Higher Education Supply, Local Competition and Economic Performance: Evidence from Italy*

Elena Cottini (Universita' Cattolica di Milano),
Paolo Ghinetti (Universita' del Piemonte Orientale)
Simone Moriconi (Iéseg School of Management)

This draft: January 2018

Abstract

This paper exploits an own built dataset on the history of higher education institutions in Italy during 1861-2010, to analyse local competition effects in the supply of higher education. We measure local supply of higher education as the number of teaching units (faculties) at the province level, and analyse how local supply responds to the supply of neighbouring provinces. Our identification strategy relies on instrumental variables that exploit initial conditions of the Italian higher education supply, inherited by the states that existed before the Italian unification. We find that local competition effects are sizeable: the more conservative estimates show that having 8 more faculties in the neighbourhood reduces the local supply by about 1 faculty. Such competition forces are mostly concentrated within the same field of study, the same region, and a spatial reach of 90 Km. Finally, we estimate the economic returns of higher education supply, and show that the opportunity cost of competition forces may be sizeable, in terms of foregone per capita value added. However, this cost is compensated by positive economic externalities coming from the higher education supply in the neighborhood.

Key Words: local competition; higher education supply; historical data; initial conditions.
JEL codes: I23, I28, N00, R1

*Address: Elena Cottini, Department of Economics and Finance, Universita' Cattolica di Milano, Largo Gemelli 1, 20123 Milano (MI), Italy; elena.cottini@unicatt.it; Paolo Ghinetti, Dipartimento di Studi per l'Economia e l'Impresa, Universita' del Piemonte Orientale, Via Perrone 18, Novara (NO), Italy: paolo.ghinetti@uniupo.it; Simone Moriconi, Department of Economics and Quantitative Methods, Iéseg School of Management and LEM, 1 parvis de La Dèfense - 92044 Paris - La Dèfense Cedex (France); s.moriconi@ieseg.fr.

1 Introduction

Throughout the past century, OECD countries experienced a major expansion of their higher education (HE) systems, both in terms of students enrolled and number of institutions involved in the market. This process, particularly during the last few decades was the result of a local diffusion and differentiation of HE supply (e.g. in terms of autonomy, territorial attachment, etc. See Eurydice, 2008), aiming at featuring HE systems which could match the increasing development needs at the regional and sub-regional (e.g. provincial) level (OECD, 1999). As a result of this process, HE providers nowadays operate in more competitive environments, which explains different performances of HE institutions (Aghion et al., 2010), as well as cross-regional differentials in regional productivity and income per capita (Gennaioli et al., 2013; Valero and Van Reenen, 2016).

The present paper contributes to this debate by investigating competition in the provision of HE services at the local level. HE providers are nowadays regarded as a major source of local development. They may attract private investments, and enhance local human capital. HE institutions, particularly in highly centralised states, provide key local agencies that bring together within the territory national interests in science and technology, industrial performance, education and skills, health, social inclusion and culture (OECD, 2008). For these reasons, the modern approach to the provision of HE is a bottom-up one, which moves from the needs of regional or even provincial development: the establishment of a HE provider in a province, may greatly benefit the local stakeholders, by diverting resources there, perhaps at the expense of neighbouring provinces which see reduced opportunities to open their own HE institution.

The primary goal of this paper is to analyse the existence of local competition in the supply of HE, by estimating the response of HE supply in an Italian province, to the supply of neighbours. Our analysis accounts for the fact that local HE supply may have feedback effects on HE supply of neighbours, or they can be both determined by external factors. To address such reverse causality and endogeneity issues, we exploit exogenous variation coming from the geographical distribution of HE supply inherited by pre-unitarian Italian states (i.e. the initial conditions of the Italian HE supply), interacted with state-level reforms of the Italian university system. Italy is indeed the country with the longest history of university education in the Western world. The first *Studium* was opened in Bologna, and dates back to 1088; many other institutions were established by the XII century.[1] We exploit the

[1] In the Middle-Age a *Studium* was an autonomous and secular organisation among students that chose and personally funded teachers through a donation collection system (*collectio*). In most cases, a *Studium* anticipated the foundation of a University (Pini, 2000). Besides Bologna, by the beginning of the XII

expansion of pre-unitarian HE supply triggered by national HE reforms during the Italian history to identify competition forces. We analyse whether competition forces operate within or across broad fields of studies and discuss the boundaries of competition effects. In the final part of the paper, we build upon the recent literature that looks at the economic effect of HE, to evaluate the economic returns of HE supply in terms of value added per capita. This allows us to measure the opportunity cost (in terms of foregone value added per capita) of HE competition, and compare this cost with the direct economic spillovers coming from HE supply of neighbours.

To these purposes, we constructed the History of Italian Universities (HIU) dataset, a cross-province panel which follows over time 110 Italian provinces between 1861, which is the year of Italian unification, and 2010, when a major reform radically changed the Italian university system. For each province, we record the local supply of HE institutions, measured as the number of faculties - the university teaching unit responsible for all BSc or MSc courses within a given field of study - registered within each province in a given year (in total, and by area of science). We then match the cross-province panel dataset to its mirror images and register pairwise relationship between province couples based either on a contiguity, or a distance criterion, or both to estimate competition effects, following the main approaches by the economics literature (see e.g. Davies and Vadlamannati, 2013; Parchet, 2014).

A desirable feature of our data is the long time span (150 years), which goes back at the onset of Italian unification, which provides substantial variability in the supply of HE at the province level. We control for local unobserved heterogeneity by including province and region-by-year fixed effects. Our data also include province level indicators (i.e. total and active population, size of the population cohort aged 0-14, relative size of agricultural/industry and services sectors), which provide useful control variables throughout the analysis. We can identify the competition effects by an instrumental variables procedure that exploits exogenous variation associated with the initial conditions of Italian higher education interacted with HE reforms that occurred during Italian history. This empirical strategy relies on two identifying assumptions. The first is that the cross-province distribution of the supply of higher education at the onset of Italian unification is exogenous. This "pre-unitarian" HE supply was indeed inherited by the states that existed before the Italian unification i.e. determined by exogenous factors such as culture, politics, institutions, and geography (Squicciarini and Voigtländer, 2015). The second identifying assumption is that single Italian provinces due not have enough "voice" (due to their weak political relevance) to influence

century, on the Italian territory there were active *Studia* in Modena (1175), Napoli (1224), Perugia (1308), Siena (1240), and Roma (1303). In Parma the first *Studium* established in 962, was closed in the XII century and then re-opened in 1601 (See Brizzi and Romano, 2007 for details).

state-level decisions. By taking the interactions of state-level reforms with pre-unitarian provinces, we assume that each pre-unitarian HE system is affected by national reforms, but the intensity of treatment varies across systems.

We find evidence that HE reforms deeply shaped supply patterns in provinces where HE institutions would already be present at the onset of Italian history. The reform that liberalised access to university education in 1969 is particularly relevant, as it created strong incentives to expand local HE supply. Our main empirical results point to strong competition effects at the province level: our more conservative estimates suggest that having 8 faculties more in the neighbourhood reduces local HE supply by over 1 faculty. This is a non-negligible impact, considering that the expansion of higher education that took place in the second half of the 20th century led to an increase by 4 faculties per province, on average. We show that local displacement forces mostly operate within the same field of study, the same region, and a 90 Km linear distance. Finally, our estimates show that the economic returns of one more faculty in a province is $0.8\% - 1.4\%$ of local value added. This suggests that the opportunity cost of competition forces may be sizeable, in economic terms. However, our estimates also point out that this cost is over-compensated by direct positive externalities coming from HE supply in the neighbourhood.

The rest of the paper is organised as follows. In the next section we provide background information about the Italian HE system, as well as we frame our analysis in the literature, from both a theoretical and empirical point of view. In Section 3 we describe the data used in the analysis. The empirical analysis and main results of competition in HE supply are discussed in Section 4. Section 5 discusses the opportunity cost of competition effects in terms of economic value of HE supply. Section 6 concludes.

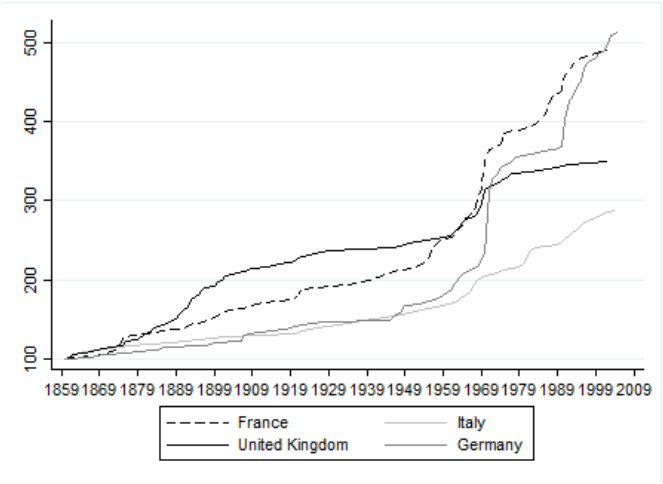
2 Background and literature

2.1 Institutional and historical setting

The Italian provision of higher education is the prototypical European one. Figure 1 documents the evolution of HE supply in France, Italy, Germany and UK. Starting from the mid-nineteenth century, all four countries are characterised by a common pattern i.e. a positive and moderate increasing trend in the supply of higher education, with a marked increase in the pace of expansion starting from the 1960s.[2] In all four countries, HE is regulated by

[2]Madsen and Murtin (2017) find also a very similar trend in education attainments for the UK. Their time series starts much earlier than ours i.e. in the XIIIth century, and refers to total years of schooling (i.e.

