Higher Education Supply, Neighbourhood Effects and Economic Welfare

Elena Cottini, Paolo Ghinetti, Simone Moriconi

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

Higher Education Supply, Neighbourhood Effects and Economic Welfare

Abstract

This paper uses our own built historical data for Italy to analyse the effects of the modern supply of higher education institutions on economic welfare, measured in terms of real income per capita. We perform an empirical analysis at the province level implementing a fixed effect instrumental variables estimator for over 150 years. We use as instruments the interaction of the province-level supply of higher education before the Italian political unification with university reforms that occurred in Italian history. We find evidence of displacement forces between institutions in sufficiently close neighbourhoods. We argue that quantifying the cost of such displacement forces is crucial to properly evaluating the welfare effects of higher education institutions. We find that a province exposed to the average supply (measured as the number of local institutions) has a 4.9% higher real income per capita compared to a province with no supply. Once we properly account for the foregone value of displacement forces, we show that economic returns are very localised at the province level.

JEL-Codes: I230, I280, N000, R100.

Keywords: neighbourhood effects, higher education supply, historical data, initial conditions, economic welfare.

Elena Cottini
Department of Economics and Finance
Università Cattolica del Sacro Cuore
Largo Gemelli 1
Italy - 20123 Milano (MI)
elena.cottini@unicatt.it

Paolo Ghinetti
Dipartimento di Studi per l'Economia e
l'Impresa, Università del Piemonte Orientale
Via Perrone 18
Italy – Novara (NO)
paolo.ghinetti@uniupo.it

Simone Moriconi
Department of Economics and Quantitative Methods
Iéseg School of Management and LEM
1 parvis de La Défense
France – 92044 Paris – La Défense Cedex
s.moriconi@ieseg.fr

1 Introduction

The development of modern systems of higher education dates back to the outset of the industrial revolution, as increased support for public education generated a deep transformation of upper-tail human capital. At the time, the modern industrial sector needed considerable and diffused “knowledge capital” that would contribute to long-run economic growth (Hanushek and Woessmann, 2015). Starting from the second half of the IXX^{th} century, rising productivity resulting from the spread of industry would call for a transition from the élite knowledge of the preindustrial era to widespread higher education favouring mass schooling in Western societies (Madrick, 2002).

Not much is known about the welfare effects of the establishment of modern higher education institutions in developed countries. It is generally believed that economic benefits are associated with the creation of large stocks of human capital. These stocks are local, as workers and students have limited geographical mobility and produce localised benefits, similar to the economic advantages of large plant openings (Moretti, 2004, 2011; Greenstone et al., 2010; Hoxby, 2016). This regional perspective supported the engagement by OECD countries to tailor the expansion of their higher education supply to regional growth throughout the XX^{th} century (OECD, 2007; OECD, 2014). However, it disregards the view that the economic benefits of higher education institutions may go beyond regional boundaries. This is also true for economic costs. Neighbouring regions insist on the same scarce resources (e.g., students and investments), so opening a higher education institution in a specific location reduces the possibility of opening competing institutions nearby. Overall, quantifying actual benefits and costs is a crucial task to evaluating the welfare effects of higher education institutions. From a policy perspective, it is a first necessary step to understand whether the rise of the modern supply of higher education may have paradoxically contributed to the persistent differentials in income per capita that have prevented regional economic convergence since the end of WWII (Gennaioli et al., 2013).

This paper analyses the welfare effects of the rise of the modern supply of higher education in Italy. Italy has the longest history of university education in the western world. We offer a detailed historical account of the constant expansion of higher education supply in Italy, starting from the establishment of élite knowledge institutions in the Italian territory during the Middle Ages. Importantly, we evaluate the effects of the rise of higher education supply on welfare. We conduct an empirical analysis at the subregional level, namely, the province, which is equivalent to the *département* in France or the county in the US or UK. This is the oldest governance level in the Italian territory and historically responsible for the provision of public goods (Federico et al., 2019). We perform our welfare evaluation using the

predicted coefficients from a twofold empirical exercise. First, we estimate neighbourhood effects, i.e., the association between the higher education supply in nearby provinces and the local supply. Second, we estimate province-level returns of local higher education supply and of the supply in neighbourhoods in terms of real income per capita. We use estimated coefficients to predict the actual distribution of economic returns and the counterfactual distribution in the absence of neighbourhood effects. By comparing the two distributions, we are able to separate local economic returns from the economic externalities of neighbouring provinces.

For these purposes, we constructed an original historical dataset covering the History of Italian Universities (HIU). We use as a measure of higher education supply the total number of “faculties” in a province in a given year. In the Middle Ages, the faculty was the lower-level higher education institution in charge of teaching provisions (including the design of curricula) in a given field of study (De Ridder-Symoens and Rüegg, 1992). The faculty maintained the same scope in the Italian higher education system until 2010, when a major reform radically changed the governance of universities.¹

Overall, HIU data cover 110 Italian provinces for between 1861 and 2010. The cross-province panel structure features a long time span and a fine geographical disaggregation. This makes the fixed effect (FE) estimator the natural econometric benchmark for our analysis. We also implement an instrumental variable (IV) approach to address remaining concerns related to reverse causality and time-varying omitted variable bias at the province level. We construct instruments *à la* Bartik as the interaction of the initial cross-province distribution of faculties in Italy with a vector of university reforms that occurred in Italian history. The initial distribution of faculties is key to our research design: the political unification of the Italian territory into the sovereign state of Italy started in 1861 and was completed in 1871 with the capture of Rome.² Italian provinces inherited their higher education institutions from the ancient sovereign states. We argue that this initial supply persistently affected the creation of new faculties after Italian unification. University reforms shaped the higher education supply of each province differently depending on this initial distribution. Throughout the analysis, we perform a series of tests and extensive robustness checks recommended by the literature that discusses the validity of Bartik instruments (Goldsmith-Pinkham et al., 2020; Adão et al., 2020). This sensitivity analysis reassures us that our results are not driven by persistent omitted factors correlated with the initial distribution of faculties.

¹The 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.

²The newly born state was a monarchy (Kingdom of Italy) until 1946, when following an institutional referendum, a new republican constitution was adopted.

The main result of this paper is that the higher education supply has sizeable and significant long-term effects on economic welfare. Our estimates predict that a province exposed to the average supply of faculties in 2010 presents a 4.9% higher real income per capita than a province benefiting from a zero supply. On average, in 2010, a province was exposed to 25 faculties, with 5 being local ones and the remaining 20 developing nearby. Nevertheless, total returns seem equally shared between local effects and geographical externalities, as one local faculty has a four-fold greater marginal effect than a faculty in a nearby province. However, we argue that standard computations of geographical externalities neglect displacement effects of faculties located in close neighbourhoods: having one more faculty nearby reduces local supply by 0.3 faculties. We show that once we properly account for the corresponding foregone value, economic returns are highly localised at the province level. Finally, we uncover very considerable inequalities between small peripheral provinces that enjoy limited economic returns from higher education and large provinces where the supply of faculties explains up to 25% of local income per capita.

Our paper builds a bridge between two strands of the literature that investigate the economic effects of higher education institutions in developed countries. The first strand focuses on the preindustrial era, i.e., an epoch where upper-tail education was reserved to a minority of élite interested in scientific progress (Mokyr, 2005). Papers from this literature adopt a historical perspective inspired by Unified Growth Theories (Galor, 2005). They analyse the role of upper-tail human capital as a determinant of preindustrial economic development (e.g., in terms of population growth, urban population density, and living standards) and the emergence of industrialisation (e.g., measured in terms of factory or manufacturing employment). See Squicciarini and Voigtländer, 2015; Cantoni and Yuchtman, 2014; and Becker et al., 2011). More recently, DelaCroix et al. (2020) studied how the mobility of scholars across European universities contributed to the rise of upper-tail human capital during the Middle Ages and early modern eras. The second strand of the literature analyses the effect of modern universities on productivity and growth in postindustrial societies (Andersson et al., 2004; Valero and Van Reenen, 2019; Aghion et al., 2009).³ These studies focus on the second half of the XX^{th} century, an epoch where public education policies expanded and decentralised undergraduate education to make it geographically and socially more accessible. These studies focus on outcomes related to productivity, economic growth and

³Andersson 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, while Valero and Van Reenen (2019) use a novel dataset with regional panel information on nearly 15000 universities across 78 countries for the post- WWII period. Aghion et al. (2009) exploit variation in the appointment of US Congressmen in Appropriations Committees to identify the effect of education investments on economic growth for a panel of 48 US states during 26 cohorts. The authors find positive effects of education investments on growth, whose size depends on a state’s distance from the technological frontier.

institutional quality. The 150 years covered by HIU data highlight the long-term effects of higher education institutions from the onset of the industrial era to today. This long period covers the full set of reforms that favoured the transition from élite to mass education. Our measure of higher education supply encompasses both disciplines considered élite knowledge and applied disciplines. At the start of our time period, only the former would have full academic status; the latter would still foster specialised competencies and skills of a higher level than those handled by an individual with average education or literary skills. However, they would be granted full academic status only during the XXth century.

As a second important contribution, we provide an accurate quantification of the long-term welfare effects of higher education supply. Our fine-grained province-level data enable us to advance the measurement of geographical externalities. In particular, we account for the welfare cost associated with the fact that the existence of faculties in a sufficiently close neighbourhood (e.g., in the same region, within a certain distance, etc.) prevents opening a similar faculty in a province. This is intuitively similar to how province-level competition attracts a new production plant (see, e.g., Greenstone et al., 2010).⁴ We are the first to point out this *displacement cost* in terms of local welfare, which determines a negative geographical externality. The scant empirical literature on the welfare effects of higher education institutions discusses only ‘gross’ positive externalities that neglect displacement forces from the neighbourhood. Andersson et al. (2004) find gross positive externalities are very limited in the case of Sweden such that economic returns are strongly localised. Valero and Van Reenen (2019) find major positive externalities instead, which account for approximately half of local economic returns at the regional level. We uncover gross positive externalities of similar magnitudes as those of Valero and Van Reenen (2019). However, as we consider negative displacement costs, the ‘net’ positive externalities are much smaller and account only for a quarter of average economic returns (approximately 1% of local income per capita in the average province).

We also make a third 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

⁴The competition mechanism is inherently different, however. In the case of plants, competition is the outcome of their own location decisions. In international trade and industrial organisation, this competition is triggered by the entry of new firms into the market (see, e.g., Ottaviano, 2011 and Berry and Reiss, 2007 for syntheses). In our case, this mostly involves yardstick competition, as the central government chooses the location of faculties. In this respect, our paper also broadly relates to the literature that studies the long-run changes of education institutions (see, e.g., Woesmann, 2003; Schutz et al., 2005; and Braga et al., 2013).

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. As in this literature, we propose 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 exogenous university reforms that occurred in Italian history. This guarantees that our instruments have considerable power in the first stage. It also enables us to include in the empirical specifications an extensive set of fixed effects at the province and region-by-year levels. Additionally, Andersson et al. (2004) and Valero and Van Reenen (2019) estimate fixed effects models at the subnational level; however, they lack suitable instruments.

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 analysis and main results of neighbourhood effects are presented in Section 4. Section 5 presents the welfare analysis. Section 6 concludes.

2 A province-level supply of higher education

The province is our benchmark unit of observation. This is the natural geographical level for measuring human capital accumulation and analysing industrialisation, production and wage patterns (see, e.g., Squicciarini and Voigtländer, 2015; Ciccone and Peri, 2006, 2011; and Bratti and Leombruni, 2014). In the first part of this section, we describe the role of provinces in the provision of higher education in Italy before and after its political unification. We then explain why a province-level analysis provides a theoretical rationale to consider neighbourhood effects in the provision of higher education.

2.1 Supply of higher education in preunitarian provinces

At the time of its political unification, Italy had 69 provinces. Not all of them became part of the sovereign state of Italy in 1861. Eight provinces joined in 1866 (Belluno, Padova, Rovigo, Treviso, Udine, Venezia, Verona, and Vicenza). Mantova joined in 1868. Roma joined in 1870, completing Italian unification. Research on the history of Italian universities demonstrates that the initial supply of higher education institutions across these provinces was very heterogeneous (Brizzi and Romano, 2007).

Table 1 summarises information on this initial supply in many relevant dimensions. In Column [1], we list the 21 provinces with at least one faculty at the time they joined

Italy. In Column [2], we report the year in which the first faculty appeared. In most cases, this had already occurred by the XVth century. In Column [3], we list the states wherein the first faculty was established. We call these “Initial States”. The vast majority of Initial States were Duchies, Counties, Principalities, Houses, Republics or Communes, which covered the area of only one province. In most initial states, university education was organised into four faculties: Literary Studies (which included Philosophy), Law, Medical Studies, and the Natural Sciences (which included Mathematics and Physics). In addition to these, autonomous faculties of Theology were present in the states under the control of the Vatican (“Papal States”).

In Column [4], we report the last state to which the province belonged to before Italian unification (we call this the “Preunitarian State”). The comparison between the Initial State and the Preunitarian State sheds light on the burgeoning political process taking place between the XI^{th} and IXX^{th} centuries. Smaller Initial States would lose their political independence (e.g., House of Este in Ferrara was absorbed by the Papal States, and the Republic of Genova was absorbed by the Kingdom of Sardinia). Some larger Initial States would be gradually dismantled and disappear (e.g., the Holy Roman Empire) or go through major political transformations due to successions of political dynasties (as in the Kingdom of Sicily, Kingdom of Sardinia, and Papal States).⁵ This process would have major consequences for the design of upper tail education at the province level. As an example, the Spanish Empire would create new education institutions to form local bureaucrats in the occupied provinces (e.g., Cagliari and Catania). It would export its own academic model and impose its own different accreditation rules, with which local institutions would have to comply, to be granted academic recognition (Roberts, 1996).

Column [5] reports the total number of faculties already present in each preunitarian province at the time they joined Italy (we call them preunitarian faculties). There are 87 preunitarian faculties in the Italian territory. Province-level supply ranges from a minimum of two faculties (in Milano, Perugia, Pesaro-Urbino, and Siena) to a maximum of eight faculties (in Torino). These faculties cover the four academic disciplines that the Italian higher education system had retained from preunitarian states (i.e., the Humanities, Mathematics and the Natural Sciences, Law, and Medicine; see De Ridder-Symoens and Rüegg, 1992). In addition, there are faculties in applied or technical fields, called for by Scientific and

⁵The Kingdom of Sicily would turn into the Kingdom of Naples under the Angevin dynasty and then the Kingdom of Two Sicilies under the Bourbons. The Kingdom of Sardinia would initially be a member of the Council of Aragon and then became a part of the Spanish Empire in 1402. In 1720, it became part of the Duchy of Savoy until Italian unification. Papal States changed their domain over time, at their greatest extent, in the 18th century, including most of central Italy and the Legations of Ravenna, Ferrara and Bologna.

Commercial Revolutions. These would appear as specialised, autonomous schools within the faculties of the Natural Sciences (e.g., for Architecture or Engineering) or technical schools that formally operate outside of the perimeter of university education (e.g., Economics and Trade). Nevertheless, these were by all means equivalent to university faculties, as they would deliver a diploma equivalent to a university degree. This diploma would be necessary to work in the Italian public sector or access field-specific highly qualified professions. Many of these specialised schools would be opened in the main capital cities of each Preunitarian State (see online appendix A1 for details).

Finally, Column [6] reports the initial assessment of these higher education institutions contained in Law 3725/1859 (called Casati from the Minister of Education at the time). This law set out the rules for accrediting preexisting institutions into the new Italian university system.⁶ In nine provinces, higher education supply was ranked “A”, as preunitarian supply fully complied with the highest quality standards set out by the Italian government in terms of teaching and research. In several provinces, preunitarian supply was ranked “B”. These faculties would not meet the highest quality standards but provided useful second-tier regional institutions. Faculties belonging to the Universities of Sassari and Macerata fell short of the minimum requisites and were ranked “C”. The Casati Law also acknowledged private higher education provision in the provinces of Ferrara, Pesaro Urbino and Perugia, where the local authorities used to directly patronise universities (Brizzi and Romano, 2007).⁷

Overall, Table 1 suggests that the higher education system of the Italian Kingdom emerged from the very beginning as an aggregation of the largely heterogeneous education systems the several states present in the Italian territory put in place starting from the Middle Ages. Figure A-I in the online appendix shows that this preunitarian supply was also dispersed geographically between the northern, central and southern regions.

