paris.

Parchet, R. (2019). Are local tax rates strategic complements or strategic substitutes? *American Economic Journal: Economic Policy, 11*(2), 189–224.

Pepe, L. (2007). Storia dell'Università di Ferrara. *Storia delle università italiane, vol. III*, G. P. Brizzi, P. del Negro, A. Romano, eds. CISUI, Sicania.

Petracchi, A. (1962). *Le Origini dell' Ordinamento comunale e provinciale italiano: storia della legislazione piemontese sugli enti locali dalla fine dell'antico regime al chiudersi dell' età cavouriana (1770–1861)*. Pozza: Pubblicazioni dell'Istituto per la scienza dell'amministrazione pubblica. Edizioni N.

Roberts, J., Rodriguez-Cruz, A., and Herbst, J. (1996). Exporting Models. In de Ridder-Symoens, H., editor, A History of the University in Europe. Vol. II: Universities in Early Modern Europe (1500–1800), pp. 256–284. Cambridge University Press.

Schofer, E., & Meyer, J. W. (2005). The worldwide expansion of higher education in the twentieth century. *American Sociological Review, 70*(6), 898–920.

Silvestri, A. (2006). La nascita delle Facoltà di Ingegneria e di Architettura in Italia. Cuzzolin Editore.

Smith, W. (2010). *Public Universities and the Democratization of the Core Public Sphere*. Palgrave Macmillan, New York: Public Universities and the Public Sphere.

Squicciarini, M. P., & Voigtländer, N. (2015). Human capital and industrialization: evidence from the age of enlightenment. *The Quarterly Journal of Economics, 130*(4), 1825–1883.

Valero, A. and Van Reenen, J. (2019). The economic impact of universities: evidence from across the globe. *Economics of Education Review*, *68*(C), 53–67.