Figure 1: Evolution in the supply of higher education 1859-2009 in France (FR), Germany (DE), UK and Italy (IT): n. universities



Notes: authors calculations on HIU and WHED UNESCO data.

the central government, which put in place reforms that favoured the expansion of higher education (Eurydice, 2008). While it possesses a flatter profile relative to the other three countries, the growth of HE supply in Italy was substantial, and in 2009 the number of universities was 3 times larger than in 1859.[3]

In line with European standards, HE providers in Italy are highly differentiated between old and new ones, large and small, public and private ones. The faculty is the relevant HE institution. This is the teaching unit, and it is meant to govern the supply of HE in a given field of study. The genesis of a faculty is very often detached from that of a university and often tailored to respond to local demand of HE services. In the history of the Italian HE system there are many cases in which the creation of the faculty is antecedent that of the university. In pre-war period, this was due to institutional constraints as very few disciplines were formally taught within the university framework.[4] In the post-war period, HE supply in a province was often established through the creation of new faculties. Their genesis was often supported by province-level *consortia*, which gathered all competences needed to

includes primary-to-tertiary education).

[3]There are several reasons for the Italian trend, for example the persistently lower number of high school graduates compared to countries like Germany, especially in the last 40 years.

[4]At the onset of Italian history there were only medicine, law, humanities, mathematics and natural sciences. All the scientific knowledge and the social sciences were taught by schools (equivalent to single faculty institutions) recognised by the state and part of the HE system. They granted degrees in professional and technical subjects equivalent to university degrees on a legal standpoint, but were not formally considered as university education, until legislative reforms in 1923 and 1933. As an example, the Veterinary School of Milan was already there at the moment of Italian unification, while the University of Milan was established in 1923. The School was formally annexed to the university in 1924 and transformed in Faculty in 1934 (see on-line appendix).

the project: a HE provider (i.e. an existing university, not necessarily located in the same province, and willing to expand its market and differentiate its supply), financing institutions and economic stakeholders (generally banks, chambers of commerce or local investors), and local politicians (particularly at the province-level), enjoying direct connections with central government's representatives. Each *consortium* would present a faculty start-up project, which would then be evaluated by the central state.[5]

Traditionally, HE providers enjoy some degree of local autonomy, with strong attachment to the territory and its economic and political stakeholders (OECD, 2008). This has historical reasons, reflecting the fact that the Italian HE system at the time of the unification was the simple aggregation of the HE systems of the ancient pre-unitarian Italian states. The province has always been the relevant units for the provision of higher education services. On a political ground, this is the oldest governance level on the Italian territory, equivalent to French departments and to US counties, which was inherited from pre-unitarian Italian States. From a territorial perspective, these are the intermediate units between the region and the municipality, and the level where one can observe the local demand of HE and the local effects of human capital on industrialisation patterns and structure of production (Squicciarini and Voigtländer, 2015; Ciccone and Peri, 2011), or wages (Ciccone and Peri, 2006; Bratti and Leombruni, 2014).

Table 1 presents the HE supply of the Italian Kingdom inherited by pre-unitarian States. In 1870, by the end of the main wave of Italian unification (i.e. after Papal States were annexed to the Kingdom of Italy), there were more than 80 faculties already operating. This pre-unitarian supply was geographically dispersed in 21 out of the 69 provinces on the Italian territory at that time.[6] The capital(s) of each pre-unitarian states (e.g. Palermo and Napoli in the Kingdom of the Two Sicilies), had their own HE institutions. Capitals of old duchies, which used to be independent states during the Middle-Age, and annexed to a pre-unitarian state later on (e.g. the Duchies of Ferrara and Pesaro Urbino, annexed by the Papal States in

[5]In the majority of cases, the presentation of project proposal would follow specific calls and development plans by the central state itself. Following this procedure, for example, in 1993 seven new faculties were opened in the Provinces of Novara, Vercelli, and Alessandria, as separate branches of the University of Torino. These gained autonomy in 1998, by the creation of the brand-new University of Eastern Piedmont. This case is not isolated, and several universities have faculties in multiple provinces (see on-line appendix for details).

[6]At the moment of unification, the first Italian government reorganised provincial constituencies taking into account most of the maps that historically different states had built over time (Palombelli, 2012). This administrative reorganisation implied the transformation of several pre-unitarian provinces into districts that were annexed to existing provinces to obtain larger territorial and administrative units. Many districts became autonomous again later on during Italian history, so that the number of provinces in our sample increases from 69 in 1870 to 110 in 2010. In the empirical analysis we account for this process, and check the robustness of our results to alternative definitions of pre-unitarian provinces.

the XVI-XVII centuries) had their own HE institutions too. Each institution complied with the accreditation rules of its own state only. The diversity of pre-unitarian HE system would also derive from cultural fragmentation: in 1861 only the 2.5% of the Italian population would be able to speak Italian, while the rest of the population would only use their local regional language (De Mauro, 1970).

The heterogeneity of pre-unitarian HE systems emerged clearly from the assessment of HE institutions present on the Italian territory by the Law 3725/1859 of the Kingdom of Sardinia (called Casati, from the name of the Minister of Education).[7] This assessment introduced a clear ranking in the Italian university system, which is described in Column [5] of Table 1: only in 9 out of the 21 pre-unitarian provinces with an active institution, HE supply fully complied with the declared quality standard of the Kingdom of Sardinia (A-level supply in Column [5]). In several provinces, HE supply would not fully match the quality standards, but provided second-tier regional higher education (B-level in Column [5]). In two provinces (Sassari, and Macerata), local HE institutions fell short of the minimum requirements to be considered higher education (C-level in Column [5]). The Casati Law also defined the requirements for private universities, which were not financed by the state. Private provision was a typical feature of the HE systems in the old provinces of Ferrara, Urbino and Perugia, where universities were financed by the Duchies and the Vatican (Brizzi and Romano, 2007).

Italian governments before WWII put a lot of effort to harmonise pre-unitarian HE systems and create a homogeneous national HE system. Two main regulatory reforms were implemented during this period. The Law 2102/1923 (called Gentile, from the name of the Minister of Education), introduced new disciplines (e.g. Political Sciences), granted academic status to technical studies, and ensures adequate financial resources to the HE system.[8] The reform also restored the distinction between A-level and B-level HE institutions, which had been progressively excised over the years (see on-line appendix for details). The reform launched by the Law 1592/1933 (called De Vecchi-Bottai) centralised higher education, by nationalising private universities and recognising technical and applied schools as faculties with full academic status.

Post-WWII governments put in place reforms that produced a new institutional setting

[7]This law set out the rules for accrediting pre-existing institutions into the new Italian university system. It was initially applied to new territories of the Kingdom of Sardinia. The successive Matteucci Law in 1862 extended it to all territories that gradually entered the Italian Kingdom.

[8]The reform expands the perimeter of the higher education system to include specialised schools in technical discipline such as engineering and architecture, economics, management, commercial and social sciences. Before the reform, these Schools were HE institutions legally recognised by the State, but had different rules (access criteria, financing, validity of degrees, etc) and stood outside the University system.

Table 1: Pre-unitarian HE supply

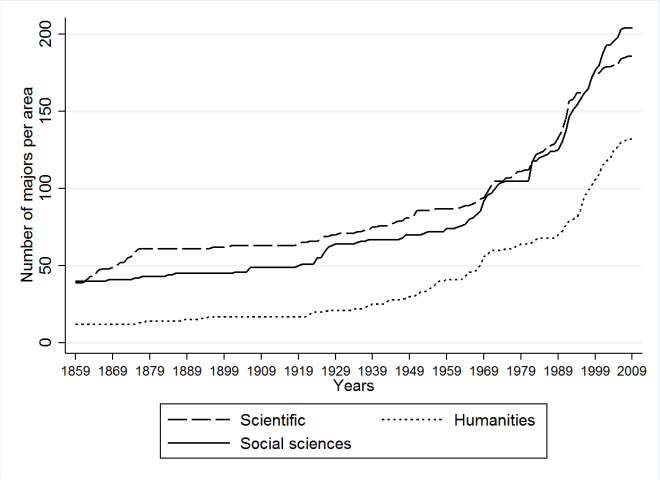
[1] Pre-unitarian Province	[2] Pre-unitarian State	[3] HE year of appearance	[4] No. of Faculties	[5] HE assessment Casati Law
Bologna	Papal States	1088	6	A
Cagliari	Kin. of Sardinia	1620	4	B
Catania	Kin. of two Sicilies	1445	4	B
Ferrara	Papal States	1391	4	private
Genova	Kin. of Sardinia	1481	5	B
Macerata	Papal States	1540	3	C
Messina	Kin. of two Sicilies	1838	3	B
Milano	Lombardy-Venetia	1791	2	A,A
Modena	Duchy of Modena	1175	4	B
Napoli	Kin. of two Sicilies	1224	7	A
Palermo	Kin. of two Sicilies	1806	4	A
Padova	Lombardy-Venetia	1407	4	A
Perugia	Papal States	1308	2	private
Pisa	Gran Duchy of Tuscany	1343	6	A
Parma	Duchy of Parma	962	5	B
P.Urbino	Papal States	1671	2	private
Pavia	Kin. of Sardinia	1361	5	A
Roma	Papal States	1303	5	A
Siena	Gran Duchy of Tuscany	1240	2	B
Sassari	Kin. of Sardinia	1765	3	C
Torino	Kin. of Sardinia	1404	8	A,B

Note: In Column [3], HE year of appearance refers to the year of the first studium in the province. In Milano HE supply appears with the autonomous School of Veterinary Studies. Column [5] reports the quality assessment of HE institution(s) in the province according to the evaluation framework set by the Casati Law. This evaluation refers to public institutions only. HE provision was private in Ferrara, Perugia, and P.Urbino. In Torino, the Casati Law assigns A-score to the University of Torino, and a B score to the Polytechnic. In Milano, it assigns A-score to both the Scientific Academy, and the Polytechnic (see on-line appendix for details). Source is History of Italian Universities (HIU) Data.

of the national university system. This was previously organised to serve the "elite", later on was asked to provide higher education for the "mass". The Law 910/1969 liberalised higher education access to all students with a 5 year diploma of secondary education, including those from technical schools (that before 1969 did not allow university enrolment). The consequent rise in enrolment rates put a lot of pressure on the national university system to adapt HE supply to the massive increase in demand. Indeed, the Law 766/1973 legislates the opening of new faculties and increases the number of faculty hires in public universities. The Law 382/1980 reorganised the internal governance of universities as well as the recruitment and career of university professors. Successive Laws 392/1989 and 245-341/1990 reorganise the system of allocation of funding to higher education.[9] The Law 59/1997 (called Bassanini

[9]The Law establishes a dedicated "Ministry of Universities and Scientific and Technological Research", and sets up triennial development plans for universities (see on-line appendix for details).