2.2 Supply of higher education in Italian provinces

The above discussion suggests that the Italian higher education system at the end of the 19th century was very polycentric and decentralised. Successive attempts to centralise it at the state level during the XX^{th} century were largely ineffective (Fioravanti et al., 2000). Both in old and recent times, the province has always been the relevant territorial jurisdiction for the supply of higher education (and local public services in general). Petracchi (1962)

⁶The Casati Law was initially applied to new territories of the Kingdom of Sardinia. The Matteucci Law extended it to all territories that gradually entered Italy (see Table A-I in the online appendix).

⁷DelaCroix et al. (2020) show that the Casati ranking is strongly consistent with their index of the notability of a university. This uses a very different metric based on human capital estimates of the top 5 academic scholars who spent time at a university from 1740-1800.

Table 1: Preunitarian higher education supply

[1] Preunitarian Province	[2] Year when First Faculty Appeared	[3] Initial State	[4] Preunitarian State	[5] Preunitarian Faculties	[6] Ranking Casati Law
Bologna	1088	Holy Roman Empire	Papal States	6	A
Cagliari	1620	Kingdom of Sardinia (Spanish Empire)	Kingdom of Sardinia	4	B
Catania	1445	Kingdom of Sicily (Spanish Empire)	Kingdom of Two Sicilies	4	B
Ferrara	1391	House of Este	Papal States	4	Private
Genova	1481	Republic of Genoa	Kingdom of Sardinia	5	B
Macerata	1540	Papal States	Papal States	3	C
Messina	1838	Kingdom of Two Sicilies	Kingdom of Two Sicilies	3	B
Milano	1791	Duchy of Milan (Habsburg Empire)	Lombardy-Venetia	2	A,A
Modena	1175	Commune of Modena	Duchy of Modena and Reggio	4	B
Napoli	1224	Kingdom of Sicily	Kingdom of Two Sicilies	6	A
Palermo	1806	Kingdom of Sicily (Spanish Empire)	Kingdom of Two Sicilies	4	A
Padova	1407	Commune of Padova	Lombardy-Venetia	4	A
Perugia	1308	Commune of Perugia	Papal States	2	Private
Pisa	1343	Republic of Pisa	Gran Duchy of Tuscany	6	A
Parma	962	County of Parma	Duchy of Parma and Piacenza	5	B
Pesaro Urbino	1671	Papal States	Papal States	2	Private
Pavia	1361	House of Visconti	Kingdom of Sardinia	5	A
Roma	1303	Papal States	Papal States	5	A
Siena	1240	Republic of Siena	Gran Duchy of Tuscany	2	B
Sassari	1765	Kingdom of Sardinia	Kingdom of Sardinia	3	C
Torino	1404	Principality of Savoy-Acaia	Kingdom of Sardinia	8	A,B

Notes: Column [1] reports the Preunitarian province. Column [2] reports the year of the first faculty in the province. Columns [3] and [4] report the Initial and Preunitarian State to which the province belonged. In Column [5], we report the number of faculties observed in the province at the moment of entry into the Italian Kingdom. Column [6] 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.

defines provinces as “big associations of municipalities devoted to the protection of the rights of each of them and to the management of their collective moral and material interests”. Among these moral and material interests, there was indeed the local demand for higher education, which would be conveyed into the creation of new faculties. This was very often detached from that of universities, and there are many cases in which the genesis of a faculty is antecedent to the creation of the university it belongs to.⁸

Eurydice (2000) mentions eight general interest reforms of the Italian university system. These provide the most significant shifts to the institutional process of faculty creation, which we exploit in the empirical analysis. In the pre-WWII period, Law 2102/1923 (called Gentile, from the Minister of Education) was the first attempt to create an organic framework for Italian higher education and featured the opening of new faculties and single faculties in social sciences and STEM (i.e., Science, Technology, Engineering and Mathematics) fields.⁹ Law 1592/1933 (called De Vecchi-Bottai) with ancillary interventions created a more overarching and centralised higher education system such that technical schools would eventually either reach full university status (e.g., Bocconi University) or be absorbed as faculties of existing universities. Post-WWII reforms produced a new institutional setting. The national university system, which was previously organised to serve the “élite”, later on was designed to provide schooling for the “masses”. Law 910/1969 liberalised access to students with a 5-year diploma of secondary education from technical schools. The consequent rise in demand put pressure on the national university system to adapt its supply. Indeed, Law 766/1973 legislated the opening of new faculties and increased the number of faculty hires. Law 382/1980 reorganised the internal governance of universities as well as the recruitment and careers of university professors. Successive Laws 392/1989 and 245-341/1990 modified the allocation of funding.¹⁰ Law 59/1997 granted universities more financial and teaching autonomy.

As a result of these reforms, new faculties would continue to be opened in provinces without a preexisting supply of higher education. Provinces would not have the political power to systematically influence these general state-level reforms, e.g., affect the total supply of faculties in the country.¹¹ However, they play a key institutional role in determining

⁸An example is the University of Eastern-Piedmont, which was founded in 1998 from seven faculties opened in 1993 as separate branches of the University of Torino in the Provinces of Novara, Vercelli, and Alessandria. Further details can be found in the online appendix.

⁹The reform also restored the distinction between A-level and B-level higher education institutions, which had been progressively excised over the years (see Table 2 in the online appendix for details).

¹⁰The Law established a dedicated “Ministry of Universities Scientific and Technological Research” and set up triennial development plans for universities (see the online appendix for details).

¹¹The weak political relevance of provinces for national politics dates back to the Republican Constitutional Law, which does not mention their exact tasks and legislative competencies (cfr. art. 114; see also Fabrizio,

the territorial distribution of the higher education supply. Province-level *consortia* would present a start-up project for new faculties or universities, which would then be evaluated by the central state (Breccia and Salustri, 2017). Province-level politicians gather academic competencies and financing institutions (banks, chambers of commerce or local investors). The central government simply acts as a regulator and sets administrative requirements, being rather agnostic about the distribution of faculties in the Italian territory.

Figure 1: Diffusion of faculties in Italy

Notes: Authors' calculation based on HIU data using a map of Italian provinces in 2010. See Table A-1 for a disaggregation by field of study.

Figure 1 gives a historical overview of the total number of faculties in 1870 by the time the Italian unification was complete. This is compared to the number of faculties in 2010 (at the 2010 province disaggregation level for comparability purposes). As early as 1870, the distribution of faculties was concentrated in the capitals of preunitarian states. In 2010, two-thirds of the provinces had at least one faculty. Provinces without faculties are the most recent, established in the post-WWII period or more recently. Few provinces had 15 faculties or more (Bari, Bologna, Milano, Napoli, and Roma). Figure 1 visually suggests that the increase in the number of faculties is widespread throughout the Italian territory.

2008a, 2008b, and 2008c for a historical overview). In addition, their broad dispersion in the Italian territory implies that each province covers only a negligible fraction of the total population.

Figure 2 offers more insight into the evolution of higher education supply over the period of 1870-2010. As detailed in Panel (a), the rise of higher education supply clearly emerges after 1969 and involved the humanities, STEM, and the social sciences in a roughly similar way. Panels (b) and (c) suggest that the direction of local responses to the reforms was homogeneous: the total number of faculties increases in all NUTS1 regions and in all quartiles of the province population. Finally, Panel (d) shows that the evolution of supply in Italy from the mid-nineteenth century shares a common pattern with that of France, Germany and the UK. All four countries are characterised by an increase in the number of universities, especially starting from the 1960s. While Italy possesses a flatter profile relative to the other three countries,¹² this pattern features the transition from élite to mass education, which took place in Europe and the US during the same period (Eurydice, 1999; Smith, 2010)

2.3 Neighbourhood effects on the supply of higher education

Based on the above arguments, in this paper, we treat Italian provinces as independent units for the supply of higher education. Higher education is a quasimarket wherein local supply is determined by local market forces and territorial competition without intervention by the central government, e.g., to avoid oversupply in close proximity (Catalano and Silvestri, 1999; Agasisti and Catalano, 2007). In this respect, the Italian case is similar to that of many other countries (OECD, 2008).

This suggests that a spatial economics conceptual framework may be appropriate to discuss why neighbourhood effects may arise in the supply of higher education at the province level. In such a framework, each province autonomously sets the local provision of higher education, given constant local demand from students.¹³ This produces province-level competition in the supply of higher education, since multiple neighbouring provinces insist on the same local demand. However, competition does not occur on the price margin, as the Italian public higher education system is characterised by low and uniform tuition fees, so provinces cannot attract students by reducing the fees of their local universities.¹⁴ Competition may occur purely through neighbourhood effects associated with faculty openings: an additional faculty in the province may attract more students, thereby reducing enrolment

¹²Several reasons might explain this difference, such as the persistently fewer high school graduates compared to those of countries such as Germany, especially in the last 40 years.

¹³This approach recalls tax setting by autonomous jurisdictions (see, e.g., Parchet, 2018), which is fully symmetric to traditional spatial competition models *à la* Hotelling, where it is the supplier that chooses its location over space instead.

¹⁴Moreover, during our sample period, universities do not have a binding budget constraint: public universities are financed by the central government through lump-sum transfers, whose size increases with both students' admissions and past university expenditures.

Figure 2: Evolution of the number of faculties in Italy from 1870-2010 and EU comparison

Notes: Panel (a) reports the number of faculties in Italy by field of study. Humanities include Education, Linguistic Studies, Literature, and Psychology. The STEM (i.e., Science, Technology, Engineering and Mathematics) group includes Agricultural Studies, Architecture, Chemistry and Pharmacy, Geology and Biology, Engineering and Scientific Studies. Social Sciences include Medical Studies, Economics and Statistics, Law, and Sociopolitical Studies. Authors' calculations based on HIU and WHED UNESCO data. 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 2009. Authors' calculations based on HIU data. 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).

into faculties of neighbouring provinces, as admissions are generally free.

This conceptual framework builds on three important assumptions. The first is that students are a ‘scarce’ resource, particularly with low geographical mobility (e.g., due to high commuting costs), such that each province can access only a limited portion of the national demand. Figure 3 reports the distribution of high school graduates in Italy by distance from the university they are enrolled in. This suggests that students’ mobility is indeed very low: the majority of Italian high school graduates choose a higher education institution at a distance of less than 150 km. Mobility is higher among students with high aptitudes towards higher education, and it is particularly low among low aptitude students. This recalls well-known results for the US (see Hoxby, 2016). The second assumption is that some degree of differentiation (e.g., in teaching content) exists between faculties in neighbouring provinces. If preferences of students are heterogeneous and defined not only based on faculty distance but also on study content, then attending the closest faculty may not be the optimal choice. Provinces may take advantage of this dimension of heterogeneity and differentiate their supply. Indeed, our analysis suggests that neighbourhood effects are mostly concentrated within the same field of study (cfr. Table 10 below). The third assumption is that opening a new faculty may produce some welfare advantages to provinces. A province will open a new faculty if this maximises total welfare and taking into account that neighbouring provinces may strategically react to anticipate their expected welfare losses (due to the displacement of their own supply). We provide detailed evidence of these welfare effects in Section 5, qualifying some results from previous studies (see, e.g., Valero and Van Reenen, 2019; Aghion et al., 2010; and Andersson et al., 2004).¹⁵

3 Data and descriptive statistics

Our main dataset is the History of Italian Universities (HIU) dataset. This is an own compiled original register that contains detailed information on higher education institutions in Italy, disaggregated at the faculty level. The dataset covers all years between 1861 and 2010.

Our primary source used to compile the HIU register is Brizzi and Romano (2007), which we complemented using several sources on the history of specific universities (see, e.g., Fois, 1991 on the University of Sassari) and faculties (e.g., Silvestri, 2006 on engineering).

¹⁵Notice that provinces may also expand their supply of higher education to increase public transfers from the central government. This form of yardstick competition can provide a complementary explanation for neighbourhood effects, which we do not investigate in the present paper.

Figure 3: Students' mobility in faculty choice

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).

We also heavily relied on *Gazzetta Ufficiale della Repubblica Italiana*, a weekly publication that collects every public act taken by the government since 1861. We double-checked all information against that provided by open-source archives, i.e., Wikipedia, universities' and faculties' websites.¹⁶ 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: almost all faculties that were created in Italian history were still present in 2010.¹⁷ The final HIU register includes the name of each faculty and its current address and 15 field identifiers, which we aggregated into 3 macroareas of science (Social Sciences, STEM, and Humanities), according to the classification used by the Italian National Statistical Office (ISTAT). 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

¹⁶Since 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.

¹⁷The only two exceptions are one scientific faculty (Chemical Studies) at Ca' Foscari University of Venezia, which was closed in 1990, and 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.

according to the Casati ranking and its following updates (see online appendix B for details).

In this paper, we consider only higher education institutions offering standard BSc education, i.e., 574 faculties in 71 universities registered in the Italian territory between 1861 and 2010.¹⁸ Based on this sample, we construct a province-level panel that counts the number of faculties present in each province and year. We exclude the years 1861-1869, as the Italian unification process was undergoing in this period. The final dataset is a province-level panel that includes 11792 observations for 110 provinces between 1870 and 2010. The panel is unbalanced because the number of provinces changes throughout the time period. The main variable is the number of faculties in province i , as a total and by macroarea of science. We also record information on the total number of universities; private universities; and A-, B-, and C-level universities in province i at time t . We complement HIU data with extensive province-level information from multiple sources (e.g., the Italian censuses; see Section B3 of the online appendix for details). In particular, we use ISTAT and Google Maps API data to construct matrices that map spatial interactions between provinces based on contiguity, linear distance, travel distance, and travel time (see details in the next section).

Table 2 presents summary statistics for the main variables. Panel A describes the local higher education supply. On average, each province has a supply of over 2.5 faculties, mostly in STEM and the Social Sciences. Each province has approximately 0.5 universities, of which 15% are private, and over 75% are classified as of level A.¹⁹ Panel B reports the higher education supply of each individual neighbour j of province i . 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). A representative neighbour, on average, has approximately 2.9 faculties; other figures are similar to those displayed in Panel A. Panel C reports indicators for higher education supply in neighbourhood $-i$. On average, each province has approximately 12 faculties in the neighbourhood with approximately three in each neighbouring province (i.e., each province has approximately 4 neighbours, on average). The other figures are similar to those displayed in Panels A and B. Finally, Panel D reports province-level historical information. On average, in our sample, each province presents a yearly real income per capita equal to 838 euros (constant 1911 prices) and a population of

¹⁸In practice, we exclude universities offering only postgraduate education or single courses or enrolling foreign students only. These include Università di Scienze Gastronomiche, Università Europea di Roma, Università Normale di Pisa, Scuola IMT Alti Studi di Lucca, Scuola superiore di studi universitari e perfezionamento “S. Anna” di Pisa, Università degli studi di Roma Foro Italico, Istituto universitario di studi superiori di Pavia, Scuola internazionale superiore di studi avanzati di Trieste, Università per Stranieri di Siena, and Università per Stranieri di Perugia.

¹⁹In particular, the high percentage of A-level universities is because the De Vecchi-Bottai Reforms, starting from 1933, classify all public universities as A-level institutions. Details on the construction of this percentage indicator are available in Sections A1 (see Table A-I) and B3 of the online appendix.

Table 2: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Higher education supply in province i					
no. of faculties in i	2.515	4.153	0	38	11792
no. of humanistic faculties in i	0.479	0.99	0	11	11792
no. of stem faculties in i	1.07	1.752	0	9	11792
no. of social sciences faculties in i	0.959	1.66	0	19	11792
no. of universities in i	0.497	0.842	0	7	11792
no. of private universities in i	0.078	0.364	0	4	11792
no. of A-level universities in i	0.381	0.621	0	3	11792
Panel B: Higher education supply of neighbouring province, j					
no. of faculties in j	2.85	4.43	0	38	50803
no. of humanistic faculties in j	0.537	1.062	0	11	50803
no. of stem faculties in j	1.212	1.837	0	9	50803
no. social sciences faculties in j	1.091	1.783	0	19	50803
no. of universities in j	0.546	0.895	0	7	50803
no. of private universities in j	0.093	0.415	0	4	50803
no. of A-level universities in j	0.418	0.639	0	3	50803
Panel C: Higher education supply in the neighbourhood, $-i$					
total no. of faculties in $-i$	12.28	9.835	0	68	11792
average no. of faculties in $-i$	2.884	2.168	0	17.333	11792
average no. of humanistic faculties $-i$	0.536	0.544	0	4.333	11792
average no. of stem faculties in $-i$	1.24	0.909	0	5	11792
average no. social sciences faculties in $-i$	1.098	0.873	0	8.333	11792
average no. of universities in $-i$	0.551	0.398	0	2.667	11792
average no. of private universities in $-i$	0.089	0.196	0	1.333	11792
average no. of A-level universities in $-i$	0.422	0.31	0	1.667	11792
Panel D: Economic and demographic characteristics of province i					
Gross income per capita, real terms (in 000s)	0.838	0.84	0.15	3.933	1080
Population	566360	506697	102497	4018108	1012
Population 0-14 years old	110888	118763	10580	853007	674
Participation rate	45.456	8.379	3.978	98.424	1012
% of active individuals in industry	27.519	17.11	4.742	436.068	1012
% with tertiary education	3.983	3.554	0.416	16.804	681

Notes: Authors' calculation based on HIU data. Panel D includes information taken from 16 waves of Italian Censuses from 1861-2009, the ISTAT and Istituto Tagliacarne. Gross income per capita in real terms is expressed in euros at 1911 constant prices.

over 566 thousand individuals. Approximately 20% of individuals are less than 14 years old, 45% are active in the labour market, and approximately 27% operate in the industry sector. Only 3.9% of the population in the province has a tertiary education degree.