Figure 2: Evolution of HE supply in Italy in 1860-2010.



Notes: authors calculations on HIU data.

reform) granted universities more financial and teaching autonomy.

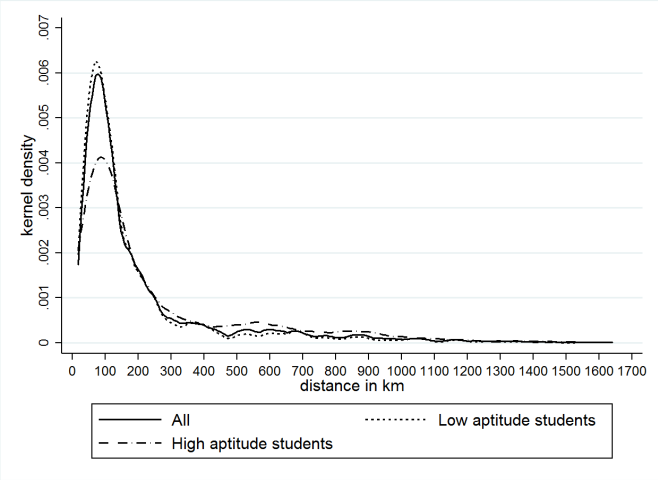
This transition from national elite to mass higher education took place during the same period everywhere in Europe (Eurydice, 1999). This is consistent with evidence presented in Figure 1 above for France, Germany, and the UK. A similar process also regarded the US (Smith, 2010). As detailed in Figure 2, the rise in HE supply after 1969, involved Humanities, STEM, and Social Sciences in a roughly similar way, notwithstanding a constantly lower HE supply in the former discipline relative to the latter two.

2.2 Literature review

The micro-foundations of our study are consistent with a monopolistically competitive location model featuring endogenous choices of students and HE institutions. Students consume higher education services and display preferences for a differentiated consumption (say a HE degree in a given field of study, or by a specific institution), and have commuting costs to the location of the HE institution. In this class of models, the spatial equilibrium depends on the trade-off between students' utility from consuming their preferred HE variety and minimising commuting costs. The spatial equilibrium features a dispersed HE supply, which is designed both to "retain" local students with high commuting costs, and attract students from close locations, by supplying their preferred HE variety.[10]

[10]At the equilibrium, the coordination costs of a spatially decentralised system more than compensate the advantages of having more varieties in the supply of HE at the local level. Accordingly, a decentralised equlibrium features too much spatial differentiation as compared to the social optimum in Hotelling or Salop location models, so that a centralised solution would be preferable from a welfare standpoint.

Figure 3: Students' mobility and choice of HE institution.



Notes: authors' calculations on data from the Italian survey of high school graduates, 2007 (ISTAT).

This dispersion is consistent with such a spatial equilibrium under the assumption that students have low mobility (i.e. high commuting cost). Figure 3 reports the distribution of high school graduates in Italy by distance from the university they enrolled to. It suggests that students' mobility is very low: the majority of Italian high school graduates chose a HE institution at a distance of about 100 Km. Mobility is higher among students with high aptitudes towards HE and it is particularly low among low aptitude students.[11] This recalls well known results for the US (see Hoxby, 2016).[12]

Our paper relates to the literature that studies the long-run changes of education institutions. Woesmann (2003) and Schtz et al. (2005) use student-level data from 39 countries to trace back international differences in students performance to key cross-country differences in institutional design (e.g. with respect to the degree of school autonomy, competition from private schools, and extent of equality in the education opportunities). Braga et al. (2013) exploit variation associated with reforms to education institutions in 24 countries over a time span of 70 years, to identify the causal effect of institutions on educational attainments. The long time span allows the authors to control for time invariant, country-specific factors. They show that education reforms that expand access to education increase average years of schooling and reduce educational inequality. Reforms that foster school accountability and

[11]We define students' aptitude based on their final grade at lower secondary schools. High aptitude students are those with the highest grade (4), while low aptitute students are those with the lowest grades (1 or 2), where grades are measured on a four point scale.

[12]Hoxby (2016) shows that the probability that a low aptitude students chooses a college at more than 250 miles of distance is less than 10% in 2010. The probability is equal to 50% for an average aptitude student and 90% for a high aptitude student.

autonomy, are also found to increase average years of schooling, but also increase inequality.[13] We contribute to this literature in several respects. The aforementioned studies relate institutional differences to educational attainments whereas we look at the input of the educational process. Also, our paper focuses on post-compulsory education institutions while existing studies look at either compulsory education only (Woesmann, 2003; Schtz et al., 2005), or both compulsory and post-compulsory education (Braga et al., 2013). Finally, the studies mentioned above use data from several countries, while we use historical data for one specific country, Italy. This allows us to avoid cross-country unobserved heterogeneity and exploit exogenous variation associated with the initial conditions of the Italian HE to investigate the effect of HE reforms occurring during Italian history on HE supply. The main objective of our analysis is to exploit this source of exogenous variation to make causal inference on the degree of local competition in the provision of HE services, which is a novel topic in the economics' literature.

The final part of the paper, where we estimate the economic returns of higher education supply, relates to the literature that looks at the long-term economic effects of education institutions. A first strand of this literature, inspired by Unified Growth Theories (Galor, 2005), adopts an historical perspective and discusses the role of initial conditions in the stock of human capital as a crucial element for industrialisation e.g. in France, Prussia, and UK (see respectively Squicciarini and Voigtländer, 2015; Becker et al., 2011; Madsen and Murtin, 2017).[14] In particular, our identification strategy is similar to Squicciarini and Voigtländer (2015), which exploit the exogenous cross-department distribution of subscriptions to the *Encyclopédie* in France to identify the impact of "upper-tail knowledge" on city growth after the onset of French industrialisation. The second strand analyses the economic returns of education institutions, also in the presence of spatial externalities (Gennaioli et al., 2013; Valero and Van Reenen, 2016; Ertur and Koch, 2007). Our empirical analysis is closer to Valero and Van Reenen (2016). They use a novel dataset, with panel information on the establishment, and regional location of nearly 15000 universities across 78 countries for the

[13]The literature is much more vast when it comes to specific policy and reforms, which however do not entail an institutional change. For example, Hoxby and Turner (2014) evaluate policy interventions that improve the access of low-income students to college related information (i.e. application process, college net costs). They show that this intervention makes low-income high performing students more likely to apply and being admitted to more colleges, especially those with higher graduation rates and offering students more instructional amenities.

[14]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. Similarly, Madsen and Murtin (2017) use 80 and 190 years lagged values of years of schooling to identify the impact of schooling on real GDP growth in the UK during 1270-2010. They show an important contribution of human capital accumulation (years of schooling) before and after the Industrial Revolution. This view is not undisputed. See for example, Clark (2005) for a discussion of the (lack of a) role of human capital accumulation for the industrial revolution in the UK.

post-war period. They estimate fixed effects models at the sub-national level and find that a doubling of universities in a region is associated with about 4 percent higher future GDP per capita. Our analysis differs as it concentrates on Italy, covering its 150 years history. It allows for observations at the NUTS3 level, measuring HE supply at the faculty level. We also complement the empirical strategy based on fixed effects with the instrumental variables' approach described above.

3 Data and descriptive statistics

3.1 The History of Italian Universities (HIU) Database

Our main data-source is an original dataset on the History of Italian Universities (HIU). This is an own compiled register that contains detailed information on institutions providing higher education in Italy, disaggregated at the faculty level. As mentioned above, this is the teaching unit, responsible for all BSc or MSc courses within a given disciplinary area. The dataset covers all years starting from 1861 (year of birth of the Kingdom of Italy) up to 2010, when the Law 30/12/2010, n. 240 (the so called "Gelmini reform" after the name of the Public Education Minister in office at that time), which eliminated the faculty from the governance of public universities.

The primary source we used to construct HIU is Brizzi and Romano (2007), which report detailed history of Italian universities starting from their foundation onwards. We integrated this information by 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). We also heavily relied on Gazzetta Ufficiale della Repubblica Italiana, a weekly publication that collects every public act taken by the government since 1861.[15] We double-checked all information against those provided by open-source archives i.e. Wikipedia, universities' and faculties' websites.[16] For the few faculties, for which we found little information on-line, we contacted the administrative representatives, so as to fill the missing information. Finally, we validated the data against the actual list of higher education institutions provided by the Italian Ministry for University Education and Research (MIUR).

The register includes the name of the faculty and its exact address; 15 faculty field identifiers, which we aggregated into 3 macro-areas of science (Social Sciences, STEM, and

[15] See Gazzetta Ufficiale della Repubblica Italiana at $http://www.gazzettaufficiale.it/30giorni/concorsi$.

[16] Since faculties no longer exist after 2010, their websites are not readily available on the web today. We retrieved them using Wayback Machine ($https://web.archive.org/$), a digital archive of the World Wide Web created by the Internet Archive.