4 Neighbourhood effects on higher education supply

In this section, we analyse neighbourhood effects on the local supply of higher education in a province, which are then incorporated in the welfare analysis in Section 5.

4.1 Empirical strategy

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 the neighbourhood. 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.²⁰ All models feature heterogeneous neighbourhood relationships, which are immune to Manski (1993)’s reflection issues.²¹

We define spatial interactions in two complementary ways. We start by assuming that higher education supply in province i depends on that of its “average neighbour”. This is defined as $\bar{F}_{-it} = \sum_{j \neq i}^{N_i} w_{ij} F_{ijt}$ (Brueckner, 2003). All neighbours have the same weight: $w_{ij} = \frac{1}{N_i}$ where N_i is the number of contiguous provinces of i . We model this relationship as follows:

$$F_{it} = \alpha + \beta \bar{F}_{-it} + \gamma_i + \delta_{r(i)t} + X_{it}\phi + \epsilon_{it}, \quad (1)$$

where β is the coefficient of interest. γ_i and $\delta_{r(i)t}$ are, respectively, province and region-by-year fixed effects (where $r(i)$ refers to the region to which province i belongs). X_{it} is a vector of time-varying province-level covariates. This includes the total number of universities and the number of private and A-level universities as time-varying province-level indicators of the quality of higher education.²² In several specifications, we also include the size of the total population, population in the 0-14 age cohort, population share in tertiary education, labour force participation rate, and size of the industry sector.²³

We also estimate a “pairwise” version of equation (1), which better accounts for the omitted determinants of province i ’s supply (e.g., related to geographical factors) that are

²⁰This is done in Table 12 below and Table A-III in the online appendix.

²¹The 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.

²²In Table A-XII, we also include interactions of the notability index developed by DelaCroix et al. (2020) with a time trend and with dummies for reforms that reduced the quality of higher education provisions in the postwar period (Bianchi, 2020).

²³This is shown in Table 7 below. We do not consider these province-level indicators for the baseline estimates because they are available from the Italian Census every ten years and, in some cases, for the postwar period only. It may also be argued that over the 150-year time period, some of these factors might not be immune to reverse causality concerns, as the structure of the economy, population and labour supply can themselves be the result of the diffusion of higher education.

spatially correlated with those of its j neighbours (see Parchet, 2018):

$$F_{ijt} = a + bF_{jt} + c_{ij} + d_{r(i)jt} + X_{ijt}f + e_{ijt}. \quad (2)$$

F_{jt} is the higher education supply in each neighbouring province j of province i . Our main indicator of higher education supply is the total number of faculties in province i and in the neighbourhood (i.e., the average neighbour $-i$ in equation (1), or each neighbour j in equation (2)). Throughout the analysis, we also use other indicators. We distinguish the extensive margin (i.e., having at least one faculty in province i and in the neighbourhood) from the intensive one (i.e., the number of faculties, where a higher education institution is available). We also measure supply as the number of universities available in province i and its neighbourhood.

Equation (2) can be estimated for all contiguous ij pairs. Each couple appears twice, with a given province shown once on the left side of equation (2) and once on the right side. Coefficient b captures the average effect of each neighbour j 's supply on the local supply of province i . c_{ij} is the fixed effect of the pair. Relative to equation (1), the inclusion of pair fixed effects accounts for the time-invariance of omitted factors that pertain to each ij couple.

Identification and estimation issues. The heterogeneity that is left in equations (1) and (2) after the inclusion of the province (or province pairs) and region-by-year fixed effects is the variability over time across provinces (or province pairs) within the same region. The identification of β and b in equations (1) and (2) is achieved through a comparison of different time-varying patterns in the number of faculties across provinces. While equations (1) and (2) assume contemporaneous effects, in our preferred empirical specification, we model a ten-year lagged effect. As discussed in Section 2 above, several actors are involved in the decision to open a new faculty, which implies a delay in the effect of neighbours on local supply.

The main endogeneity concern in models (1) and (2) is reverse causality. The number of faculties of a neighbouring set of provinces itself depends on the number of faculties of province i . For example, in equation (2), any change in F_{ijt} may induce an adjustment of F_{jt} . Moreover, unobservable time-varying determinants of one province's number of faculties (e.g., local economic conditions and 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 .

To address endogeneity concerns, we implement a Bartik instrumental variables approach that exploits variation in the total number of faculties observed across neighbouring provinces j s at time 0, IC_{j0} .²⁴ These are the initial conditions of the Italian higher education supply, determined by the burgeoning historical process discussed in Section 2.1. We interact these initial conditions with the vector of eight major reforms of the Italian university system, R_t , which took place between 1923 and 1997 (cfr. Section 2.2). For the pairwise model (2), our instrument is constructed as follows:

$$Z_{jt} = IC_{j0} * R'_t \quad \text{where } R_t = 1 \text{ if } t > \text{year of reform, } 0 \text{ otherwise.} \quad (3)$$

For the average neighbour model (1), we construct $\bar{Z}_{-it} = \bar{IC}_{-i0} * R'_t$, where $\bar{IC}_{-i0} = \sum_{j \neq i}^{N_i} w_{ij} IC_{ij0}$. In vector R_t , each university reform dummy takes a value of 0 before the reform and 1 starting from the year of the reform. This features the idea that each successive reform builds on those already implemented so that reforms have an incremental effect.

Equation (3) features a research design whereby initial conditions measure the exposure of preunitarian provinces to posterior reforms of the Italian university system. In our IV approach, general higher education reforms R_t trigger the faculty creation process over Italian history. These reforms are independent of previous provincial conditions and likely uncorrelated with past shocks. Following Goldsmith-Pinkham et al. (2020), we argue that the identification of Z_{jt} comes from the cross-province variation of IC_{j0} . Provinces inherited higher education institutions from the preunitarian Italian states. These institutions persistently affected the process of faculty creation at the province level such that the same general reform shaped the higher education supply of each province j differently depending on its initial conditions.

In Table 3, we show that the initial distribution of the number of faculties across preunitarian provinces was not correlated with observable province level characteristics before the start of our sample period. In Columns [1]-[5], we check whether initial conditions correlate with time invariant characteristics of province i at the onset of Italian unification, i.e., the percentage of the provincial population residing in an urban area with more than 6000 inhabitants, the linear distance (in km) of province i from Bologna (i.e., the city with the oldest active university), income per capita, and the share of workers in the agriculture and industry sectors. In Columns [6]-[10], we consider the pretrends of available indicators during various preunitarian spells, i.e., urban wage growth in province i from 1850-1861 (Malanima,

²⁴Time zero is 1861 for fifty-nine preunitarian provinces. We observe the initial conditions of the remaining ten provinces due to the lengthy Italian unification process. Thus, time zero is 1866 for Belluno, Padova, Rovigo, Treviso, Udine, Verona, Vicenza and Venezia; 1868 for Mantova and 1870 for Roma.

2003; Malanima, 2006); urban population growth in province i from 1800-1861 and 1600-1800; the change in the probability of having a university in province i ; and its academic notability between 1600 and 1800 (DelaCroix et al., 2020). We subject these correlations to an alternative clustering of standard errors. We report standard errors clustered at the province level between (...). Standard errors are clustered at the regional level between [...] and at the level of the preunitarian state the province belonged to between {...}.

Panel A presents the battery of correlations of neighbours' initial conditions, IC_j , with the aforementioned set of preexisting characteristics of province i . The first battery of tests in Columns [1]-[5] suggests that IC_j s are uncorrelated with the time invariant characteristics of province i . The second battery of tests of pretrends in Columns [6]-[10] is consistent with the view that IC_j s neither correlate with long-term developments, e.g., associated with the Scientific and Commercial Revolutions, nor with short-term productivity changes, e.g., due to the economic and political turmoil generated by Italian unification: the negative correlation with urban wage growth in province i from 1850-1861 becomes insignificant when standard errors are clustered at the regional or preunitarian state level.²⁵ Overall, this evidence does not seem to suggest any systematic bias in the estimation of neighbourhood effects.

In addition, the results in Panel B do not suggest strong correlations of IC_i with either levels or changes of preexisting economic characteristics. Relevant exceptions are the positive and significant correlations of IC_i to local 1600-1800 changes in the probability of having a university in Column [9] and the change of academic notability in column [10]. This confirms the importance of the historical process described in Section 2.1: the new country inherited the set of higher education institutions that preunitarian states had designed independently in the past. Note that the correlations presented in Panel B are not strictly relevant to the identification of Z_{jt} in equation (2); they will be useful when discussing the identification of the welfare effects of higher education supply instead (see Section 5 below).

Overall, the evidence in Table 3 does not point to persistent omitted factors determining the initial distribution of the number of faculties across preunitarian provinces. To be further reassured, we perform an extensive sensitivity analysis of our empirical results. We exploit the 1861-2010 time period and the richness of events that occurred in Italian history to propose an alternative definition of initial conditions (based on the pre-WWII endowment of faculties at the province level) and narrow down the contribution of specific higher education reforms (see Table 8 below). We also replicate our results using a standard shift-share instrument, which we construct as a linear combination of a vector of province-specific faculty shares by field of study $k = 1...K$ and the total faculty growth rates in each field of study

²⁵For data on urban wage feature variation at regional and higher levels, see online appendix B.

Table 3: Correlations of IC_j and IC_i with the characteristics of province i in 1870 and various pretrend periods from 1600-1861.

Correlations of initial levels				Pretrends 1600-1861						
	[1] urbanization rate in i (> 6000 inh.)	[2] distance of i from Bologna	[3] per capita income of i	[4] share of agric. in i	[5] share of industry in i	[6] $\Delta_{1850-1860}$ urban wages in i	[7] $\Delta_{1800-1861}$ urban pop. in i	[8] $\Delta_{1600-1800}$ urban pop. in i	[9] $\Delta_{1600-1800}$ University in i	[10] $\Delta_{1600-1800}$ notability in i
Panel A) IC_{j0}	-0.5845 (0.6750) [0.6248] 0.7278	-0.0010 (0.0005) [0.0005] {0.0012}	3.6176 (5.5743) [7.4082] {8.6384}	-0.0058 (0.0089) [0.0082] {0.0072}	-0.0144 (0.0177) [0.0197] {0.0157}	-1.8244 (0.8891)** [1.5803] {2.0494}	-0.0693 (0.2212) [0.3390] {0.3324}	-0.0416 (0.1708) [0.1806] {0.1684}	0.1831 (0.4197) [0.4716] {0.6217}	-0.0510 (0.0814) [0.0935] {0.1132}
Observations	231	277	277	277	277	277	264	256	277	277
Panel B) IC_{i0}	0.7863 (1.2730) [0.5645] {1.1052}	-0.0007 (0.0009) [0.0008] {0.0011}	15.7991 (11.6664) [11.2938] {11.3030}	-0.0428 (0.0239)* [0.0265] {0.0211}	0.0397 (0.0339) [0.0320] {0.0100}**	-0.6341 (1.8581) [2.0851] {2.4428}	-0.0167 (0.4441) [0.4858] {0.4921}	-0.1635 (0.0841)* [0.1614] {0.1126}	1.8387 (0.4348)** [0.4813]** {0.6009}**	1.0687 (0.2175)** [0.2146]** {0.2635}**
Observations	58	68	68	68	68	68	68	68	62	64

Notes: In Columns [1]-[5], we report correlations of initial conditions with time invariant characteristics of province i . The urbanization rate is measured in 1861 for provinces that already belonged to the Italian Kingdom by then. All other variables are measured for 1871 for provinces except Belluno, Mantova, Padova, Roma, Rovigo, Treviso, Udine, Vicenza, and Verona and for 1861 for all other provinces. In Columns [6]-[10], we report correlations of initial conditions with growth rates of province i 's variables as reported in each column. In column [9], we report correlations with a dummy equal to 1 if a university was established in province i between 1600 and 1800. In column [10], we report correlations with differentials between the 1618 and 1800 values of the notability index developed by Delacroix et al. (2020) for province i . Standard errors are clustered at the province level between (). Standard errors are clustered at the regional level between []. Standard errors are clustered at the level of the preunitarian country between {}. Significance levels: *: 10% **: 5% ***: 1%.

during Italian history (see Table 9). In line with suggestions by Goldsmith-Pinkham et al. (2020), we replicate the analysis using alternative estimators that combine the moment conditions in potentially more efficient ways. Through the analysis, we also account for the fact that the territories and borders of preunitarian provinces change over time, alongside the creation of new provinces over Italian history (see Table A-II and the corresponding discussion in Section C1 in the online appendix). We implement the correction developed by Adão et al. (2020) to account for arbitrary correlation in the error term across preunitarian provinces (see Table A-VII). In our setting, this correlation is likely to show, as in many preunitarian provinces, initial conditions featuring a zero supply of faculties. Finally, we check how influential specific subperiods or regions are by removing two decades or multiple regions at one time and accounting for potential spatial dependence in the standard errors. This extensive battery of robustness checks shows qualitatively and quantitatively similar results. This makes us more confident that our results are not driven by omitted factors or outliers (see Section 4.3.2 for details).

4.2 Neighbourhood effects: empirical results

Table 4 presents the OLS estimates of equation (1) in Panel A and of equation (2) in Panel B. Row (a) shows the effect of the number of faculties in the “average neighbour” of province i on the number of faculties in province i at time t . In Rows (b), we describe higher education supply in each province as a dummy variable, which is equal to 1 if at least one faculty is present in the province at time t and 0 otherwise (the “extensive margin” of neighbourhood effects). In Rows (c), we describe neighbourhood effects only between provinces that have at least one faculty at time t (the “intensive margin”). Finally, in Rows (d), we measure supply in terms of the number of universities. We present the results of four different specifications. In Column [1], we use OLS with province fixed effects (OLS FE). In Column [2], we add year fixed effects. In Column [3], we include region-by-year fixed effects. In Column [4], we add province-level controls. Standard errors are always clustered by province.

The estimated coefficients for the average neighbour model of Panel A turn from positive in Column [1] to negative in Column [4] for all measures in Rows (a)-(d). This is mostly due to the inclusion of the region-by-year FE, which allows us to control for large regional inequalities in the evolution of higher education provision (Di Martino and Vasta, 2017). Estimates in Row (a) suggest that an increase in the neighbourhood’s number of faculties is associated with fewer faculties in province i , and this effect is significant at the 1% level. The results of the “pairwise model” (2) in Panel B are qualitatively very similar to average neighbour estimates: an increase in supply in neighbouring province j is negatively associated

Table 4: OLS FE estimates: Neighbourhood effects on higher education supply

	[1]	[2]	[3]	[4]	Obs.
Panel A - Average neighbour approach					
(a) average no. of faculties in $-i$	0.59*** (0.11)	-0.14 (0.10)	-0.76*** (0.25)	-0.48*** (0.14)	11383
(b) at least one faculty in $-i$	0.31*** (0.11)	0.18* (0.10)	-0.24* (0.12)	-0.21** (0.09)	11383
(c) average no. faculties in $-i$ (int. margin)	1.27*** (0.32)	0.07 (0.12)	-0.86 (0.53)	-0.47 (0.30)	4064
(d) average no. of universities in $-i$	0.43*** (0.10)	0.21** (0.10)	-0.82*** (0.27)	-0.82*** (0.27)	11383
province FE	yes	yes	yes	yes	
year fixed effects	no	yes	no	no	
region-by-year FE	no	no	yes	yes	
provincial controls	no	no	no	yes	
Panel B - Pairwise approach					
(a) no. of faculties in j	0.37*** (0.06)	-0.03 (0.03)	-0.14** (0.03)	-0.08*** (0.04)	50803
(b) at least one faculty in j	0.35*** (0.06)	0.13*** (0.04)	-0.07** (0.03)	-0.04* (0.02)	50803
(c) no. of faculties in j (int. margin only)	0.73*** (0.15)	0.02 (0.05)	-0.16** (0.06)	-0.05*** (0.02)	22728
(d) no. of universities in j	0.25*** (0.06)	0.10** (0.04)	-0.13** (0.06)	-0.13** (0.06)	50803
provincial pair FE	yes	yes	yes	yes	
year fixed effects	no	yes	no	no	
region-by-year FE	no	no	yes	yes	
provincial controls	no	no	no	yes	

Notes: The four columns refers to different OLS FE specifications. In Column [1], we use province FE. In Column [2], we add year FE. Column [3] uses region-by-year fixed effects. Column [4] adds province level controls. In Rows (a) and (c), the dependent variable is the total number of faculties in province i . In Row (b), the dependent variable is a dummy equal to 1 if at least one faculty is active in province i and 0 otherwise. In row (c), we analyse the intensive margin, i.e., we restrict the analysis to provinces with a positive number of faculties (we exclude observations with zero faculties). In Row (d), the dependent variable is the total number of universities active in province i . The set of controls includes the total number of universities (not included in specification (d)), the number of A-level universities, and private universities in province i . Standard errors clustered by province are reported in parentheses. The difference in the number of observations between Panels A and B is due to the fact that, in the pairwise analysis, in any period for each province we have as many observations as the number of its pairs. Significance levels: * : 10% ** : 5% ***: 1%.

with supply in province i for all measures in Rows (a)-(d). As expected, accounting for time invariant omitted factors that pertain to each neighbouring pair (i.e., the pair fixed effects) reduces the size of the coefficients.

Tables 5 and 6 present the results of the two-stage IV FE model. We use the baseline specification with the complete set of fixed effects and provincial controls (Rows (a), Column [4] of Table 4). In Panels A and B, we report estimates for the average neighbour and pairwise approaches, respectively. In Column [1], we present the results for all 110 provinces and set the initial conditions of the 41 provinces created in Italian history to zero. In Column [2], we consider only the 69 preunitarian provinces to focus on the meaningful set of initial conditions IC_{j0} in equation (3), while in Column [3], we allow for a ten-year lagged effect of the higher education supply of neighbours. We implement a feasible GMM estimator.²⁶

Table 5 presents the first-stage estimates. The results confirm that reforms interacted with neighbours' initial conditions have a positive effect on their own supply of faculties. In Panel A, most of the action comes from the 1969 reform, which liberalised university access, while the contribution of the other reforms through the initial conditions is not estimated very precisely. This implies that particular care is needed to interpret the second-stage coefficients of the average neighbour model, as identification in the first stage might be tenuous. In Panel B, most reforms contribute to identification instead. The 1969 education reform is again the most relevant one, alongside the 1923 and 1933 reforms.²⁷ Summing the significant coefficients in Panel B suggests that the reform effort that took place during Italian history induced one neighbouring province with some preunitarian supply to open roughly one faculty on average. Table A-VIII of the online appendix shows that the effect of the instruments on the supply of faculties of neighbours is homogenous across different subgroups of regions, quartiles of the provincial population, and the quality of higher education (according to the Casati ranking) in province i .

Table 6 presents the results for the second stage. In Column [1], we report OLS FE results from Table 4 for ease of comparison. F-statistics confirm that instruments have strong predictive power in the first stage of pairwise estimates and much less so under the average neighbour approach. Instruments also pass the overidentification test. We cluster standard errors at the province level, and we verify that efficiency and instruments' power

²⁶This is more efficient than using the 2SLS estimator. We also checked the robustness of our results to a ten-year lagged effect of the instruments on the higher education supply of neighbours in the first stage to account for the fact that reforms may have a delayed effect on higher education supply. See Table A-XI (column [5]) in the online appendix for details.

²⁷Estimates in Table A-XIII in the on-line Appendix suggests that the coefficient of the 1969 reform in the first stage becomes smaller as we include in the first stage also the 1961 reform that allowed students from technical schools to enroll into college (see Bianchi, 2020 for details). Second stage results are unaffected.

Table 5: First stage of IV estimates: initial conditions, reforms and higher education supply.

	[1]	[2]	[3]
Panel A - Average neighbour approach			
$(IC_{-i})^*(L. 2102/1923)$	0.19 (0.16)	0.21 (0.16)	0.22 (0.16)
$(IC_{-i})^*(L. 1592/1933)$	-0.04 (0.10)	0.24 (0.18)	0.25 (0.17)
$(IC_{-i})^*(L. 910/1969)$	0.21*** (0.05)	0.31*** (0.08)	0.30*** (0.08)
$(IC_{-i})^*(L. 766/1973)$	0.09** (0.04)	0.06* (0.04)	0.06* (0.04)
$(IC_{-i})^*(L. 382/1980)$	0.07 (0.04)	0.03 (0.05)	0.03 (0.05)
$(IC_{-i})^*(L. 168/1989)$	0.04 (0.03)	0.04 (0.05)	0.04 (0.05)
$(IC_{-i})^*(L. 245-341/1990)$	0.19*** (0.10)	0.07 (0.10)	0.06 (0.10)
$(IC_{-i})^*(L. 59/1997)$	0.26*** (0.09)	0.20 (0.16)	0.09 (0.12)
Observations	11383	8626	8614
Panel B - Pairwise approach			
$(IC_j)^*(L. 2102/1923)$	0.20*** (0.05)	0.23*** (0.04)	0.22*** (0.04)
$(IC_j)^*(L. 1592/1933)$	0.14*** (0.03)	0.15*** (0.02)	0.16*** (0.02)
$(IC_j)^*(L. 910/1969)$	0.24*** (0.02)	0.22*** (0.02)	0.22*** (0.02)
$(IC_j)^*(L. 766/1973)$	0.06*** (0.02)	0.07*** (0.02)	0.07*** (0.02)
$(IC_j)^*(L. 382/1980)$	0.05 (0.03)	0.01 (0.03)	0.01 (0.03)
$(IC_j)^*(L. 168/1989)$	0.04** (0.01)	0.02 (0.01)	0.02 (0.01)
$(IC_j)^*(L. 245-341/1990)$	0.19*** (0.05)	0.14*** (0.04)	0.13*** (0.04)
$(IC_j)^*(L. 59/1997)$	0.27*** (0.05)	0.17*** (0.06)	0.13*** (0.04)
R sq.	0.91	0.91	0.93
Observations	50803	35510	32644

Notes: First stage IV FE estimates reported in Table 6. The dependent variable is F_{-it} in Panel A and F_{jt} in Panel B. We use the baseline specification with the complete set of fixed effects and provincial controls (Row (a), Column [4] of Table 4). In Column [1], the results are for all 110 provinces. In Column [2], the results pertain to the 69 preunitarian provinces. Column [3] allows for a 10 years' lagged effect of neighbours. All specifications include the usual set of fixed effects and provincial controls. Standard errors clustered by province are reported in parentheses. Significance levels: * : 10% ** : 5% *** : 1%.

Table 6: IV estimates: neighbourhood effects in the higher education supply (2^{nd} stage).

	[1] baseline sample		[2] preunitarian provinces		[3] preunitarian prov. (lagged 10 y.)	
	OLS FE	IV FE	OLS FE	IV FE	OLS FE	IV FE
Panel A - Average neighbour approach						
no. of faculties in $-i$	-0.48*** (0.14)	-1.14*** (0.39)	-0.09 (0.08)	-1.08** (0.49)	-0.11 (0.10)	-0.97* (0.52)
Observations	11383	11383	8626	8626	8026	8026
K-P rk Wald F-stat		7.167		3.838		3.884
Hansen J-stat (p-value)		0.366		0.778		0.759
Panel B - Pairwise approach						
no. of faculties in j	-0.08*** (0.03)	-0.16*** (0.05)	-0.04** (0.02)	-0.13** (0.06)	-0.05* (0.02)	-0.13*** (0.05)
Observations	50803	50803	35510	35510	32644	32644
K-P rk Wald F-stat		35.474		25.864		22.188
Hansen J-stat (p-value)		0.977		0.507		0.830

Notes: Regressions in column [1] exploit the baseline sample, i.e., 110 provinces for the entire period of 1871-2010. IV FE estimates are performed by feasible GMM estimation. Regressions in column [2] cover only the set of 69 preunitarian provinces for the entire period. Regressions in column [3] cover preunitarian provinces for the entire period and take 10 years lags of regressor and controls. Specifications in Panel A include province fixed effects. Specifications in Panel B include province pair fixed effects. All specifications include region-by-year fixed effects and the usual set of provincial controls. In IV estimates, for the total number of faculties in province $-i$, j is instrumented by the initial conditions (i.e., the number of faculties in $-i$, and j in the first year the province joins the newly formed Italian State) interacted by a battery of dummies for higher education reforms in Italy. The full set of first stage coefficients is shown in Table 5. The full set of second stage coefficients is shown in Table A-2 in the appendix. Standard errors clustered by province are reported in parentheses. Significance levels: * : 10% ** : 5% ***: 1%.

in the first stage increase once we account for spatial dependence in the standard errors.²⁸ The results show a negative effect of faculties in the neighbourhood on the local supply of faculties. This is more precisely estimated and stable under the pairwise model than under the average neighbour model. Pairwise IV results in column [3] suggest that having one faculty more in province j is associated with -0.13 local faculties in province i , on average. This is evidence of a sizable “displacement effect” of local higher education supply triggered by an increase in the supply of neighbouring provinces. We can better gauge the magnitude of this effect considering that each Italian province has four neighbours, on average. Estimates in column [3] suggest that a one standard deviation increase in higher education supply in each neighbour j (4.43 faculties) results in $-2.31 (= -0.13 * 4.43 * 4)$ faculties in province i . This is more than half of a standard deviation decrease in local supply (cfr. Table 2).

Finally, notice that the IV FE coefficient in column [3] is more than double (and more precisely estimated) than the corresponding OLS FE coefficient. This suggests that unobserved, time-varying factors may attenuate OLS estimates of neighbourhood effects. Examples include administrative agreements, political agreements, or the creation of forty-one new provinces in Italian history that reduce competition between existing i - j couples. We investigate some of these hypotheses in the sensitivity analysis below.

4.3 Sensitivity analysis of neighbourhood effects

Before moving to a discussion of welfare, we present an extensive sensitivity analysis to our baseline findings of neighbourhood effects. The results below are based on our preferred pairwise specification in Table 6 (cfr. Panel B, Column [3]).

4.3.1 Demand for university education

In Table 7, we refine the analysis to better account for changes in the structure of the economy overtime and the population and labour supply at the province level, which are important determinants of local demand for higher education.

In Panel A, we report the results from the benchmark pairwise specification for comparison purposes. In Panel B, we show that our estimates are not altered when we exclude the 14 provinces hosting a metropolitan city identified by the Italian Constitution (i.e., Bari, Bologna, Catania, Firenze, Genova, Messina, Milano, Napoli, Palermo, Reggio Calabria,

²⁸We present results of alternative clustering methods in Table A-VI in online appendix C2. We implement the methodology developed by Driscoll and Kraay (1998a) for average neighbour estimates and that developed by Cameron et al. (2006) for pairwise estimates.

Table 7: Neighbourhood effects: local demand controls

	[1] OLS FE	[2] IV FE	[3] obs.
Panel A - Baseline specification			32644
no. of faculties in j	-0.05* (0.02)	-0.13*** (0.05)	
K-P rk Wald F-stat		22.188	
Hansen J-stat (p-value)		0.830	
Panel B - Drop metropolitan cities			21412
no. faculties in j	-0.11** (0.04)	-0.22*** (0.08)	
K-P rk Wald F-stat		37.694	
Hansen J-stat (p-value)		0.418	
Panel C - Control for population size			32644
no. faculties in j	-0.04 (0.02)	-0.12*** (0.04)	
K-P rk Wald F-stat		20.845	
Hansen J-stat (p-value)		0.852	
Panel D - Weight by population size			32644
no. faculties in j	-0.08*** (0.02)	-0.18*** (0.05)	
K-P rk Wald F-stat		24.703	
Hansen J-stat (p-value)		0.508	
Panel E - Control for share of 0-14 year olds and with higher education			11442
no. faculties in j	-0.06** (0.03)	-0.12* (0.06)	
K-P rk Wald F-stat		31.842	
Hansen J-stat (p-value)		0.796	
Panel F - Control for participation and size of the industry sector			32644
no. faculties in j	-0.05* (0.02)	-0.12*** (0.05)	
K-P rk Wald F-stat		22.684	
Hansen J-stat (p-value)		0.809	
Panel G - Placebo: “alphabetical” neighbours			45269
no. faculties in j	0.008 (0.019)	0.066 (0.063)	
K-P rk Wald F-stat		73.895	
Hansen J-stat (p-value)		0.325	

Notes: In Panel A, baseline pairwise estimates are as in Table 6, Panel B, column [3]. In Panel B, provinces with metropolitan cities (Bari, Bologna, Cagliari, Catania, Firenze, Genova, Messina, Milano, Napoli, Palermo, Reggio Calabria, Roma, Torino, and Verona) are dropped from the sample. In Panel C, the total population is included in the set of controls. In Panel D, the total population is used as a weight. In Panel E, the share of the population belonging to the 0-14 cohort and the share of the population with higher education are included among the controls (available only for 1951-2010). In Panel F, the share of people active in the industry sector and the participation rate are included among the controls. In Panel G, neighbours are defined as provinces whose name starts with the same letter of the alphabet. All specifications include provincial pair fixed effects, region-by-year fixed effects, and the usual set of provincial controls. All specifications consider 10-year lags of the regressor and controls. In IV estimates, the instruments are interactions of initial conditions with higher education reforms. The full set of coefficients is shown in Table A-4. Standard errors are clustered at the province level. Significance levels: * : 10% ** : 5% ***: 1%.

Roma, Torino and Verona). This reassures us that our results are not driven by omitted demand for higher education, e.g., determined by increasing urbanisation during the studied time period. Changes in the demographic and economic structure of provinces may also matter. Next, we consider province-level indicators drawn from the Italian census data.²⁹ In Panel C, we show that the results persist once we include as a control the total population (in log). In Panel D, we confirm the results when we use the total population as a weight of each individual province observation instead. In Panel E, we include the population share in the 0–14 cohort and the population share with tertiary education. We were able to recollect this information only from 1951 onwards and not for all provinces. This produces a substantial loss of observations, which slightly reduces the precision of the estimated coefficients but confirms the main results. Demand for higher education may also follow the expansion of the industrial sector and labour market participation during the 20th century. In Panel F, we account for this and include the participation rate and the population share active in the industry sector. Our results are also robust to the inclusion of these controls.³⁰ Finally, in Panel G, we perform a placebo exercise to check whether our results truly capture neighbourhood effects. We impose a nonsense “alphabetical contiguity” relationship and define provinces as neighbours if their name starts with the same letter of the alphabet, regardless of their actual geographical location (e.g., Alessandria is coded as a neighbour of Agrigento, despite being located over 1000 km away). The estimates show that the negative coefficient of the number of faculties in j now disappears. This reassures us that the negative coefficient estimated in our main analysis truly relates to geographical proximity.

4.3.2 Alternative IV specifications and estimators

We now present results from an extensive set of sensitivity tests to our exposure research design, in the spirit of Goldsmith-Pinkham et al. (2020).

In Table 8, we consider alternative combinations of initial conditions and reforms to minimise the exposure of Z_{jt} to omitted factors for the 150 years of Italian history. In Column [1], we confirm our results considering only the interaction between the instrument and the 1969 reform, which displays a more sizeable impact on first-stage estimates in Table 5. The larger F-stat compared to the baseline regression is consistent with the pivotal role this reform played in the expansion of higher education in Italian history. However, the

²⁹These data are gathered once every ten years in accordance with the census. We assigned each province-level data point to the corresponding year and successive ones until the next data point was available from the corresponding census wave.

³⁰In the online appendix, Table A-XI (column [3]), we check the robustness of our results to the inclusion of region-by-year dummies and the number of universities (total, A-level, and private) of the neighbour.