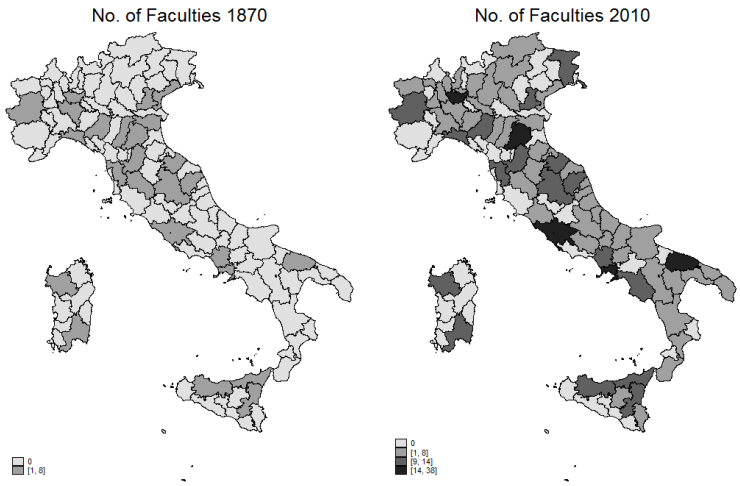
Humanities), according to the classification used by the Italian National Statistical Office, ISTAT. The register includes the year of establishment of the faculty, which is the year when the faculty is formally established as provider of higher education. HIU also includes name and address of each university, its governance structure (private or public), and assessments by the national government (in A,B,C-level). All over-time changes e.g. to the governance of the faculty or the university, or in the government assessments are also recorded in the data (See Appendix for details).[17]

We focus on faculties that deliver standard BSc education i.e. drop from the sample HE institutions specialised in post-graduate education, and enrolling foreign students only. Our working HIU sample includes 574 faculties (in 71 universities) registered in the Italian territory at some point between 1861 and 2010. We use this HIU working sample to construct a panel cross-province version of the data which follows over-time HE supply in Italian provinces, during 150 years of Italian history. This is a province by year dataset, which, for each province i existing at time t, counts the number of faculties present in its territory. The panel dimension is unbalanced, because the number of provinces changes throughout the time span. We exclude from the main analysis the first 10 years of Italian history, as the unification process was still ongoing, featuring the addition of new provinces to the Italian Kingdom. The final cross-province panel HIU sample includes 11792 observations for 110 provinces between 1870 and 2010. The main variable is the number of faculties in province i, as a total and by macro-area of science (Humanities, STEM and Social Sciences). Following the structure of the register data, we also record information on the total number of universities, total number of private universities, and total number of A,B,C-level universities in province i at time t.

Figure 4 gives an historical overview on the total number of faculties in 1870 as compared to 2010, at the 2010 province level. We observe a major expansion and geographical dispersion of higher education, which reflects the pattern of development described in institutional reports (OECD, 1999; OECD, 2008). As early as 1870, the pre-unitarian distribution featured a HE supply that was concentrated in the capitals of pre-unitarian states, with the highest number of faculties in a province being 8 (in Torino, compare Table 1). In 2010, almost all provinces have "their own" local HE supply. Those with zero faculties are the most recent provinces, established as autonomous administrative units in the post-war period. A few provinces display a HE supply of 15 faculties or more (Bari, Bologna, Milano, Napoli,

[17]Some of this information is not directly relevant to the present analysis e.g. the dates of some institutional changes that occurred at the faculty level. Notice that our data may not map faculty closures as precisely as faculty openings. However none of our historical sources mentions significant waves of closure of HE institutions during Italian history. Also HIU data do not record overtime changes in the addresses of universities and faculties. These information are collected as for 2010.

Figure 4: Diffusion of faculties in Italy



Note: authors' calculation on HIU data

and Roma). All of them include on their territory a metropolitan city.

Figure 4 suggests a spatial component in the expansion of HE supply during Italian history. We construct several matrices to model spatial interactions in the supply of higher education between neighbouring provinces. The first one is a *contiguity* matrix: we define as neighbours provinces that share the same border, regardless of their distance. A second matrix is based on *linear distances*: we define neighbouring provinces on the basis of their linear distance, regardless of whether they are contiguous or not. For most part of the analysis, we use and combine these two matrices. We also provide robustness checks using alternative distance matrices based upon travel distance, and travel time between province couples (details are in Appendix A.2).

Table 2 presents summary statistics for the main variables. Panel A describes local HE supply. On average, each province has a supply of over 2.5 faculties, mostly concentrated in STEM and Social Sciences, and about 0.5 universities, of which the 15% are private, and over the 75% are classified as A-level.[18] In next panels we describe the supply of higher education of neighbouring provinces. Panel B reports HE supply of "average neighbour" $-i$, computed as the average of the HE supply of all neighbours of province i, based on the contiguity matrix. These figures are obviously very similar to those displayed in Panel A. A representative neighbour, on average has about 2.9 faculties, mostly in STEM and social sciences, and roughly 0.5 universities, mostly public and A-level. Panel C reports instead the HE supply of each individual neighbour j. These figures are very similar to those displayed in

[18]The high percentage of A-level universities is due to the fact that the De Vecchi-Bottai Reforms, starting from 1933 classify all public universities as A-level.

Table 2: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Local HE Supply					
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 fac. 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: HE Supply of average neighbour, $-i$					
no. of faculties in $-i$	2.884	2.168	0	17.333	11792
no. of humanistic faculties $-i$	0.536	0.544	0	4.333	11792
no. of stem faculties in $-i$	1.24	0.909	0	5	11792
no. social sciences fac. in $-i$	1.098	0.873	0	8.333	11792
no. of universities in $-i$	0.551	0.398	0	2.667	11792
no. of private universities in $-i$	0.089	0.196	0	1.333	11792
no. of A-level universities in $-i$	0.422	0.31	0	1.667	11792
Panel C: HE Supply of neighbour, 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 fac. 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

Note: authors' calculation on HIU data

Panels A and B. However, allowing pairwise relationships between provinces i and j, increases the sample size by roughly four times (as each province i has roughly four neighbours js, on average), and avoids averaging down the HE supply of neighbour (as shown by comparing the maximum no. of faculties in Panels B and C).

4 Competition in HE supply

4.1 Empirical strategy

4.1.1 Empirical model

We use our province by year panel dataset to define a model of horizontal spatial interactions where the supply of higher education F in the i-th Italian province at time t is influenced by the supply of higher education in the neighbouring provinces. In our baseline model, they are those provinces that share a border with the i-th province, and may belong to the same

region or to a different one.

The spatial interaction can be modelled in several ways, featuring the standard approaches in the literature. First, we assume that HE supply in province i reacts to HE supply of its "average neighbour" (Brueckner, 2003). This is the weighted average of HE supply among the neighbours of the ith province, which is defined as $\overline{F}_{-it} = \sum_{j \neq i}^{N} w_{ij} F_{ij}$, where w_{ij} is a set of weights. We model this relationship through the following linear spatial competition model:

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

where β is the coefficient of interest as it captures local strategic interactions among close provinces. γ_i and $\delta_{r(i)t}$ are, respectively, province and region-by-year fixed effects (where $r(i)$ refers to the region province i belongs to at time t). X_{it} is a vector of time varying province level covariates. This includes the total number of universities, number of private and A-level universities in the province. In several specifications we also include province level demographic and economic controls that account for potentially spatially correlated factors.[19]

Province and region-by-year fixed effects in equation (1) control for a great deal of unobserved heterogeneity. However, to better account for the omitted determinants of one province's HE supply that are spatially correlated with those of the neighbours (e.g. related to geographical factors), we also estimate a "pairwise" version of equation (1) which takes the following form (see Parchet, 2014):

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

Equation (2) is estimated for all pairs of neighbouring provinces i-j. F_{jt} is the HE supply in each neighbouring province j of province i, c_{ij} is the fixed effect of the pair, $d_{r(i)t}$ are the region-by-year fixed effects and X_{it} is the usual set of province level controls. The main advantage of equation (2) relative to (1) is that the inclusion of pair fixed effects allows to better account for the time-invariant omitted factors that pertain to each $i-j$ couple. Each couple appears twice, with a given municipality being once on the left-hand side and once on the right-hand side of the equation. Each pair contributes separately to the estimates, and b captures the average response of province i to HE supply of neighbour j.

The extensive set of dummies in equations (1) and (2) account for all time-invariant

[19] In Table 8 below we account for the size of population (total and in the 0-14 age cohort), and size of the industry sector. We did not include these variables to the baseline specification as they are available from the Italian Census every ten years, in some cases for the post-war period only. Moreover, due to their persistence over time, they are not immune to endogeneity concerns.

characteristics at the level of each province, or province pairs. In both specifications, the region-by-year fixed effects control for all events at the region level that affect all the provinces in that region simultaneously and identically. The heterogeneity that is left after the inclusion of the fixed effects is basically the variability over time across provinces within the same region. Identification of β and b in equations (1) and (2) respectively is obtained through the comparison of different time-varying patterns in the number of faculties across provinces.

We estimate these two models using various measures of HE supply. Our main indicator 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)). We also distinguish the extensive margin of HE supply (i.e. having at least one faculty in province i and in the neighbourhood) from the intensive one (i.e. number of faculties, where HE supply is available). We alternatively measure HE supply in terms of number of universities available in province i and its neighbourhood.

The models (1) and (2) are in the standard fixed effects regression format, and can be easily estimated by OLS. In both models, a key aspect is the definition of neighbouring provinces, which determines the choice of the matrix of local interactions. As mentioned above, in the baseline specification, we defined as neighbours of i all provinces that share a border with the i-th province. In equation (1) this definition of neighbourhood is used to construct the term $\overline{F}_{-it} = \sum_{j \neq i}^{N} w_{ij} F_{ij}$. Here, weights are settled to reproduce sample means, i.e. any neighbour is given the same weight: $w_{ij} = \frac{1}{N}$. Throughout the analysis, we depart from this baseline matrix and adopt alternative neighbourhood's definitions based on e.g. linear distance, travel distance, and travel time measures.