Table 8: Alternative distances between IC and reforms

	[1] IV FE	[2] IV FE	[3] IV FE	[4] IV FE	[5] IV FE	[6] IV FE
no. of faculties in j	-0.19** (0.07)	-0.18** (0.08)	-0.19** (0.07)	-0.19** (0.09)	-0.17** (0.07)	-0.11* (0.06)
Reforms	1969 only	1923 only	1933 only	1969 only	1969 only	1969 only
Initial conditions	1861	1861	1861	1861	1938	1938
Time Span	1870-2010	1870-2010	1870-2010	1950-2010	1950-2010	1964-1974
First stages						
$(IC_j^{1861})*(L. 1969)$	0.55*** (0.06)					
$(IC_j^{1861})*(L. 1923)$		0.49*** (0.06)				
$(IC_j^{1861})*(L. 1933)$			0.50*** (0.06)			
$(IC_j^{1938})*(L. 1969)$				0.32*** (0.04)	0.26*** (0.03)	0.09*** (0.01)
K-P rk Wald F-stat	72.594	57.635	62.259	74.388	88.567	106.942
Observations	32644	32644	32644	14086	22148	3375

Notes: All specifications include the usual set of controls for region-by-year and pair fixed effects. All specifications consider 10 years lags of the regressor and controls, and the time span reported in the table refers to first stage estimates. Standard errors clustered by province are reported in parentheses. Significance levels: * : 10% ** : 5% ***: 1%

time between 1861 and 1969 is considerable, and many events happened over these years. The following columns address this issue. In Columns [2] and [3], we reduce the distance between initial conditions and reforms by taking as unique shocks the major reforms of 1923 and 1933, respectively. The signs, size and significance of the coefficients are preserved, as well as the power of the instrument for the first stage. Next, we consider again the 1969 reform but shorten the time span and the distance between initial conditions and reforms. In Column [4], we focus on the post-WWII period only, i.e., 1950-2010. In Column [5], we consider the 1950-2010 period and take as initial conditions the number of faculties in the province in 1938 (i.e., the year before the start of WWII). In Column [6], we narrow the time period to 1964-1974, i.e., five years before and five years after the 1969 reform. Our results focus on these cases.

In Table 9, we check the robustness of our results obtained when we implement the following shift-share version of our instrument:

$$B_{jt} = \sum_k ic_{jk0} * g_{kt}, \quad (4)$$

where ic_{jk0} is the share of faculties in the field of study $k = 1...K$ in province j at time

0, and g_{kt} is the growth rate of the number of faculties in the field of study k at time t . In column [1], we show results for $K = 1$, i.e., when B_{jt} exhibit a similar variation structure as Z_{jt} . In this case, ics feature location shares at time 0, while g is the growth rate in the total number of faculties without distinguishing by field of study. In Column [2], we report results obtained when B_{jt} is the inner product of the discipline-location shares and the discipline specific growth rates when $K = 3$, i.e., STEM, Social Sciences and Humanities. In Column [3], we directly compare the relevance of reforms to their discipline-specific realisations by including both Z_{jts} and B_{jt} in the first stage. For comparison purposes, in Column [4], we consider the baseline first stage with the Z_{jts} only. The high F-statistics in Columns [1] and [2] suggest that B_{jt} has high explanatory power in the first stage.³¹ Additionally, the estimated coefficient of B_{jt} is large and significant in Columns [1] and [2] but becomes five times smaller in Column [3]. Moreover, the lower F-stats in Columns [3] and [4] suggest that including several instruments at the same time weakens the power of the first stage. Accordingly, alongside standard GMM estimates, in both columns, we also report Fuller (FUL) estimates and the GMM generalisation of the limited information maximum likelihood estimator (CUE/LIML), which are robust to many potentially weak instruments. Overall, the size and significance of the second-stage estimates across Columns [1]-[4] are preserved and confirm the baseline results.

In Section C of the online appendix, we report the results of other important robustness checks of our IV estimator. Thus far, we have allowed for the territories and borders of the 69 provinces that existed as administrative units in 1861 to change alongside the creation of 41 new provinces over Italian history. This modelling choice allows us to very precisely measure local neighbourhood effects. However, it may not fully serve the purpose of exogeneity, as governance decisions regarding provinces' territories and borders may depend on the territorial provision of higher education services. In Table A-II, we check the robustness of our results to alternative definitions of preunitarian provinces that preserve their preunitarian territory and borders. In Table A-X, we verify that our results are not driven by specific subperiods or regional outliers and drop twenty-year time spans and groups of regions for at a time from the sample. In Table A-VI, we verify that our results hold as we account for potential spatial dependence using Driscoll and Kraay (1998b) and two-way clustered standard errors. In Table A-VII, we confirm our results, as we implement the correction developed by Adão et al. (2020) to account for arbitrary correlation in the error term across provinces with similar initial conditions.³² In Table A-XI, we confirm our result after a final

³¹In Table A-IX in online appendix C, we show that this is generally the case, even under the average neighbourhood approach, as we replicate baseline Table 6 using the shift-share $\bar{B}_{-it} = \sum_{j \neq i}^{N_i} w_{ij} B_{jt}$.

³²As the correction developed by Adão et al. (2020) requires a specification with only one instrument, we apply it using the standard shift-share approach, which uses as instrument B_{jt} with $K=1,3$. The estimates

Table 9: Comparison to the shift-share instrument using multiple estimators

	[1] IV FE	[2] IV FE	[3] IV FE			[4] IV FE		
			GMM	FUL	CUE/LIML	GMM	FUL	CUE/LIML
no. of faculties in j	-0.18** (0.08)	-0.14** (0.07)	-0.13*** (0.05)	-0.19** (0.07)	-0.09** (0.03)	-0.13*** (0.05)	-0.19** (0.07)	-0.12*** (0.04)
Hansen J-stat (p-value)	—	—	0.528	0.528	0.648	0.830	0.830	0.839
First stages								
B_{jt} (K=1)	728.63*** (100.97)							
B_{jt} (K=3)		186.56*** (22.97)		36.87*** (12.38)				
$(IC_j^{1861})^*(L. 1923)$				0.21*** (0.04)			0.22*** (0.04)	
$(IC_j^{1861})^*(L. 1933)$				0.16*** (0.02)			0.16*** (0.02)	
$(IC_j^{1861})^*(L. 1969)$				0.19*** (0.02)			0.22*** (0.02)	
$(IC_j^{1861})^*(L. 1973)$				0.11*** (0.02)			0.07*** (0.02)	
$(IC_j^{1861})^*(L. 1980)$				-0.01 (0.03)			0.01 (0.02)	
$(IC_j^{1861})^*(L. 1989)$				0.01 (0.01)			0.02 (0.01)	
$(IC_j^{1861})^*(L. 1990)$				0.12*** (0.03)			0.13*** (0.04)	
$(IC_j^{1861})^*(L. 1997)$				0.14*** (0.03)			0.13*** (0.04)	
K-P rk Wald F-stat	52.071	65.94		19.958			22.188	
Observations	32644	32644		32644			32644	

Notes: All specifications include the usual set of controls for region-by-year and pair fixed effects. All specifications consider 10-year lags of the regressor and controls. The total number of faculties in province j is instrumented by the initial conditions interacted with the national growth rates of faculties. In the computation of the shift-share instrument, in Column [1] we consider the total number of faculties in j without distinguishing by field of study (i.e., $K = 1$). In columns [2] and [3], we consider the number of faculties in j by field of study and distinguish between the STEM, Social Sciences, and Humanities disciplines (i.e., $K = 3$). Standard errors clustered by province are reported in parentheses. Significance levels: * : 10% ** : 5% *** : 1%.

set of robustness checks.³³

4.3.3 Extensions

Before moving to welfare analysis, we propose three insightful extensions to the baseline results on neighbourhood effects.

Cross-disciplinary neighbourhood effects In Table 10, we compare neighbourhood effects within and between fields of study. In practice, we group all faculties into three major groups and run three sets of regressions (Panels A-C) where our dependent variables are the local supply in the Humanities, STEM, and the Social Sciences, respectively. In Columns [1] and [3], we show OLS FE and IV FE estimates, where the explanatory variable is the neighbour’s higher education supply in the same field of study. In Columns [2] and [4], we include the neighbour’s supply in the other two fields. To ease interpretation, Column [5] reports the same estimates as Column [4] with standardised coefficients. The results show that neighbourhood effects are concentrated within the same field of study, while cross-field effects are never statistically significant. This supports the view that higher education provisions are substitutes within the same field of study only such that having, e.g., a STEM faculty in province j displaces STEM higher education supply in province i but not local supply in the Humanities or Social Sciences.

Historical and regional factors In Table 11, we report results from heterogeneity exercises based on important historical and regional factors. In Panel A, we consider WWII, which also delimits the prerepublican and republican eras of Italian history. In Column [1], we consider the prewar period as before 1936, the year of the Rome-Berlin Axis treaty. In Column [2], we consider the postwar period as after 1949, the year of the North Atlantic treaty. The results hold for both periods, although the negative coefficient is more precisely estimated for the postwar period, where the rise of higher education supply is concentrated (cfr. Figure 2). Second, we split our sample into northern and central-southern regions, as these are characterised by long-lasting differences in economic and institutional development (Panel B). The results in Columns [1] and [2] confirm the size and significance of the baseline result for both macroregions, particularly for IV estimates. In Panel C, we argue that each

in Table A-VII confirm those reported in Table 9, Columns [1] and [2].

³³We include among the set of instruments the 2nd degree spatial lag of the higher education supply of the local province in the spirit of De Giorgi et al., 2010. We also consider a ten-year lagged effect of instruments in the first stage, extend the set of controls for neighbouring provinces and regions, and account for the possibly confounding role of multicampus universities.

Table 10: Cross-disciplinary neighbourhood effects: humanities, STEM, social sciences

	[1]	[2]	[3]	[4]	[5]	[6]
	OLS FE	OLS FE	IV FE	IV FE	IV FE	Obs.
Panel A - Humanities						32644
no. of faculties _j , Humanities	-0.06*	-0.07*	-0.02	-0.07*	-0.08*	
	(0.04)	(0.04)	(0.07)	(0.04)	(0.04)	
no. of faculties _j , STEM		0.01		-0.02	-0.03	
		(0.02)		(0.03)	(0.05)	
no. of faculties _j , Social Sciences		0.01		-0.00	-0.01	
		(0.02)		(0.01)	(0.02)	
K-P rk Wald F-stat			4.879	11.141	11.141	
Hansen J-stat (p-value)			0.365	0.255	0.255	
Panel B - STEM						32644
no. of faculties _j , STEM	-0.09***	-0.11***	-0.08*	-0.05**	-0.09**	
	(0.03)	(0.03)	(0.04)	(0.02)	(0.04)	
no. of faculties _j , Humanities		0.02		-0.03	-0.03	
		(0.02)		(0.02)	(0.02)	
no. of faculties _j , Social Sciences		0.02		0.00	0.00	
		(0.02)		(0.01)	(0.02)	
K-P rk Wald F-stat			24.294	11.141	11.141	
Hansen J-stat (p-value)			0.675	0.781	0.781	
Panel C - Social Sciences						32644
no. of faculties _j , Social sciences	-0.04*	-0.04	-0.16**	-0.07**	-0.12**	
	(0.02)	(0.03)	(0.07)	(0.03)	(0.05)	
no. of faculties _j , STEM		0.02		-0.01	-0.03	
		(0.03)		(0.02)	(0.04)	
no. of faculties _j , Humanities		-0.01		0.03	0.03	
		(0.04)		(0.02)	(0.03)	
K-P rk Wald F-stat			23.301	11.141	11.141	
Hansen J-stat (p-value)			0.904	0.498	0.498	

Notes:. Baseline pairwise estimates are shown as in Table 6, Panel B, Column [3]. OLS FE estimates are shown in Columns [1] and [2]. IV FE estimates are shown in Column [3] with one endogenous regressor, i.e., the no. of faculties in the respective discipline (Humanities in Panel A, STEM in Panel B, and Social Sciences in Panel C), in the neighbouring province. In Column [4], IV FE estimates with three endogenous regressors are shown, i.e., the number of Humanities, STEM and Social Sciences faculties in the neighbouring province. The specification in Column [5] is the same as that in Column [4], but regressors are standardised to have a zero mean and unity standard deviation. Among IV estimates, the instruments are interactions of initial conditions (the total number of faculties in each discipline in the first year the province joins the newly formed Italian State) with a battery of dummies for higher education reforms in Italy. See Table A-3 for the first stage estimates. All specifications include provincial pair fixed effects, region-by-year fixed effects, and the usual set of provincial controls. Standard errors clustered by province are reported in parentheses. Significance levels: * : 10% ** : 5% *** : 1%.

Table 11: Heterogeneous neighbourhood effects: historical and regional factors

Panel A - Pre-WWII vs. post-WWII period				
	[1] prewar		[2] post-war	
	OLS FE	IV FE	OLS FE	IV FE
no. of faculties in j	-0.02 (0.02)	-0.13 (0.10)	-0.04* (0.02)	-0.10** (0.05)
Observations	14867	14867	14330	14330
K-P rk Wald F-stat		26.836		30.752
Hansen J-stat (p-value)		—		0.756
Panel B - North vs. Centre-South				
	[1] North		[2] Centre-South	
	OLS FE	IV FE	OLS FE	IV FE
no. of faculties in j	-0.05 (0.03)	-0.14** (0.07)	-0.03 (0.03)	-0.09** (0.04)
Observations	16687	16687	15957	15957
K-P rk Wald F-stat		18.982		35.290
Hansen J-stat (p-value)		0.652		0.439
Panel C - Intraregional vs. interregional competition				
	[1] intraregional		[2] interregional	
	OLS FE	IV FE	OLS FE	IV FE
no. of faculties in j	-0.11** (0.05)	-0.19*** (0.06)	0.03 (0.03)	-0.06 (0.07)
Observations	18832	18832	13683	13683
K-P rk Wald F-stat		21.096		11.696
Hansen J-stat (p-value)		0.649		0.957

Notes: Baseline pairwise estimates are shown. In Panel A, pre-WWII and post-WWII periods are defined as before 1936 and after 1949, respectively. In Panel B, northern regions belong to NUTS1 “Northeast” and “Northwest”, while Centre-South regions belong to NUTS1 “Centre”, “South” and “Islands”. In Panel C, intraregional competition occurs between provinces in the same NUTS2 region, while interregional competition occurs between provinces belonging to different NUTS2 regions. In IV estimates, the instruments are interactions of initial conditions (total number of faculties in the first year the province joins the newly formed Italian State) with higher education reforms in Italy in each subperiod. First stage estimates are available upon request. All specifications include provincial pair fixed effects, region-by-year fixed effects, and the usual set of provincial controls. Standard errors are clustered at the province level. Significance levels: * : 10% ** : 5% *** : 1%

neighbouring province j may or may not belong to the same administrative region as region i . Thus, we analyse interactions between provinces that belong to the same region in Column [1] and interactions between provinces that belong to different regions in Column [2]. Interestingly, neighbourhood effects are mostly concentrated within the same region. This suggests that substitutability between faculties is lower when they are located in different regions even though they are in neighbouring provinces.

The 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. In Table 12, we propose a complementary definition of a neighbourhood based on the linear or travel distance $d_{i,j}$ 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 $d_{i,j} < D$, where D is a supposed distance bandwidth for neighbourhood effects (see Parchet, 2018). Then, in Panel A, we consider travel distance (by car) with bandwidths $D = 120, 240, 360, 480$ (see Columns [1]-[4]). The results suggest that neighbourhood effects are concentrated within a bandwidth of 120 kilometres of travel. This is consistent with the evidence given in Figure 3 above showing that the majority of Italian students choose a higher education institution within 150 km of their residence. As we measure travel as of today, estimates in Panel A do not account for the changing state of travel infrastructures (i.e., roads) over the sample period. In Panel B, we consider i-to-j linear distances instead of bandwidths $D = 90, 180, 270, 360$ km. In this case, neighbourhood effects are concentrated within a linear distance of 90 kilometres. This is the “spatial reach” of neighbourhood effects in linear distance terms. Negative coefficients also appear over longer distances, but they are much smaller and generally nonsignificant. In the online appendix, we report results as we provide complementary definitions of neighbourhoods, e.g., combining information on distance and contiguity or accounting with spatial restrictions imposed by Italian geography.³⁴

³⁴In Table A-XI, we focus on contiguous provinces in the spatial reach of 90 km in linear distance terms (cfr. column 6). In Table A-IV, we propose three schemes that weight provinces j by distance in a neighbourhood matrix featuring a gravity model. In Table A-III, we presents results for the spatial reach of neighbourhood effects as we define $d_{i,j}$ and D s in terms of (travel) time rather than distance. Finally, in Table A-V, we deal with spatial restrictions related to the Italian territory, i.e., we drop internal Italian provinces, and account for the supply of higher education of neighbouring provinces from foreign countries.

Table 12: Spatial reach of neighbourhood effects (in Km)

Panel A - Travel distance									
	[1] within 120 km		[2] between 120 and 240 km		[3] between 240 and 360 km		[4] between 360 and 480 km		
	OLS FE	IV FE	OLS FE	IV FE	OLS FE	IV FE	OLS FE	IV FE	
no. of faculties in j	-0.046** (0.018)	-0.136*** (0.044)	-0.019 (0.014)	0.021 (0.014)	-0.016 (0.012)	-0.024* (0.013)	0.009 (0.011)	-0.002 (0.007)	
Observations	40069	40069	91090	91090	100647	100647	78450	78450	
K-P rk Wald F-stat		37.398		151.815		155.445		75.294	
Hansen J-stat (p-value)		0.217		0.318		0.508		0.443	
Panel B - Linear distance									
	[1] within 90 km		[2] between 90 and 180 km		[3] between 180 and 270 km		[4] between 270 and 360 km		
	OLS FE	IV FE	OLS FE	IV FE	OLS FE	IV FE	OLS FE	IV FE	
no. of faculties in j	-0.039** (0.018)	-0.095** (0.039)	-0.045*** (0.014)	0.003 (0.017)	0.003 (0.007)	-0.014 (0.010)	-0.013 (0.011)	0.001 (0.010)	
Observations	43500	43500	88512	88512	99184	99184	77114	77114	
K-P rk Wald F-stat		45.850		149.615		167.054		85.479	
Hansen J-stat (p-value)		0.239		0.739		0.414		0.176	

Notes:. Baseline pairwise estimates. In IV estimates, the total number of faculties in province j is instrumented by the initial conditions interacted by higher education reforms in Italy. Standard errors clustered by province are reported in parentheses. Significance levels: * : 10% ** : 5% *** : 1%.

5 Welfare analysis

We are now ready to incorporate the neighbourhood effects into an evaluation of the welfare effects of higher education supply at the province level. By doing so, we quantify total province-level returns and separate local returns from geographical externalities ‘net’ of displacement costs.

5.1 The economic returns of higher education supply

We now estimate the following model:

$$V_{it} = \phi_1 F_{it-10} + \phi_2 F_{-it-10} + X_{it-10} \phi_3 + \varphi_i + \xi_{r(i),t} + \mu_{it}. \quad (5)$$

V_{it} is a function that proxies the (natural) log transformation of gross income per capita in province i at time t , in real terms, Y_{it} . In most specifications, we simply set $V_{it} = \ln(1 + Y_{it})$. Alternatively, we use the inverse hyperbolic sine transformation of Y_{it} , which removes the additive constant from the $\ln()$ function. F_{it-10} is the total number of faculties in province i at time $t - 10$. F_{-it-10} is the total number of faculties in the neighbourhood of province i at time $t - 10$. X_{it-10} is a vector of province-level controls, including the rate of population growth and the shares of the active population in the industry and agricultural sectors in province i at time $t - 10$.³⁵ All information about income per capita and province-level controls come from the Italian census data, which are collected roughly every 10 years. This means that for the 150-year sample period, we observe these province-level variables at most 15 times, i.e., we have an unbalanced panel with $T = 15$ covering the period of 1870-2010. Accordingly, we include 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 are ϕ_1 and ϕ_2 : these measure income semielasticities to have one more faculty in province i and its neighbourhood, respectively. Two endogeneity issues complicate their estimation. First, omitted factors may motivate both an increase in income per capita and the opening of faculty. Second, reverse causality goes from economic performance to the supply of faculties: richer and/or more productive provinces may express more (or less) demand for higher education. Thus, the opening of a faculty may result from this demand.

³⁵In their analysis of the human capital determinants of the industrial revolution in Prussia, Becker et al. (2011) consider the share of the population active in the industry sector as a proxy of industrial growth. In this setting, we consider this an important control, as our time period occurs after this industrial revolution.

We mitigate reverse causality concerns by taking ten-year lagged values of higher education supply indicators in equation (5). More importantly, we implement the same IV approach as before and use initial conditions to measure the exposure of preunitarian provinces to postunification university reforms. In practice, we construct Bartik instruments $Z_{it-10} = IC_{i0} * R'_{t-10}$, $Z_{-it-10} = IC_{-i0} * R'_{t-10}$, for F_{it-10} and F_{-it-10} in equation (5). In this case, as we exploit a ten-year variation, vector R'_t may capture exposure to “reform packages” in decades where multiple reforms were implemented.³⁶

We extensively discussed the general identifying assumption behind the use of these Bartik instruments in the previous section. In equation (5), for ϕ_1 and ϕ_2 to measure the exact economic returns of the higher education supply in province i and its neighbourhood, there must not be persistent omitted factors systematically correlated with the initial distribution of faculties that also determine the semielasticity of income in province i (Goldsmith-Pinkham et al., 2020). The evidence shown in Table 3 above did not suggest a systematic correlation of preunitarian higher education supply with presample levels of economic performance or prosperity (e.g., measured in terms of real income per capita, urbanization, or the share of workers in the industrial sector). More importantly, we did not find any systematic presample trends in urban wages and urban population indicators, either in the sixty years antecedent to the Italian unification or in the previous two centuries, despite these being dense historical events, as highlighted in Section 2.1. While this evidence is reassuring, any deep underlying omitted factors possibly correlated with the cross-provincial distribution of initial conditions (e.g., culture) would still prompt an interpretation of F_i and F_{-i} in equation (5) as “proximate” (rather than exact) determinants of local economic returns.³⁷

Table 13 reports estimates from equation (5). In Columns [1]- [4], we report OLS FE and IV FE results when we consider only local faculties. Table A-5 shows that many reform packages do not have explanatory power in the first stage of IV FE estimates in Column [2] (due to likely collinearity with the FE). Thus, starting from Column [3], we use only the two most relevant instruments, i.e., $IC_i * R_{21/30}$ and $IC_i * R_{61/70}$. This greatly increases the instruments’ power in the first stage. In Column [4], we use the inverse hyperbolic sin transformation as a dependent variable. Both the OLS FE and IV FE results point to positive coefficients of higher education supply with very similar magnitudes across specifications. Next, we also include faculties in the neighbourhood in the specification. The inclusion of

³⁶More precisely, dummy $R_{21/30}$ covers the decade of 1921/30, which includes only L. 2102/1923. $R_{31/40}$ covers the decade of 1931/40, which includes L. 1592/1933, and $R_{61/70}$ covers the decade of 1961/70, while includes Law 910/1969. Conversely, $R_{71/80}$ covers 1971/80 and includes a package composed of Laws 766/1973 and 382/1980. Similarly, $R_{81/90}$ includes L. 168/1989 and Laws 245-341/1990. Finally, $R_{91/00}$ covers Laws 59-127/1997.

³⁷We thank Nico Voigtländer for making this important point.

Table 13: Economic returns of higher education supply

	[1] OLS FE	[2] IV FE	[3] IV FE	[4] IV FE ASINH(V)	[5] IV FE	[6] IV FE CUE/LIML	[7] IV FE
no. of faculties in i	0.0077*** (0.0013)	0.0069*** (0.0018)	0.0104*** (0.0035)	0.0138*** (0.0046)	0.0143*** (0.0043)	0.0140*** (0.0043)	0.0118*** (0.0017)
no. of faculties in $-i$					0.0068** (0.0034)	0.0068** (0.0033)	0.0024** (0.0011)
Reforms in 1st stage	—	all	two	two	two	two	two
F_{it-20}, F_{-it-20} in 1st stage	—	no	no	no	no	no	yes
Observations	937	937	937	937	937	937	874
K-P rk Wald F-stat		11.984	16.661	16.661	3.707	3.707	51.128
Hansen J-stat (p-value)		0.029	0.184	0.166	0.370	0.382	0.385

Notes: In Column [4], the dependent variable is the inverse hyperbolic sine transformation of real income per capita (ASINH(Y)). In all other specifications, the dependent variable is $\ln(1 + Y_{it})$. The total number of faculties in the province and in the neighbourhood are lagged 10 years. All specifications include growth in the local population and the size of the industry and agricultural sectors as controls. They also include the province FE and region by year dummies. In specifications [3]-[7], the two reform packages in vector R are $R_{23/30}$ and $R_{61/70}$ only. Standard errors clustered at the province level are always reported in parentheses. Significance levels: * : 10% ** : 5% *** : 1%.

a second endogenous regressor reduces the value of the F-stat in the baseline feasible GMM estimates in Column [5]. Thus, in Column [6], we report CUE/LIML estimates, which are median-unbiased in overidentified models. In Column [7], we include the 20-year lagged values of F_i and F_{-i} , which increases the instruments' power in the first stage. Overall, estimated coefficients ϕ_1 and ϕ_2 in Table 13 are positive, and their size and statistical significance remain broadly stable across the different specifications. In Table A-6, we further confirm our findings as we correct for potential spatial dependence (Driscoll and Kraay, 1998a); we adopt a narrower definition of neighbourhood that considers only the spatial reach of 90 km; we weight each province by its presample population level; finally, we drop provinces hosting metropolitan cities belonging to the top quartile of the initial population distribution. These two exercises in particular are meant to address the concern that larger, or economically most advanced provinces, with a historically larger higher education supply may simultaneously experience increases in income per capita (e.g., due to skill-biased economic growth induced by industrialisation after 1880) and jointly push for nationwide education reforms.³⁸

To interpret coefficients, we take as our preferred specification the one in Table 13, Column [7], which displays an F-statistic of reassuring levels. Taken at face value, these suggest that, on average, increasing higher education supply by 10% in the province (i.e., by 0.25 faculty) raises local income per capita by 0.29%(= 0.25 * 0.0118) at the sample mean (equal to 0.84). A 10% increase in the supply of higher education in the neighbourhood (i.e., by 1.2 faculties) has a quantitatively equivalent effect, as it raises local income per capita by 0.28%(= 1.2*0.0024). These magnitudes are very close to the estimated effects from existing

³⁸Similarly, local unobservables could drive both development and demand for education by these advanced provinces. Our results also hold when we gradually drop provinces characterised by the largest supplies (see Table A-XIV in the online appendix).

studies, despite the different data and estimation strategies employed (see, e.g., Valero and Van Reenen, 2019).³⁹

5.2 Effects on welfare

We are now ready to quantify the total contribution of higher education supply to the welfare of the local province, measured in income terms. To do this, we have to take into full consideration how the effects of faculties in the neighbourhood complement the effect of local faculties. Our analysis thus far has highlighted two opposite externalities. The first is measured by ϕ_2 in equation (5). This reflects a positive effect of the higher education supply of neighbouring provinces on local real income per capita, which is well known in the literature. The second is shown by $b < 0$ in equation (2). 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 equations (2) and (5). 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.:

$$\tilde{V}_i = \hat{\phi}_1 F_i + \hat{\phi}_2 F_{-i}, \quad (6)$$

where F_i , F_{-i} is the observed number of faculties in province i and its neighbourhood, and $\hat{\phi}_1$ and $\hat{\phi}_2$ are the estimated economic returns. If there were no faculties in the neighbourhood, province i would neither benefit from positive income externalities nor suffer from negative competition from those faculties. Counterfactual returns in this scenario are as follows:

$$\tilde{\tilde{V}}_i = \hat{\phi}_1 \hat{F}_i = \hat{\phi}_1 (F_i - \hat{b} \sum_{j \neq i} F_{ij}) = \hat{\phi}_1 (F_i - \hat{b} F_{-i}), \quad (7)$$

where $\hat{b} < 0$ is the neighbourhood effect predicted from equation (2), and $\hat{F}_i$ measures the counterfactual number of faculties in province i in the absence of neighbourhood effects. This is higher than observed supply, as relaxing competitive pressures from contiguous provinces acts as an incentive for local provision. Net externality Ext_i is the difference between (6)

³⁹Valero and Van Reenen (2019) measure higher education supply in university units (rather than our faculty units), within NUTS2 regions (instead of our NUTS3 units), and across several countries (rather than our focus one a single country). They estimate that increasing the supply of universities by 10% is associated with a 0.4% increase in regional GDP per capita, which is very close to our figure.

and (7):

$$Ext_i = \tilde{V}_i - \tilde{\tilde{V}}_i = (\hat{\phi}_2 + \hat{\phi}_1 \hat{b}) F_{-i}. \quad (8)$$

This is the sum between the positive value of direct welfare externalities $\hat{\phi}_2 F_{-i}$ and the negative value of displacement externality $\hat{\phi}_1 \hat{b} F_{-i}$.

We compute predicted welfare effects by incorporating $\hat{b}$ and $\hat{\phi}$ s from our preferred specifications into equations (6)-(8).⁴⁰ Table 14 summarises statistics about predicted welfare and corresponding income effects. On average, provinces benefiting from the average 2010 supply of higher education (25 faculties, of which 5 are local and 20 are in contiguous provinces) have a local real income per capita $\tilde{Y}/Y = 0.115/2.31 = 4.9\%$ higher than those that do not benefit from any faculty at all. This features nonnegligible returns of higher education supply on average. The predicted values $\hat{\phi}_1 F_i$ and $\hat{\phi}_2 F_{-i}$ in Table 14 feature an approximately equal contribution of local and neighbours' supply of faculties. However, once we correctly take into account the opportunity cost of cross-provincial displacement forces $\hat{\phi}_1 \hat{b} F_{-i}$, the net welfare contribution of local faculties is $\tilde{\tilde{V}} = 0.085$ on average. This is four times larger than the contribution of faculties in the neighbourhood, as $Ext = 0.02$. In practice, an average local supply of five faculties accounts for approximately 80% of the economic returns in a province (equal to 3.9% of local income per capita), while net externalities from an average supply of 20 faculties in the neighbourhood explain the remaining 20% (equal to 1% of income per capita of the local province).⁴¹

Table 14: Simulation exercise: summary of welfare and income effects

Variable	Mean	Std. Dev.	Min.	Max.	N
$\hat{\phi}_1 F_i$	0.055	0.074	0	0.441	110
$\hat{\phi}_2 F_{-i}$	0.05	0.035	0	0.148	110
$\hat{\phi}_1 \hat{b} F_{-i}$	-0.03	0.021	-0.094	0	110
$\tilde{V}_i = \hat{\phi}_1 F_i + \hat{\phi}_2 F_{-i}$	0.105	0.08	0	0.476	110
$\tilde{\tilde{V}}_i = \hat{\phi}_1 F_i - \hat{\phi}_1 \hat{b} F_{-i}$	0.085	0.075	0	0.461	110
$Ext_i = \hat{\phi}_2 F_{-i} + \hat{\phi}_1 \hat{b} F_{-i}$	0.02	0.014	0	0.057	110
Y_i	2.31	0.56	1.33	3.49	107
$\tilde{Y}_i = (e^{\tilde{V}_i} - 1)$	0.115	0.096	0	0.61	110
$\tilde{\tilde{Y}}_i = (e^{\tilde{\tilde{V}}_i} - 1)$	0.092	0.09	0	0.586	110

Notes: Predicted effects from equations (6)-(8) in the text. Values refer to 2010.

In Figure 4, we draw the cross-province distribution of predicted economic returns and net externalities in % of local income per capita. Panel (a) looks at the geographical distri-

⁴⁰These are IV FE estimates in Table 6, Panel B, Column [3] for equation (2) and Table 13, Column [7] for equation (5).

⁴¹In fact, on average, $\tilde{Y}/Y = 0.092/2.31 = 0.039$, while $(\tilde{Y} - \tilde{\tilde{Y}})/Y = 0.023/2.31 \approx 0.01$.

bution of actual returns. These tend to be negligible (less than 1% of local income per capita; transparent grey provinces) or small (between 1% and 3%, light grey areas) in peripheral provinces that only benefit from faculties of a few small neighbours. Small provinces with local universities enjoy local returns of up to 5% of local income per capita (grey provinces). Returns are larger in Italian provinces that benefit from externalities from major metropolitan areas (between 5% and 10%, dark grey provinces). Large provinces such as Milano, Roma and Napoli benefit from local university hubs that explain up to 25% of local income per capita (dark grey provinces).