Models (1) and (2) assume contemporaneous spatial competition. This overlooks some inertia, which may occur in the process of opening a new faculty. As discussed in Section 2 above, several actors are involved in the decision of opening a new faculty, which implies that the reaction of a province to its neighbours in the supply of higher education is likely not to be immediate. While it is difficult to define an appropriate lags' structure, we also experimented specifications with lagged effects. In our preferred empirical specification, we allow the spatial competition effect to be modelled as a ten years' lag: a given province at time t reacts to the supply of HE in the neighbourhood at time $t - 10$.

4.1.2 Identification and estimation issues

The main issue in the estimation of equations (1) and (2) arise because the number of faculties of a neighbouring set of provinces itself depends on the number of faculties of province i: any change of F_{ijt} may induce each neighbour to adjust F_{jt}. This gives rise to a reverse

causality problem. Moreover, many time-varying determinants of one province's number of faculties, such as local economic conditions or local demand for higher education, are likely to be unobservable and spatially correlated, such that $cov(\epsilon_{ijt}, F_{ijt}) \neq 0$. This would be for example the case of spatially correlated shocks to a neighbouring province. Even if such unobservable shocks were region specific, they would not be captured by region-by-year fixed effect unless all neighbours were in the same region as province i. This creates an omitted variables' issue that, together with reverse causality may bias the OLS estimation of (2). The average neighbour model (1) presents the same endogeneity problems, plus measurement error, due to the fact that the term $\overline{F}_{-it}$ is an aggregation of the individual F_{ijt}s.

To deal with endogeneity problems we follow an empirical strategy based on instrumental variables. The key for identification is to isolate variations in number of faculties of competing provinces that can be plausibly considered exogenous. We propose as instruments the initial geographical distribution of faculties of each neighbouring province (ICj), interacted with state-level university reforms (R_t). Our vector of instruments is $Z_{jt} = R_t * ICj$, where $R_t = 1$ if $year \geq t$ and 0 otherwise. These interactions convey the idea that, while state-level reforms affect in principle all provinces, the same general reform may have differential effects and shape HE supply differently, depending on the pre-unitarian endowment of higher education. As usual, instruments need to satisfy two conditions. First, they have to be exogenous to HE supply in province i, that is $cov(Z_{jt}, \epsilon_{ijt}) = 0$ in equation (2). Second, they should be relevant, i.e. they should imply enough variation in F_{jt}. The exogeneity assumption must hold for both initial conditions and reforms.

The geographical distribution of faculties in pre-unitarian Italian States is most likely to be exogenous. As discussed in Section 2 above, this reflected the HE policies of pre-unitarian states which were politically, culturally and linguistically fragmented. They were often in conflict, and had their own institutions so that they were not coordinating at all their decisions in any area of public provision, including higher education. As a result of the extension of the Casati framework to all provinces that progressively joined the Italian State, the HE system of the Italian Kingdom gradually accredited all the pre-existing HE institutions, with their specificities without imposing any substantial change. Hence, the Italian university system was inherited from the past, and the initial distribution of faculties across provinces was exogenous. While it is true that for any province the initial conditions are correlated with the subsequent development of the university system (relevance of the instrument), what is key for identification is that the initial distribution of faculties of the j and i provinces are uncorrelated. This is indeed the case, for the reasons we discussed in Section 2 and reprised here.

Since initial conditions are fixed over time, they would blur into the pair fixed effects. For this reason we use law interventions in the university system as another source of exogenous variation, which is fixed across provinces, but varies over time. We focus on general reforms of the university system, i.e. laws with a general purpose and not those intended to regulate some specific need of a limited set of universities. The Gentile reform in 1923 is the first attempt to create an organic national university system, coherent in its objectives and with a clear structure. The Devecchi-Bottai reform 1933, together with subsequent ancillary interventions, created a more overarching and centralised HE system. The 1969 and 1973 reforms liberalised university access to students from technical schools and legislated the opening of new faculties to cope with the increase in HE demand. The reforms implemented in the 1980s reorganised the internal governance of universities and the allocation of funding to higher education. The Bassanini reform in 1997 granted universities more financial and teaching autonomy.

In this case, the key point for identification is whether (upper-level) state decisions are arguably exogenous to the (lower-level) province decisions. The exogeneity holds under two assumptions. The first identifying assumption is that state-level reforms are not driven by some unobserved time-varying factors that also affect the number of faculties in the province i and its neighbours. This means assuming that region-by-year dummies capture all the aggregate component of province-specific shocks. The second requirement is that individual provinces do not systematically affect state-level higher education policies (no reverse causality). This rules out the possibility that a province willing to modify its supply of HE (measured by the number of faculties) has enough policy power to influence the State decisions about the general setting of the university system. This is likely to be the case here. On the one hand, provinces are the relevant jurisdictions for decisions on the supply of local public services such as higher education. On the other hand, they have very limited legislative competences and their political relevance itself having always been questioned during Italian history.[20] It is rather unlikely that any province has enough "voice" and bargaining power to direct general and national interest reforms that modified the governance of Italian universities, and implied a general expansion of the university system. On the Italian territory there is a sufficiently high number of provinces and the population is dispersed enough, to prevent specific province from having enough political power to influence

[20]Petracchi (1962) defines provinces "a 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". However the Republican Constitutional Law fails to mention their exact tasks and competences. It only states (art. 114) that they are autonomous bodies (as well as municipalities, metropolitan cities, and regions) with their own statutes, powers and functions according to the principles established by the Constitution (See also Fabrizzi, 2008a,2008b,2008c for historical overviews).

state-level decision. Furthermore, we address potential feedback effects from large provinces to national policy-making, by performing a robustness check where we exclude provinces hosting a metropolitan city (see Table 8 below).

Interacted with the initial conditions, these reforms are exogenous shocks that modified the incentives and the net benefits of changing the supply of higher education at the province level. As we explained above, we associate to each reform a dummy variable which equals one from the year of the reform onwards and zero before. The idea is that, at any t, each province and its neighbours face an university system that is the result of the stratification of current reforms and those inherited from the past. This layering process is measured by the number of ones in the set of dummies (i.e. the number of laws active) in a given year. Basically, identification comes by the number of reforms at which at time t a province has been exposed since the first year it appeared in the sample, interacted by its initial endowment of faculties.

An interesting feature that emerges from Figure 2, is that the number of faculties increases slowly until 1969, with a sharp and steadily increasing trend emerging afterwards. On the whole, our 2SLS model exploits one big discontinuity in 1969 plus additional smaller discontinuities. Post-1969 reforms further increased the autonomy of opening new faculties at the province level. In the pre-1969 period, the environment was more stable, but, if anything we expect more effects from the 1923 reform, which increased autonomy and the opportunity to open new faculties especially in traditional fields like scientific ones.

4.2 Main results

Table 3 presents the estimates of equation (1), based upon the average neighbour approach. We present four sets of estimates. In Row (a), we estimate the effect of the average number of faculties in the neighbourhood of province i (i.e. the total number of faculties present in the neighbouring provinces divided by the number of neighbours) on the total number of faculties in province i at time t. In Row (b), we isolate the extensive margin of competition in HE supply i.e. we describe HE supply in province i as a dummy variable, which is equal to 1 if at least one faculty is present in province i at time t, 0 otherwise. In the same way, we define the extensive margin of HE supply in the province's neighbourhood. In Row (c), we focus on the intensive margin of competition i.e. describe the effect of competition, only on provinces that have at least one faculty in place in provinces i and j at time t. Finally, in Row (d) we investigate competition effects when we measure HE supply in terms of the number of universities (rather than faculties) operating in each province.

Table 3: Competition in higher education: average neighbour approach

	[1]	[2]	[3]	Obs.
(a) average no. of faculties in $-i$	0.59***	–0.76***	–0.48***	11383
	(0.11)	(0.25)	(0.14)	
(b) at least one faculty in $-i$	0.31***	–0.24*	–0.21**	11383
	(0.11)	(0.12)	(0.09)	
(c) average no. faculties in $-i$ (int. margin)	1.27***	–0.86	–0.47	4064
	(0.32)	(0.53)	(0.30)	
(d) average no. of universities in $-i$	0.43***	–0.82***	–0.82***	11383
	(0.10)	(0.27)	(0.27)	
province FE	yes	yes	yes	
region-by-year FE	no	yes	yes	
provincial controls	no	no	yes	

Notes:. In Rows (a) and (c) the dependent variable is the total number of majors 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, 0 otherwise. In Row (d), the dependent variable is the total number of universities active in province i. The set of controls include 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. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

We adopt three different specifications. In Column [1], we present results from simple OLS, with province fixed effects. In Column [2], we include region-by-year fixed effects. In Column [3], we add province level controls. Standard errors are always clustered by province. Results from the preferred specification in Columns [3] show a negative coefficient, generally significant at conventional levels. Estimates in Row (a) suggest that a one standard deviation increase in neighbourhood's HE supply (equivalent to 2.17 faculties in the average neighbour) is associated with a reduction of HE supply in province i by $(2.17 * 0.48 =)1.04$ faculty, this effect being significant at the 1% level. This negative effect is present both on the extensive and the intensive margin of HE supply in Rows (b) and (c), while becoming less precisely measured in the latter. The negative effect is also present in Row (d), as we measure HE supply in terms of no. of universities in each province: an increase in the neighbourhood's HE supply by 1 university is associated with a reduction of the local HE supply by over 0.8 universities.