Panel (b) shows the geographical distribution of the net externalities in % terms of local income. Everywhere in the Italian territory, the higher education supply produces nonnegative net externalities, suggesting that economic returns are larger than displacement costs in all provinces. Cross-province figures also confirm that net externalities are generally not too large. In most border provinces (particularly in the northeast), they account for less than 1% of local income per capita (transparent grey provinces). Around most urban areas, they explain between 1% and 3% of local income per capita (light grey areas). In the neighbourhood of large university centres such as Roma and Napoli, net externalities predict up to 5% of the local income per capita (grey provinces).

Overall, these results suggest that local returns from higher education supply may be sizeable, especially around major urban areas. They also suggest that caution is needed when assessing the magnitude of spillovers from a neighbourhood. Once we account for the economic cost associated with displacement effects, net returns appear strongly localised. Disregarding these forces may severely overestimate economic externalities.

6 Conclusions

In this paper, we used an own-built historical dataset for Italy with fine-grained province-level information to evaluate the welfare effects of higher education supply. We exploited variation associated with the initial conditions of university faculties and national higher education reforms and showed that a province exposed to an average supply of 25 faculties (5 local and 20 located in the neighbourhood) has a 4.9% higher real income per capita than a province exposed to a zero supply. By accounting for displacement effects between faculties located in nearby provinces, we were also able to find that economic returns of higher education supply are highly localised: on average, local supply accounts for approximately 80% of the economic returns in a province, while net externalities from faculties in the neighbourhood explain only the remaining 20%.

Figure 4: Geographical distribution of predicted returns

Notes: Values are expressed in % of local real income per capita Y_i of province i (gross terms). All values refer to 2010 levels. Information on income is not available for the provinces of Barletta, Carbonia-Iglesias, Fermo, Monza-Brianza, Medio-Campidano and Ogliastro. Detailed values by province are shown in Table A-7.

These results do not carry over to specific policy proposals regarding the territorial distribution of a country's supply. The choice between, e.g., a relatively dispersed vs. agglomerated setting, depends on underlying country- and region-level parameters such as the alignment of objectives of local and national policy-makers (Glaeser, 2010). Rather, our findings provide an economic rationale for the fact that local higher education providers better match local development needs. According to this interpretation, the economic benefits of expanding the higher education supply are to be considered inherently local, similar to the advantages of large plant openings. Moreover, these results are also consistent with the view that regional disparities in higher education supply may partly be responsible for observed cross-regional differentials in human capital, productivity and income per capita in OECD countries. In fact, we uncover very large inequalities between small peripheral provinces that enjoy limited economic returns from higher education and large provinces wherein the supply of faculties explains up to 25% of local income per capita. A possible caveat to our results is that we have found neighbourhood effects in a centralized higher education system such as the Italian one, where institutions only compete to attract enrollments from neighbouring areas. It seems important to quantify neighbourhood effects in countries where

spatial competition coexists with other competition channels e.g. decentralized tuition fees in public universities (as in the UK) or in a large private education sector (as in France). Finally, further research is needed to point out the exact mechanism(s) that determine aggregate returns, e.g., human capital accumulation, and/or changes to industry composition and technologies.

Acknowledgements

We are grateful to David De la Croix for sharing his data on academic notability and to Giovanni Peri and Nico Voigtländer for their useful suggestions. We are also indebted to Dirk Foremny and Marine de Talancé for their very helpful discussions. We also acknowledge the Istituto Tagliacarne for providing 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. 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*. We are particularly grateful to Pietro De Ponti, Ginevra Gallassi, Giuseppe Grasso, George Jacob and Roberta Terranova for providing assistance at various stages of our research. The usual disclaimer applies.

References

- Adão, R., Kolesár, M., Morales, E., and 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. Statistical Software Components, Boston College Department of Economics.
- 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.
- Aghion, P., Boustan, L., Hoxby, C., and Vandenbussche, J. (2009). The causal impact of education on economic growth: Evidence from U.S. volume 1 of *Harvard University working paper*.
- Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., and Sapir, A. (2010). The governance and performance of universities: evidence from europe and the US. *Economic Policy*, 25(1):7–59.
- Andersson, R., Quigley, J. M., and Wilhelmson, M. (2004). University decentralization as regional policy: the Swedish experiment. *Journal of Economic Geography*, 4(4):371–388.
- Becker, S. O., Hornung, E., and Woessmann, L. (2011). Education and catch-up in the industrial revolution. *American Economic Journal: Macroeconomics*, 3(3):92–126.
- Berry, S. and Reiss, P. (2007). Empirical models of entry and market structure. volume 3, chapter 29, pages 1845–1886. Elsevier, 1 edition.
- Bianchi, N. (2020). The indirect effects of educational expansions: Evidence from a large enrollment increase in university majors. *Journal of Labor Economics*, 38(3):767–804.
- Braga, M., Checchi, D., and Meschi, E. (2013). Educational policies in a long-run perspective. *Economic Policy*, 28(73):45–100.
- 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.
- Breccia, A. and Salustri, S. (2017). "New" new economic geography: firm heterogeneity and agglomeration economies. *E-Review*, 5.

- Brizzi, G.P. Del Negro, P. and Romano, A. e. (2007). *Storia delle Università in Italia*. CISUI.
- Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International Regional Science Review*, 26(2):175–188.
- Cameron, A., Gelbach, J., and Miller, D. (2006). Robust inference with multi-way clustering. NBER Technical Working Papers 0327, National Bureau of Economic Research, Inc.
- Cantoni, D. and Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, 129(2):823–887.
- 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.
- Ciccone, A. and 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.
- De Giorgi, G., Pellizzari, M., and Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2):241–75.
- De Ridder-Symoens, H. and Rüegg, W., e. (1992). *Universities in the Middle Ages*. A History of the University in Europe, Vol. 1. Cambridge, UK: Cambridge University Press,.
- DelaCroix, D., Docquier, F., Fabre, A., and Stelter, R. (2020). The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000-1800). CEPR Discussion Papers 14509, C.E.P.R. Discussion Papers.
- Di Martino, P. and Vasta, M. (2017). *Ricchi per caso. La parabola dello sviluppo economico italiano*. Di Martino, M. Vasta (eds.). Mulino.
- Driscoll, J. and Kraay, A. (1998a). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4):549–560.

- Driscoll, J. and Kraay, A. (1998b). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4):549–560.
- Eurydice (1999). Two decades of reform in higher education in Europe: 1980 onwards. Eurydice studies, Eurydice.
- Eurydice (2000). *Two Decades of Reform in Higher Education in Europe: 1980 Onwards*. Eurydice studies. Eurydice European Unit.
- 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.
- Federico, G., Nuvolari, A., and 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.
- Fioravanti, G., Moretti, M., and Porciani, I. (2000). *L’istruzione universitaria, 1859-1915*. Ministero per i Beni e le Attività culturali. Ufficio centrale per i Beni archivistici, Roma.
- Fois, G. (1991). *L’Università di Sassari nell’Italia liberale*. Università di Sassari.
- Galor, O. (2005). Unified growth theory. Development and comp systems, EconWPA.
- Gennaioli, N., Porta, R. L., de Silanes, F. L., and Shleifer, A. (2013). Human Capital and Regional Development. *The Quarterly Journal of Economics*, 128(1):105–164.
- Glaeser, E. (2010). *Agglomeration Economics*. National Bureau of Economic Research Conference Report. University of Chicago Press.
- Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, (forthcoming).
- Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598.

- 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.
- Madrick, J. (2002). *Why economies grow*. Basic Books.
- Malanima, P. (2003). Measuring the italian economy. 1300-1861. *Rivista di storia economica*, (3):265–296.
- Malanima, P. (2006). An Age of Decline. Product and Income in Eighteenth-Nineteenth Century Italy. *Rivista di storia economica*, (1):91–134.
- Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–542.
- Mokyr, J. (2005). Long-Term Economic Growth and the History of Technology. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1 of *Handbook of Economic Growth*, chapter 17, pages 1113–1180. Elsevier.
- Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121(1-2):175–212.
- Moretti, E. (2011). *Local Labor Markets*, volume 4 of *Handbook of Labor Economics*, chapter 14, pages 1237–1313. Elsevier.
- OECD (2007). *Higher Education and Regions: Globally Competitive, Locally Engaged. Final Report, OECD, Paris*. OECD Paris.
- OECD (2008). *OECD Review of Tertiary Education. Final Report, OECD, Paris*. OECD Paris.
- OECD (2014). *The State of Higher Education 2014. A. Glass, (eds.). The OECD Higher Education Programme (IMHE), Paris*. OECD Paris.
- Ottaviano, G. (2011). 'new' new economic geography: firm heterogeneity and agglomeration economies. *Journal of Economic Geography*, 11(2):231–240.
- Parchet, R. (2018). Are local tax rates strategic complements or strategic substitutes? *American Economic Journal:Economic Policy (forthcoming)*.

- 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)*. Pubblicazioni dell'Istituto per la scienza dell'amministrazione pubblica. N. Pozza.
- Roberts, John; Rodriguez Cruz, A. M. H. J. (1996). "Exporting models". pages 256–284. Cambridge University Press.
- Schutz, G., Ursprung, H., and Woessmann, L. (2005). Education policy and equality of opportunity. CESifo Working Paper Series 1518, CESifo Group Munich.
- 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*. Public Universities and the Public Sphere. Palgrave Macmillan, New York.
- Squicciarini, M. P. and 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.
- Woessmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2):117–170.

Appendix A Additional Tables

Table A-1: Numbers of faculties in Italy by field of study: 1870 and 2010

Faculty	1870	2010
Humanities	16	127
Education	1	34
Languages	1	24
Literature	12	54
Psychology	2	15
Scientific and Medical studies	51	186
Agriculture	9	37
Chemistry&Pharmacy	18	31
Geology&Biology	1	3
Scientific studies	15	48
Architecture	3	23
Engineering	5	44
Social Sciences	42	210
Medical studies	19	39
Economics&Statistics	1	68
Law	21	55
Sociopolitical studies	1	48

Notes: There are five faculty fields that first appeared in Italian universities after 1870 and include Education (1876), Foreign Languages (1954), Geology and Biology (1993) and Psychology (1971).

Table A-2: Omitted neighbourhood variables and IV: full set of coefficients

	[1] baseline sample		[3] preunitarian provinces		[4] preunitarian prov. (lagged 10 y.)	
	OLS FE	IV FE	OLS FE	IV FE	OLS FE	IV FE
Panel A - Average neighbour approach						
total no. faculties in $-i$	-0.48*** (0.14)	-1.14*** (0.39)	-0.09 (0.08)	-1.08** (0.49)	-0.11 (0.10)	-0.94* (0.50)
no. of universities	1.48*** (0.55)	1.36** (0.57)	1.82*** (0.64)	2.00*** (0.63)	1.62** (0.61)	2.11*** (0.55)
no. of A-level universities	1.06*** (0.34)	0.86** (0.34)	0.87** (0.34)	0.31 (0.41)	0.95*** (0.31)	0.59 (0.39)
no. of private universities	2.38*** (0.57)	2.20*** (0.56)	1.93*** (0.53)	1.41** (0.57)	2.28*** (0.71)	1.71** (0.65)
Panel B - Pairwise approach						
total no. faculties in j	-0.08*** (0.03)	-0.16*** (0.05)	-0.04** (0.02)	-0.13** (0.06)	-0.05* (0.02)	-0.12*** (0.04)
no. of universities	2.17*** (0.52)	2.20*** (0.49)	2.34*** (0.56)	2.48*** (0.49)	2.17*** (0.54)	2.37*** (0.45)
no. of A-level universities	0.65** (0.27)	0.64*** (0.24)	0.48* (0.26)	0.46** (0.21)	0.56** (0.26)	0.57** (0.22)
no. of private universities	1.95*** (0.54)	1.79*** (0.50)	1.47*** (0.45)	1.29*** (0.38)	1.82*** (0.59)	1.77*** (0.54)

Notes:. The full set of coefficients refers to OLS FE and IV FE estimates reported in Table 6. All specifications include province fixed effects (Panel A), provincial pair fixed effects (Panel B), and region-by-year fixed effects. In IV estimates, the instruments are interactions of initial conditions with higher education reforms, as reported in Table 5. Standard errors are clustered at the province level. Significance levels: * : 10% ** : 5% *** : 1%.

Table A-3: Neighbourhood effects across and within fields of study: first stage

	Humanities (HH)		Stem (ST)		Social Sciences (SS)	
$(IC_j \text{ in HH})^*(L. 2102/1923)$	0.34*** (0.11)	-0.18*** (0.06)		0.09 (0.08)		1.05*** (0.08)
$(IC_j \text{ in HH})^*(L. 1592/1933)$	0.08 (0.06)	-0.23*** (0.06)		0.26*** (0.07)		0.41*** (0.07)
$(IC_j \text{ in HH})^*(L. 910/1969)$	0.16** (0.07)	-0.22** (0.10)		0.49*** (0.04)		0.06 (0.11)
$(IC_j \text{ in HH})^*(L. 766/1973)$	0.04 (0.03)	-0.12** (0.05)		0.12*** (0.04)		0.30*** (0.10)
$(IC_j \text{ in HH})^*(L. 382/1980)$	0.01 (0.02)	0.03 (0.02)		-0.05 (0.04)		0.01 (0.05)
$(IC_j \text{ in HH})^*(L. 168/1989)$	-0.05** (0.02)	0.06** (0.03)		-0.34*** (0.06)		0.11** (0.04)
$(IC_j \text{ in HH})^*(L. 245-341/1990)$	0.41*** (0.11)	0.40*** (0.09)		-0.53*** (0.08)		0.14** (0.07)
$(IC_j \text{ in HH})^*(L. 59/1997)$	0.58*** (0.14)	0.36*** (0.11)		-0.26*** (0.08)		0.13 (0.13)
$(IC_j \text{ in ST})^*(L. 2102/1923)$		0.14*** (0.02)		0.04 (0.04)	0.15*** (0.04)	-0.23*** (0.04)
$(IC_j \text{ in ST})^*(L. 1592/1933)$		-0.03 (0.02)		-0.10** (0.05)	0.15*** (0.02)	0.09** (0.04)
$(IC_j \text{ in ST})^*(L. 910/1969)$		-0.05 (0.04)		-0.18*** (0.03)	0.22*** (0.02)	0.15*** (0.04)
$(IC_j \text{ in ST})^*(L. 766/1973)$		-0.12*** (0.03)		-0.04* (0.02)	0.12** (0.05)	0.23*** (0.07)
$(IC_j \text{ in ST})^*(L. 382/1980)$		-0.01 (0.01)		-0.04 (0.02)	0.00 (0.02)	0.04** (0.02)
$(IC_j \text{ in ST})^*(L. 168/1989)$		-0.02 (0.01)		0.05** (0.02)	0.02 (0.02)	0.00 (0.02)
$(IC_j \text{ in ST})^*(L. 245-341/1990)$		0.11*** (0.03)		-0.02 (0.04)	0.05 (0.03)	-0.11*** (0.03)
$(IC_j \text{ in ST})^*(L. 59/1997)$		0.17*** (0.04)		-0.03 (0.03)	0.07 (0.05)	-0.12** (0.06)
$(IC_j \text{ in SS})^*(L. 2102/1923)$		-0.17*** (0.03)	0.10*** (0.03)	0.02 (0.03)		0.11*** (0.04)
$(IC_j \text{ in SS})^*(L. 1592/1933)$		0.09*** (0.03)	0.11*** (0.03)	0.10** (0.04)		-0.09** (0.04)
$(IC_j \text{ in SS})^*(L. 910/1969)$		0.18*** (0.06)	0.09*** (0.02)	0.03 (0.04)		-0.00 (0.06)
$(IC_j \text{ in SS})^*(L. 766/1973)$		0.17*** (0.03)	0.02 (0.01)	0.01 (0.03)		-0.29*** (0.07)
$(IC_j \text{ in SS})^*(L. 382/1980)$		-0.00 (0.01)	0.03 (0.02)	0.08*** (0.02)		-0.06** (0.03)
$(IC_j \text{ in SS})^*(L. 168/1989)$		-0.06*** (0.02)	0.09*** (0.03)	0.14*** (0.04)		-0.01 (0.02)
$(IC_j \text{ in SS})^*(L. 245-341/1990)$		-0.22*** (0.05)	0.16*** (0.03)	0.29*** (0.06)		0.09** (0.04)
$(IC_j \text{ in SS})^*(L. 59/1997)$		-0.15*** (0.05)	-0.00 (0.02)	0.04 (0.04)		0.09 (0.07)
R sq.	0.82	0.89	0.96	0.97	0.87	0.93
N	32644	32644	32644	32644	32644	32644

Notes:. The first stage IV FE estimates are reported in Table 10. All specifications include provincial pair fixed effects, region-by-year fixed effects, and the usual set of provincial controls. Standard errors clustered by province are reported in parentheses. Significance levels: * : 10% ** : 5% ***: 1%.