In Table 4 we present estimates of regression model (2) estimated for all pairs of border provinces (all provinces in the neighbourhood of i). In these estimates pair fixed effects replaced the province fixed effects. Results confirm that HE supply of neighboring provinces has a negative effect on the local HE supply. Comparison with average neighbour estimates in Table 3 suggests that accounting for the pair fixed effects reduces the magnitudes of the coefficients. Results in Row (a), column (3) suggest that, on average, a one standard

Table 4: Competition in higher education: pairwise approach

	[1]	[2]	[3]	Obs.
(a) no. of faculties in j	0.37***	–0.14**	–0.08***	50803
	(0.06)	(0.03)	(0.04)	
(b) at least one faculty in j	0.35***	–0.07**	–0.04*	50803
	(0.06)	(0.03)	(0.02)	
(c) no. of faculties in j (int. margin only)	0.73***	–0.16**	–0.05***	22728
	(0.15)	(0.06)	(0.02)	
(d) no. of universities in j	0.25***	–0.13**	–0.13**	50803
	(0.06)	(0.06)	(0.06)	
provincial pair FE	yes	yes	yes	
region-by-year FE	no	yes	yes	
provincial controls	no	no	yes	

Notes:. In Rows (a) and (c) the dependent variable is the total number of majors active 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, 0 otherwise. In Row (d), the dependent variable is the total number of universities active in province i. The set of controls include 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. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

deviation increase in HE supply in a neighbouring province (equivalent to 4.43 faculties) is associated with a reduction of HE supply in province i by $(4.43 * 0.08 =)0.35$ faculties. This negative effect is present both on the extensive and the intensive margin of HE supply, and when we measure HE supply in terms of no. of universities rather than no. of faculties in each province.

As we discussed above, we resort to 2SLS to tackle endogeneity concerns. Tables 5 and 6 show results for the two stages. In Panels A and B we report estimates for the average neighbour and pairwise approaches, respectively. We always include the complete set of fixed effects (FE) and provincial controls, as in column [3] of Tables 3 and 4. We present different 2SLS FE specifications, as well as their OLS fixed effects (OLS FE) counterpart to ease comparison. In Column [1], we report estimates from the baseline specification (Row (a) in Tables 3 and 4). In Column [2], we focus on the 69 "pre-unitarian provinces" that existed as territorial and administrative units at the moment of the Italian unification. This is meant to avoid any reverse causality concern e.g. from public policy, to the creation of new provinces.[21] Finally, in Column [3], we acknowledge that the reaction of province i to HE supply of province j is not instantaneous and allow for a 10 years' lagged response of each province to its neighbours.

[21]In the baseline set of estimates in column [1] we used all 110 provinces. Accordingly, in 2SLS estimates we set to zero the initial conditions of the 41 provinces that were created during the Italian history.

Table 5: 2SLS estimates: initial conditions, reforms and HE supply ($1st$ stage).

	[1]	[2]	[3]
Panel A) average neighbour approach			
(IC_{-i})*(L. 2102/1923)	0.19	0.21	0.25
	(0.16)	(0.16)	(0.16)
(IC_{-i})*(L. 1592/1933)	–0.04	0.24	0.26
	(0.10)	(0.18)	(0.17)
(IC_{-i})*(L. 910/1969)	0.21***	0.31***	0.30***
	(0.05)	(0.08)	(0.08)
(IC_{-i})*(L. 766/1973)	0.09**	0.06*	0.06*
	(0.04)	(0.04)	(0.04)
(IC_{-i})*(L. 382/1980)	0.07	0.03	0.03
	(0.04)	(0.05)	(0.05)
(IC_{-i})*(L. 168/1989)	0.04	0.04	0.04
	(0.03)	(0.05)	(0.05)
(IC_{-i})*(L. 245-341/1990)	0.19***	0.07	0.06
	(0.10)	(0.10)	(0.10)
(IC_{-i})*(L. 59/1997)	0.26***	0.20	0.09
	(0.09)	(0.16)	(0.12)
Observations	11383	8626	8614
Panel B) pairwise approach			
(IC_j)*(L. 2102/1923)	0.11***	0.12***	0.15***
	(0.04)	(0.04)	(0.04)
(IC_j)*(L. 1592/1933)	0.10***	0.13***	0.15***
	(0.03)	(0.03)	(0.02)
(IC_j)*(L. 910/1969)	0.22***	0.25***	0.24***
	(0.02)	(0.02)	(0.02)
(IC_j)*(L. 766/1973)	0.05***	0.05**	0.06***
	(0.02)	(0.02)	(0.02)
(IC_j)*(L. 382/1980)	0.05***	0.02	0.01
	(0.02)	(0.01)	(0.02)
(IC_j)*(L. 168/1989)	0.04***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)
(IC_j)*(L. 245-341/1990)	0.21***	0.13***	0.13***
	(0.04)	(0.03)	(0.03)
(IC_j)*(L. 59/1997)	0.26***	0.18***	0.14***
	(0.04)	(0.05)	(0.03)
Observations	50803	35510	35060

Notes: First stage of IV FE estimates reported in Table 6. 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%.

First stage estimates in Table 5 shows that reforms interacted with the initial HE conditions have a positive impact on HE supply in neighbouring provinces. Significance is generally higher in Panel B compared to Panel A, due to the smaller standard errors with pairwise estimation. Summing up the significant coefficients from pairwise estimates suggests that reform effort that took place during Italian history induced each neighbouring province

Table 6: 2SLS estimates: competition in higher education supply (2^{nd} stage).

	[1] baseline sample		[2] pre-unitarian provinces		[3] pre-unitarian prov. (lagged 10 y.)	
	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE
Panel A) average neighbour approach						
total no. faculties in $-i$	–0.48***	–1.14***	–0.09	–1.08**	–0.11	–0.94*
	(0.14)	(0.39)	(0.08)	(0.49)	(0.10)	(0.50)
Observations	11383	11383	8626	8626	8614	8614
K-P rk Wald F-stat		7.167		3.838		3.905
K-P rk LM-stat		25.392		15.106		17.199
(p-value)		0.001		0.057		0.028
Hansen J-stat		7.636		4.017		4.475
(p-value)		0.366		0.778		0.724
Panel B) pairwise approach						
total no. faculties in j	–0.08***	–0.16***	–0.04**	–0.13**	–0.05*	–0.12***
	(0.03)	(0.05)	(0.02)	(0.06)	(0.02)	(0.04)
Observations	50803	50803	35510	35510	35060	35060
K-P rk Wald F-stat		35.474		25.864		22.231
K-P rk LM-stat		61.609		42.057		43.617
(p-value)		0.000		0.000		0.000
Hansen J-stat		1.644		6.282		3.620
(p-value)		0.977		0.507		0.822

Notes:. Regressions in column [1] exploit the baseline sample i.e. 110 provinces during the entire period 1871-2010. Regressions in column [2] cover only the set of 69 pre-unitarian provinces for the entire period. Regressions in column [3] cover pre-unitarian provinces during 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 2SLS estimates the total no. of faculties in province $-i$, j is instrumented by the initial conditions (i.e. number of faculties in $-i$, j in 1861) interacted by a battery of dummies for higher education reforms in Italy. The full set of coefficients is in Table B-2 below. The full set of first stage coefficients is in Table 5. Standard errors clustered by province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

with a pre-unitarian HE supply to open about one faculty. Unsurprisingly, the most relevant reform is $L.910/1969$, which liberalised university access.

Table 6 presents results for the second stage. Coefficients estimated under pairwise approach in Panel B are smaller than their average neighbour counterpart in Panel A. However, they are also more precisely estimated and more stable across OLS and 2SLS specifications. F-statistics confirm that instruments have strong predictive power in the first stage of pairwise estimates, much less so under the average neighbour approach. The Hansen test confirms instruments provide valid exclusion restrictions in the second stage. Pairwise estimates in column [3] point to sizeable competition effects: HE supply in a province, with e.g. 4 neighbours, decreases by 1.12 faculties, when each neighbour sets up 2 new faculties. This negative effect can be interpreted as supply of HE services being perceived as close substi-

tutes by their consumers (i.e. students) when these are provided in neighbouring provinces. From now on, for sake of brevity we focus on our preferred specification and report only pairwise estimates, where all specifications include province pair fixed effects, region by year fixed effects, and the usual set of controls. [22]

4.2.1 Competition across and within field of study

In Table 7, we analyse the issue of competition in HE supply within and between fields of study. We group all faculties available in our sample into three major groups i.e. Humanities, STEM (i.e. Science, Technology, Engineering and Mathematics), and Social Sciences.[23] We run three sets of regressions for cross and within discipline competition in HE supply. Table 7 includes three Panels A-C, where our dependent variable is the local supply of HE in Humanities, STEM, and Social Sciences, respectively. For each panel, we present five different specifications. In Columns [1] and [3], we perform OLS FE and 2SLS FE estimates, considering as explanatory variable the neighbour's HE supply in the own field of study. In Columns [2] and [4] we include neighbour's HE supply in the other two fields. Finally, estimates in Column [5] are identical to those reported in Column [4], but coefficients are standardised to compare their magnitudes.

Results provide support to the view that HE supply in neighbouring provinces has a negative effect on the local HE supply within the same field of study. Conversely, there is not much evidence of negative effects between different fields of study. HE services provided by neighbouring provinces are perceived as close substitutes within the same field of study, while it seems difficult to consider e.g. a STEM faculty as a close substitute of a Humanities faculty. This suggests that differentiating HE supply from that of neighbours can be an effective policy to avoid competition.

4.2.2 Controlling for HE demand

In Table 8 we refine the analysis and account for factors related to HE demand, which vary at the level of the province and, if omitted, may affect the estimation of competition effects. In Panel A, we report the results from our preferred specification (Table 6, Panel B, Column [3]) to ease comparability of results. In Panel B, we check whether our results are driven

[22] Estimates with the average neighbour approach are qualitatively similar, but, for the reasons explained above, in general less precise. They are available upon request by the authors.

[23] Humanities include Education, Linguistic Studies, Literature, and Psychology. The STEM 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 Socio-Political Studies.

Table 7: Cross-disciplinary competition: humanities, stem, social sciences

	[1]	[2]	[3]	[4]	[5]	[6]
	OLS FE	OLS FE	2SLS FE	2SLS FE	2SLS FE	Obs.
Panel A) Humanities						32644
faculties$_j$, Humanities	–0.06*	–0.07*	–0.02	–0.07*	–0.08*	
	(0.04)	(0.04)	(0.07)	(0.04)	(0.04)	
faculties$_j$, STEM		0.01		–0.02	–0.03	
		(0.02)		(0.03)	(0.05)	
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	
K-P rk LM-stat			20.558	41.826	41.826	
(p-value)			0.008	0.007	0.007	
Hansen J-stat			7.641	24.824	24.824	
(p-value)			0.365	0.255	0.255	
Panel B) STEM						32644
faculties$_j$, STEM	–0.09***	–0.11***	–0.08*	–0.05**	–0.09**	
	(0.03)	(0.03)	(0.04)	(0.02)	(0.04)	
faculties$_j$, Humanities		0.02		–0.03	–0.03	
		(0.02)		(0.02)	(0.02)	
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	
K-P rk LM-stat			33.890	41.826	41.826	
(p-value)			0.000	0.007	0.007	
Hansen J-stat			4.878	15.791	15.791	
(p-value)			0.675	0.781	0.781	
Panel C) Social Sciences						32644
faculties$_j$, Social sciences	–0.04*	–0.04	–0.16**	–0.07**	–0.12**	
	(0.02)	(0.03)	(0.07)	(0.03)	(0.05)	
faculties$_j$, STEM		0.02		–0.01	–0.03	
		(0.03)		(0.02)	(0.04)	
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	
K-P rk LM-stat			34.300	41.826	41.826	
(p-value)			0.000	0.007	0.007	
Hansen J-stat			2.790	20.361	20.361	
(p-value)			0.904	0.498	0.498	

Notes:. Baseline pairwise estimates as in Table 6, Panel B, Column [3]. OLS FE estimates in Columns [1] and [2]. 2SLS FE estimates 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, Social Sciences in Panel C), in the neighbouring province. In Column [4] 2SLS FE estimates with three endogenous regressors i.e. the no. of faculties in Humanities, STEM and Social Sciences in the neighbouring province. The specification in Column [5] is the same as in Column [4], but regressors are standardised to have zero mean and unity standard deviation. In 2SLS estimates, the instruments are interactions of initial conditions (total no. of faculties in each discipline in 1861) with a battery of dummies for higher education reforms in Italy. See Table B-3 below for the first stage estimates. All specifications include provincial pair fixed effects, region-by-year fixed effects, the usual set of provincial controls for the local province. Standard errors clustered by province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 8: HE supply Competition: local demand effects

	[1] OLS FE	[2] 2SLS FE	[3] obs.
Panel A) baseline specification			34912
total no. faculties in j	–0.05*	–0.12***	
	(0.02)	(0.03)	
K-P rk Wald F-stat		22.241	
K-P rk LM-stat		43.612	
p-value		0.000	
Hansen J-stat		3.635	
p-value		0.821	
Panel B) drop metropolitan cities			22814
total no. faculties in j	–0.11**	–0.22***	
	(0.04)	(0.07)	
K-P rk Wald F-stat		37.514	
K-P rk LM-stat		25.099	
p-value		0.001	
Hansen J-stat		7.139	
p-value		0.415	
Panel C) control for population size			34912
total no. faculties in j	–0.04	–0.12***	
	(0.03)	(0.04)	
K-P rk Wald F-stat		20.689	
K-P rk LM-stat		40.629	
p-value		0.000	
Hansen J-stat		3.328	
p-value		0.853	
Panel D) control for participation and size of the industry sector			34912
total no. faculties in j	–0.05*	–0.12***	
	(0.02)	(0.04)	
K-P rk Wald F-stat		22.759	
K-P rk LM-stat		43.717	
p-value		0.000	
Hansen J-stat		3.779	
p-value		0.805	
Panel E) control for share of 0-14 years old			11442
total no. faculties in j	–0.05*	–0.09*	
	(0.03)	(0.05)	
K-P rk Wald F-stat		31.930	
K-P rk LM-stat		40.092	
p-value		0.000	
Hansen J-stat		2.576	
p-value		0.765	
Panel F) placebo: "alphabetical" neighbors			48337
total no. faculties in j	0.009	0.063	
	(0.020)	(0.061)	
K-P rk Wald F-stat		76.556	
K-P rk LM-stat		44.544	
p-value		0.000	
Hansen J-stat		8.124	
p-value		0.322	

Notes: In Panel A baseline pairwise estimates 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, Verona) are dropped from the sample. In Panel C, the total population is included in the set of controls. In Panel D, the share of active people in the industry sector, and the participation rate are included among the controls. In Panel E, the share of population belonging to the 0-14 cohort is included among the controls (available only for period 1951-2010). In Panel F, neighbors 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. In 2SLS estimates, the instruments are interactions of initial conditions with higher education reforms. The full set of coefficients is in Table B-4 below. Standard errors clustered at the province level. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

by the 14 Italian metropolitan cities.[24] Some of these cities (e.g. Milano) were not primary university centres in the 19th century, but became so only afterwards as a result of their population growth. We re-run our regressions after excluding them and our results are not altered. In the same spirit, in Panel C, we check how our results change if we include as a control the total population (in log).[25] Our results survive to its inclusion.

Demand for higher education may actually follow the gradual expansion of the industrial sector, and labour market participation during the 20th century. To account for that, in Panel C we include as controls the participation rate and the population share active in the industry sector. Also the inclusion of these controls makes no difference. Demographic changes in the composition of population may also determine changes in HE demand. In Panel D, we include the population share in the $0-14$ cohort. We were able to recollect this information only from 1951 onwards, and with some missing information for some provinces. For this reason the inclusion of this control produces a substantial loss of information. Our results however survive also to this robustness check. Finally, one may argue that our results are driven by some underlying omitted factors, which we are not able to control. If this was the case, the estimated competition effect would be likely to survive, whatever the neighbourhood relationship hypothesised. In Panel F we perform a placebo exercise imposing a non-sense alphabetical neighbourhood relationship: provinces are considered neighbours if their name starts with the same letter of the alphabet, regardless of their actual neighbourhood relationships (e.g. Alessandria is coded as neighbour of Agrigento, despite being over 1000 Kms aparts, and other provinces starting with an "A" only). The estimates show that the negative coefficient of the number of faculties in j disappears. This reassures us that the competition effect is specific to the neighbourhood matrix we have chosen.

4.2.3 The spatial reach of competition effects

Up to now we defined neighbours the provinces who share a border. Still, even provinces that do not share a border can be very close to each other in terms of distance. We used available Google Maps applications to compute alternative matrices based upon linear distance, travel distance and travel time between province capitals (see Appendix A.2 for details). Estimates using these alternative matrices allow us to discuss the "spatial reach" of competition, that

[24]These are identified by the Italian Constitution: Bari, Bologna, Catania, Firenze, Genova, Messina, Milano, Napoli, Palermo, Reggio Calabria, Roma, Torino, Verona.

[25]Notice that population data are drawn from the Italian Census data, which occurred about once every ten years during the period 1861-2011. We assigned each population-by-province data point to the successive years, until a new wave of census data is available.

Table 9: Competition in HE supply: linear distance matrix (in Km) and spatial reach

	[1] within 90 Km		[2] between 90 and 180 Km		[3] between 180 and 270 Km		[5] between 270 and 360 Km	
	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE
total no. faculties in j	–0.039**	–0.095**	–0.045***	0.003	0.003	–0.014	–0.013	0.001
	(0.018)	(0.039)	(0.014)	(0.017)	(0.007)	(0.010)	(0.011)	(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
K-P rk LM-stat		47.364		49.465		51.681		53.463
p-value		0.000		0.000		0.000		0.000
Hansen J-stat		9.190		4.349		7.142		10.233
p-value		0.239		0.739		0.414		0.176

Notes:. Baseline pairwise estimates. In 2SLS estimates, the total no. of faculties in province j is instrumented by the initial conditions interacted by higher education reforms in Italy. See Table B-7 below for details. Standard errors clustered by province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

is the distance up to which HE supply in a province responds to HE supply set elsewhere.

In Table 9 we report estimates using the definition of neighbourhood based on exogenous linear distance between provinces. Columns [1] to [5] show how HE supply of each province reacts to the HE supply of provinces located at different linear distances, on average. Results suggest that HE supply in a province reacts more to the supply of closer provinces. The effect is concentrated within a linear distance of 90 kilometres. Negative coefficients appear also for higher distances, however the effects are much smaller in size, and generally non statistically significant. Any evidence of HE supply competition disappears for linear distances beyond 270 kilometers.

In Table B-5 in the appendix, we report results for matrices based on travel distance and travel time. In Panels A-B of Table B-5 we define neighbours on the basis of the travel distances, and travel time between the provinces' capitals. Using these alternative measures we find that competition effect is concentrated among a travel distance of 120 kilometres and a travel time of 80 minutes. The spatial reach of strategic interactions can thus be estimated to be somewhat within these ranges. The effect disappears for travel distances beyond 360 kilometers and travel time exceeding 3 hours.

4.2.4 Heterogeneous effects

We perform three additional exercises to investigate the heterogeneity in the competition of HE supply. First, we report the results dividing the sample into two periods across World War II (WWII): a pre-war period is defined as before year 1936, while post-war as after year 1950. Results, reported in Panel A of Table 10, confirm that HE supply in neighbouring provinces has a negative coefficient both for pre-war period (Column [1]) and the post-war period (Column [2]), although being less precise in the pre-war period.

Second, to capture long-lasting differences in development of HE system within Italy we also perform an analysis by two macro regions. In Panel B, we present results for Northern regions which belong to NUTS1 North-East and North-West (Column 1), and Centre-South regions belonging to NUTS1 Centre, South and Islands (Column 2). The estimates show that the negative effect is still present in both Columns [1] and [2], being more precise for 2SLS FE compared to OLS FE estimates.