Table A-4: Neighbourhood effects: controls for local demand

	[1] OLS FE	[2] IV FE
Panel A - Baseline specification		32644
total no. faculties in j	-0.05* (0.02)	-0.13*** (0.04)
no. of universities	2.22*** (0.54)	2.42*** (0.45)
no. of private universities	1.78*** (0.59)	1.72*** (0.54)
no. of elite universities	0.55** (0.27)	0.56** (0.23)
Panel B - Drop metropolitan cities		21412
tot. no. faculties in j	-0.11** (0.04)	-0.22*** (0.08)
no. of universities	1.92*** (0.50)	2.15*** (0.39)
no. of private universities	0.27 (0.52)	0.14 (0.43)
no. of elite universities	0.77*** (0.28)	0.69*** (0.23)
Panel C - Control for population size		32644
total no. faculties in j	-0.04 (0.02)	-0.12*** (0.04)
no. of universities	2.09*** (0.59)	2.30*** (0.50)
no. of private universities	1.63*** (0.57)	1.70*** (0.52)
no. of elite universities	0.47* (0.26)	0.54** (0.22)
total population (log)	1.34* (0.76)	0.69 (0.68)
Panel D - Weight population size		32644
total no. faculties in j	-0.08*** (0.02)	-0.18*** (0.05)
no. of universities	1.88*** (0.56)	2.29*** (0.40)
no. of private universities	2.28*** (0.45)	1.92*** (0.35)
no. of elite universities	0.61*** (0.19)	0.61*** (0.17)
Panel E - Control for share of 0-14 year olds and tertiary educated		11442
total no. faculties in j	-0.06** (0.03)	-0.12* (0.06)
no. of universities	1.98*** (0.58)	2.04*** (0.55)
no. of private universities	1.49*** (0.53)	1.39*** (0.50)
no. of elite universities	0.00 (.)	0.00 (.)
0-14 cohort share	-0.02 (0.07)	-0.02 (0.07)
share with tertiary education	1.79** (0.81)	1.94** (0.74)
Panel F - Control for share active in industry sector		32644
total no. faculties in j	-0.05* (0.02)	-0.12*** (0.05)
no. of universities	2.24*** (0.54)	2.47*** (0.45)
no. of private universities	1.76*** (0.55)	1.68*** (0.50)
no. of elite universities	0.54** (0.25)	0.55** (0.21)
share active in the industry sector	-0.01 (0.01)	-0.01 (0.01)
participation rate	0.01 (0.02)	0.01 (0.02)
Panel G - Placebo: "alphabetical" neighbours		45269
total no. faculties in j	0.008 (0.019)	0.066 (0.063)
no. of universities	1.810*** (0.589)	1.828*** (0.557)
no. of elite universities	0.953*** (0.265)	1.034*** (0.245)
no. of private universities	2.360*** (0.613)	2.187*** (0.595)

Notes: The full set of coefficients refers to OLS FE and IV FE estimates reported in Table 7. All specifications include provincial pair fixed effects and region-by-year fixed effects. In IV estimates, the instruments are interactions of initial conditions with higher education reforms. Standard errors are clustered at the provincial level. Significance levels: * : 10% ** : 5% *** : 1%.

Table A-5: Economic returns of higher education supply (1st stage).

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
A) Dependent variable F_{it}							
$IC_i * R_{21/30}$	0.3177*** (0.0806)	0.4080*** (0.0868)	0.4080*** (0.0868)	0.4387*** (0.0991)	0.4387*** (0.0991)	0.0016 (0.0355)	
$IC_i * R_{31/36}$	0.1206*** (0.0279)						
$IC_i * R_{61/70}$	0.2002*** (0.0531)	0.3286*** (0.0765)	0.3286*** (0.0765)	0.3464*** (0.0833)	0.3464*** (0.0833)	0.0101 (0.0396)	
$IC_i * R_{71/80}$	0.0500 (0.0366)						
$IC_i * R_{81/90}$	0.0304 (0.0477)						
$IC_i * R_{91/00}$	0.1878 (0.1240)						
$IC_{-i} * R_{21/30}$				0.0546 (0.0545)	0.0546 (0.0545)	-0.0096 (0.0136)	
$IC_{-i} * R_{61/70}$				0.0280 (0.0623)	0.0280 (0.0623)	-0.0090 (0.0278)	
F_{it-20}						1.0341*** (0.0776)	
F_{-it-20}						-0.0008 (0.0158)	
B) Dependent variable F_{-it}							
$IC_i * R_{21/30}$				-0.1404 (0.1919)	-0.1404 (0.1919)	-0.0014 (0.0798)	
$IC_i * R_{61/70}$				-0.0559 (0.1557)	-0.0559 (0.1557)	0.1098 (0.1215)	
$IC_{-i} * R_{21/30}$				0.2376* (0.1363)	0.2376* (0.1363)	0.0364 (0.0379)	
$IC_{-i} * R_{61/70}$				0.4672*** (0.1179)	0.4672*** (0.1179)	0.1600* (0.0821)	
F_{it-20}						-0.0949 (0.0883)	
F_{-it-20}						0.9566*** (0.0783)	
Observations	937	937	937	937	937	874	
– within 90 Km	–	–	–	No	No	No	

Notes: First stage IV FE estimates are reported in Table 13. All specifications include province fixed effects and region-by-year fixed effects, population growth, the participation rate and the size of the industry sector. Standard errors clustered by province are reported in parentheses. Significance levels: * : 10% ** : 5% *** : 1%.

Table A-6: Economic returns of higher education supply: robustness checks

	[1] IV FE DKRAAY	[2] IV FE	[3] IV FE POPW	[4] IV FE	[5] IV FE DKRAAY	[6] IV FE	[7] IV FE POPW	[8] IV FE
no. of Faculties in i	0.0137*** (0.0024)	0.0116*** (0.0036)	0.0158*** (0.0053)	0.0134** (0.0051)	0.0113*** (0.0010)	0.0117*** (0.0017)	0.0126*** (0.0013)	0.0101*** (0.0027)
no. of Faculties in $-i$	0.0065*** (0.0011)	0.0061** (0.0026)	0.0073* (0.0039)	0.0064* (0.0038)	0.0023*** (0.0003)	0.0037** (0.0014)	0.0018 (0.0011)	0.0022** (0.0011)
Reforms in 1st stage	two	two	two	two	two	two	two	two
F_{it-20}, F_{-it-20} in 1st stage	no	no	no	no	yes	yes	yes	yes
$-i$ within 90 Km	no	yes	no	no	no	yes	no	no
Observations	937	937	937	892	874	874	874	832
K-P rk Wald F-stat	10.926	12.684	3.623	2.741	78.010	14.676	66.255	44.433
p-value Hansen	0.190	0.406	0.132	0.705	0.029	0.508	0.200	0.630

Notes: In all specifications, the dependent variable is $\ln(1 + Y_{it})$. The total number of faculties in the province and neighbourhood is lagged 10 years. All specifications include growth in the local population and the size of the industry and agricultural sectors as controls. They also include the province FE and region by year dummies. In all specifications, the two reform packages used to construct the instruments are $R_{23/30}$ and $R_{61/70}$ (all first stage estimates are available from the authors upon request). In Columns [3] and [7], we use the provincial population in 1861 as a weight (POPW). In Columns [4] and [8], we drop metropolitan cities belonging to the top quartile of the initial population distribution (i.e., Milano, Napoli, and Torino. See Table 1). Standard errors are always reported in parentheses. In Columns [1] and [5], they are robust to potential spatial dependence (DKRAAY, Driscoll and Kraay (1998a)), while in all other specifications, they are clustered at the province level. Significance levels: * : 10% ** : 5% *** : 1%.

Table A-7: Welfare analysis of higher education supply.

Province	F_{-i}	F_i	$\hat{F}_i$	$\tilde{Y}_i/Y_i$	$\tilde{\tilde{Y}}_i/Y_i$	$(\tilde{Y}_i - \tilde{\tilde{Y}}_i)/Y_i$
Agrigento	11.00	0.00	1.39	1.92	1.10	0.81
Alessandria	36.00	3.00	7.55	5.17	3.52	1.65
Ancona	29.00	5.00	8.66	5.17	3.89	1.28
Aosta	15.00	5.00	6.90	3.51	3.01	0.50
Ascoli Piceno	24.00	1.00	4.03	3.42	2.18	1.24
Aquila	51.00	7.00	13.44	11.85	8.54	3.31
Arezzo	42.00	1.00	6.31	4.80	2.91	1.89
Asti	17.00	0.00	2.15	1.63	1.02	0.60
Avellino	42.00	0.00	5.31	6.91	3.90	3.00
Bari	5.00	16.00	16.63	13.31	12.91	0.39
Bergamo	39.00	6.00	10.93	6.00	4.42	1.58
Biella	15.00	0.00	1.90	1.39	0.88	0.51
Belluno	18.00	0.00	2.27	1.65	0.94	0.70
Benevento	37.00	4.00	8.68	9.65	6.79	2.86
Bologna	26.00	15.00	18.29	8.49	7.39	1.09
Brindisi	21.00	2.00	4.65	5.03	3.76	1.28
Brescia	16.00	6.00	8.02	4.05	3.39	0.66
Bolzano	6.00	3.00	3.76	1.55	1.38	0.16
Cagliari	0.00	11.00	11.00	6.51	6.51	0.00
Campobasso	20.00	6.00	8.53	6.52	5.30	1.22
Caserta	32.00	10.00	14.04	14.84	12.55	2.29
Chieti	13.00	7.00	8.64	6.16	5.37	0.79
Caltanissetta	22.00	0.00	2.78	3.73	2.13	1.60
Cuneo	14.00	0.00	1.77	1.25	0.71	0.53
Como	6.00	2.00	2.76	1.51	1.25	0.26
Cremona	53.00	1.00	7.70	5.88	3.49	2.38
Cosenza	7.00	6.00	6.88	5.56	5.06	0.50
Catania	12.00	11.00	12.52	10.81	9.87	0.94
Catanzaro	9.00	3.00	4.14	3.26	2.70	0.56
Enna	33.00	5.00	9.17	9.91	7.74	2.17
Forli-Cesena	12.00	5.00	6.52	3.11	2.64	0.47
Ferrara	22.00	8.00	10.78	6.42	5.34	1.08
Foggia	10.00	6.00	7.26	6.70	5.92	0.79
Firenze	38.00	12.00	16.80	8.97	7.26	1.70
Frosinone	51.00	4.00	10.44	7.98	5.76	2.22
Genova	15.00	11.00	12.90	6.75	6.05	0.70

Province	F_{-i}	F_i	$\widehat{F}_i$	$\widetilde{Y}_i/Y_i$	$\widetilde{\widetilde{Y}}_i/Y_i$	$(\widetilde{Y}_i - \widetilde{\widetilde{Y}}_i)/Y_i$
Gorizia	22.00	0.00	2.78	2.15	1.35	0.80
Grosseto	24.00	0.00	3.03	2.50	1.43	1.07
Imperia	0.00	0.00	0.00	0.00	0.00	0.00
Isernia	24.00	1.00	4.03	3.87	2.67	1.20
Crotone	9.00	0.00	1.14	1.61	1.02	0.59
Lecco	5.00	0.00	0.63	0.44	0.28	0.16
Lecce	1.00	8.00	8.13	6.33	6.26	0.07
Livorno	11.00	0.00	1.39	1.13	0.65	0.48
Lodi	43.00	0.00	5.43	4.25	2.65	1.60
Latina	47.00	0.00	5.94	5.43	3.37	2.05
Lucca	31.00	0.00	3.92	2.93	1.67	1.26
Macerata	17.00	10.00	12.15	7.29	6.40	0.89
Messina	22.00	11.00	13.78	12.20	10.51	1.69
Milano	21.00	33.00	35.65	15.64	14.66	0.98
Mantova	40.00	0.00	5.05	3.53	2.00	1.53
Modena	25.00	7.00	10.16	5.05	4.06	0.99
Massa Carrara	12.00	0.00	1.52	1.38	0.80	0.59
Matera	23.00	2.00	4.91	4.54	3.35	1.19
Napoli	20.00	23.00	25.53	24.95	23.01	1.94
Novara	47.00	3.00	8.94	6.22	4.08	2.14
Nuoro	20.00	0.00	2.53	2.61	1.64	0.97
Oristano	20.00	0.00	2.53	2.89	1.82	1.07
Otranto	0.00	0.00	0.00	0.00	0.00	0.00
Palermo	11.00	11.00	12.39	10.21	9.38	0.83
Piacenza	33.00	3.00	7.17	4.37	3.01	1.36
Padova	11.00	13.00	14.39	7.03	6.52	0.50
Pescara	14.00	3.00	4.77	3.54	2.90	0.64
Perugia	33.00	11.00	15.17	10.24	8.37	1.86
Pisa	19.00	11.00	13.40	7.21	6.31	0.90
Pordenone	15.00	0.00	1.90	1.36	0.86	0.50
Prato	26.00	0.00	3.29	2.35	1.48	0.88
Parma	16.00	12.00	14.02	6.67	5.98	0.69
Pistoia	33.00	0.00	4.17	3.32	2.08	1.24
Pesaro-Urbino	17.00	9.00	11.15	6.28	5.45	0.82
Pavia	41.00	8.00	13.18	8.57	6.51	2.06

Province	F_{-i}	F_i	$\hat{F}_i$	$\tilde{Y}_i/Y_i$	$\tilde{\tilde{Y}}_i/Y_i$	$(\tilde{Y}_i - \tilde{\tilde{Y}}_i)/Y_i$
Potenza	31.00	6.00	9.92	8.74	6.67	2.07
Ravenna	37.00	0.00	4.68	3.61	2.05	1.56
Reggio Calabria	3.00	4.00	4.38	3.60	3.38	0.22
Reggio Emilia	17.00	4.00	6.15	3.38	2.67	0.71
Ragusa	12.00	0.00	1.52	1.78	1.13	0.66
Rieti	63.00	0.00	7.96	7.52	4.64	2.88
Roma	14.00	38.00	39.77	19.19	18.43	0.75
Rimini	13.00	1.00	2.64	1.49	1.09	0.40
Rovigo	32.00	0.00	4.04	3.12	1.77	1.35
Salerno	26.00	10.00	13.29	11.57	9.66	1.91
Siena	39.00	8.00	12.93	7.75	5.94	1.82
Sondrio	19.00	0.00	2.40	1.63	0.93	0.70
La Spezia	21.00	0.00	2.65	2.07	1.30	0.77
Siracusa	11.00	1.00	2.39	2.41	1.68	0.73
Sassari	0.00	11.00	11.00	7.63	7.63	0.00
Savona	14.00	0.00	1.77	1.31	0.83	0.49
Taranto	20.00	4.00	6.53	6.16	5.00	1.17
Teramo	10.00	5.00	6.26	4.48	3.89	0.60
Trento	14.00	7.00	8.77	4.19	3.73	0.46
Torino	4.00	14.00	14.51	7.19	7.00	0.19
Trapani	11.00	0.00	1.39	1.88	1.08	0.80
Trani	24.00	0.00	3.03	2.67	1.68	0.99
Trieste	0.00	12.00	12.00	4.99	4.99	0.00
Treviso	18.00	0.00	2.27	1.74	0.99	0.74
Udine	5.00	10.00	10.63	5.06	4.83	0.22
Varese	37.00	6.00	10.68	6.14	4.83	1.32
Verbania	10.00	0.00	1.26	1.11	0.70	0.41
Vercelli	28.00	1.00	4.54	2.97	2.02	0.94
Venezia	23.00	5.00	7.91	4.33	3.37	0.96
Vicenza	25.00	0.00	3.16	2.26	1.29	0.97
Verona	25.00	7.00	10.16	5.44	4.38	1.07
Viterbo	44.00	6.00	11.56	8.61	6.61	2.00
Vibo Valentia	6.00	0.00	0.76	0.98	0.62	0.36
Total	20.25	4.75	7.31	5.06	4.06	1.00

Notes: All values refer to 2010 levels. Information on Y_i is not available for provinces of Barletta, Carbonia-Iglesias, Fermo, Monza-Brianza, Medio-Campidano and Ogliastro.