Finally, estimates in Table 6 consider homogeneous interactions between neighbouring provinces, regardless of the NUTS2 region they belong to. In Panel C we analyse heterogeneous interactions between provinces that belong to the same region (Column [1]), and provinces that belong to different regions (Column [2]). Interestingly enough, the competi-

Table 10: HE supply Competition: heterogeneity

Panel A) pre-WWII vs. post-WWII period				
	[1] pre-war		[2] post-war	
	OLS FE	2SLS FE	OLS FE	2SLS FE
total no. faculties in j	–0.03	–0.11	–0.04*	–0.10**
	(0.02)	(0.09)	(0.02)	(0.05)
Observations	17283	17283	14330	14330
K-P rk Wald F-stat		32.639		30.752
K-P rk LM-stat		19.372		41.734
p-value		0.000		0.000
Hansen J-stat		–		2.637
p-value		–		0.756
Panel B) North vs. Centre-South				
	[1] North		[2] Centre-South	
	OLS FE	2SLS FE	OLS FE	2SLS FE
total no. faculties in j	–0.05	–0.14**	–0.04	–0.09**
	(0.03)	(0.06)	(0.03)	(0.04)
Observations	17785	17785	17275	17275
K-P rk Wald F-stat		19.167		39.900
K-P rk LM-stat		22.602		22.540
p-value		0.004		0.004
Hansen J-stat		4.948		6.881
p-value		0.666		0.441
Panel C) intra-regional vs. inter-regional competition				
	[1] intra-regional		[2] inter-regional	
	OLS FE	2SLS FE	OLS FE	2SLS FE
total no. faculties in j	–0.11**	–0.18***	0.03	–0.06
	(0.04)	(0.06)	(0.04)	(0.07)
Observations	20234	20234	14687	14687
K-P rk Wald F-stat		21.651		12.205
K-P rk LM-stat		37.903		24.197
p-value		0.000		0.002
Hansen J-stat		5.331		1.947
p-value		0.620		0.963

Notes: Baseline pairwise estimates. In Panel A, pre-WWII and post-WWII periods are defined as before 1936 and after 1950, respectively. In Panel B, North regions belong to NUTS1 "North-East and "North-West", while Centre-South regions belong to NUTS1 "Centre", "South" and "Islands". In Panel C, intra-regional competition is between provinces in the same NUTS2 region, while inter-regional competition is between provinces belonging to different NUTS2 regions. In 2SLS estimates, the instruments are interactions of initial conditions (total no. of faculties in 1861) with higher education reforms in Italy in each sub-period. See Table B-6 below 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 at the province level. Significance levels: $*$: 10% $**$: 5% $***$: 1%

tion effect is mostly concentrated within the same region. This suggests that substitutability between HE institutions is lower when these are located in different regions, even though they are in neighbouring provinces.

4.3 Sensitivity analysis

4.3.1 Alternative definitions of pre-unitarian provinces

In our preferred specification, we define as pre-unitarian provinces that existed as administrative units at the onset of Italian history. However, we allow their territory and borders to change overtime, alongside the creation of new provinces during Italian history.

We take this modelling choices to be as precise as possible in the measurement of the local neighbourhood effects. However this is itself not immune to criticisms. It is well known that early Italian governments re-aggregated most provinces of pre-unitarian states in larger ones on the basis of political, economic, administrative and demographic considerations (Palombelli, 2012). Thus, our definition of pre-unitarian province may not fully serve the purpose of exogeneity, provided that aggregation might be carried out considering also the territorial provision of HE services. It may also be argued that it is not only the administrative dimension of province that matters, but the territorial subdivision. In that case exogeneity would require to restore provinces' pre-unitarian territories and borders.

In Table 11 below, column [1] we adopt a definition of pre-unitarian province that encompasses not only their existence as administrative units, but also their territory and borders in 1870. In order to do so, we assign all faculties set up in a province created during Italian history to the territory of the corresponding pre-unitarian province. We define neighbourhood relationship accordingly, on the basis of these pre-unitarian borders, which we maintain constant during Italian history.[26] Our main results are preserved. However, the decrease in size and significance of OLS estimates suggests the adoption of this definition of pre-unitarian province may induce some measurement error. In the next two columns of Table 11 we build upon the consideration that the 41 provinces created during Italian history are the same provinces of pre-unitarian states that were downgraded to districts of the 69 we use in the baseline analysis (Palombelli, 2012). If this is the case, the process of provinces' creation

[26]To make some examples, the faculties of the province of Pescara belong to the pre-unitarian province of Chieti (as the corresponding territory was then assigned to the new-born province of Pescara in 1927), which we assume always shares borders with the pre-unitarian territories of l'Aquila, Teramo and Campobasso during Italian history. Similarly we assign HE supply set up in the province of Taranto to the pre-unitarian territory of Lecce (to which the province of Taranto belongs), in the neighbourhood of the pre-unitarian provinces of Bari and Potenza (which shared borders with Lecce in 1870, but never with Taranto).

Table 11: Competition in HE supply: alternative definitions of pre-unitarian provinces

	[1] existing in 1870, borders at 1870		[2] existing in 1870, borders in 2010		[3] existing in 2010 borders in 2010	
	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE
total no. faculties in j	–0.01	–0.09**	–0.02	–0.07*	–0.08***	–0.09***
	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.03)
Observations	42813	42813	42813	42813	68410	68410
K-P rk Wald F-stat		36.130		41.221		48.375
K-P rk LM-stat		42.550		44.206		71.220
p-value		0.000		0.000		0.000
Hansen J-stat		6.046		6.797		10.327
p-value		0.534		0.450		0.171

Notes:. Baseline pairwise estimates. Borders and territory of pre-unitarian provinces are maintained constant over the entire period $1870-2011$. In column [1] pre-unitarian provinces are defined by borders and territory at 1870 i.e. the HE supply of the territory of provinces created during Italian history is reassigned to the pre-unitarian province from which the new provinces have been created. In column [2] pre-unitarian provinces are defined by borders and territory in 2010. In column [3] pre-unitarian provinces are defined as administrative units in 2010 at their 2010 borders and territories. Standard errors clustered by province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

during Italian history actually features these 41 districts gaining autonomy again i.e. restores the true pre-unitarian distribution. In column [2] we define pre-unitarian provinces the administrative units that existed in 1870, but at their 2010 borders. In column [3], we define pre-unitarian provinces all administrative units that existed in 2010 at their 2010 borders. In both cases, we maintain these definitions constant over the 150 years of Italian history. Our results are again confirmed alongside size and significance of the estimated coefficients.

4.3.2 Alternative empirical strategies

In Table 12, we check the sensitivity of our results to the adoption of alternative empirical strategies. In column [1], we perform a robustness check in the spirit of De Giorgi et al. (2010) and include in the set of instruments the 2^{nd} degree spatial lag of the HE supply of the local province. The identifying assumption is that all provinces z that share a border with province j, but not with province i, are valid instruments to identify the impact of HE supply of province j on local supply in i. Our main results are preserved however the Hansen J-test rejects the exogeneity of the instruments. This suggests that there are other factors over and beyond sharing a border that matter in determining local interactions. These factors undermine the validity of this instrument in the present setting. In column [2], we concentrate on the decade 1966-1976 only. In this way, we focus on an "event study" that uses the major liberalisation entailed by the 1969 and 1973 reforms to identify

Table 12: Sensitivity analysis: alternative empirical strategies

	[1] 2SLS FE	[2] 2SLS FE	[3] 2SLS FE	[4] 2SLS FE	[5] 2SLS FE	[6] 2SLS FE
total no. faculties in j	–0.12**	–0.19*	–0.14***	–0.13***	–0.16***	–0.12**
	(0.06)	(0.11)	(0.05)	(0.05)	(0.05)	(0.05)
Observations	109817	2653	34904	35060	32210	35328
K-P rk Wald F-stat	29.118	27.529	28.608	58.178	23.230	63.414
K-P rk LM-stat	39.760	23.703	42.723	44.423	41.329	58.088
p-vale KP	0.000	0.000	0.000	0.000	0.000	0.000
Hansen J-stat	17.212	0.014	2.540	3.823	4.908	7.434
p-value Hansen	0.028	0.906	0.924	0.800	0.427	0.385

Notes: Baseline pairwise estimates. In column [1], the set of instruments includes the 2^{nd} degree spatial lag of the local province. Regressions in column [2] cover the sub-period 1966-1976 only. In column [3] region-by-year FE for the region of the neighbouring province are included. In column [4], we include the number of universities (total, A-level, private) of the neighbouring province as additional controls. In column [5], we lag instruments ten years relative to the dependent variable in the first stage. In column [6] we define neighbours as provinces that have a shared border, within the spatial reach of 90 Km. Standard errors clustered by province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

competition effects.[27] The drop in the number of observations reduces the precision of the estimates, however our results are not altered. In columns [3] and [4], we include region-by-year dummies of the neighbouring province, and its HE supply controls, respectively. Also in this case our results are confirmed. Results do not change in column [5] either, as we consider a ten years lagged effect of the instrument on the HE supply of the neighbour in the first stage. Finally, in column [6] we consider competition between provinces that share a border and within a spatial reach of 90 Km. Our results are confirmed in this case too.

5 Economic value of HE supply

In this section, we analyse the economic value of HE supply. Besides being an interesting exercise per se, this is also useful to compute the opportunity cost of competition forces, in terms of value added per capita that is lost when not setting up a local faculty due to competition forces.

To this purpose, we estimate the following model:

$$ln(Y_{it}) = \alpha + \beta F_{it-10} + X'_{it}\gamma + \delta_i + \mu_{r(i),t} + \epsilon_{it}. \quad (3)$$

[27] This exercise can also be viewed as an additional robustness check that our results are not determined by omitted time varying factors associated with overtime changes to territory and borders of provinces. In fact the number and borders of provinces remain constant during this decade, the only exception being the creation of the Province of Isernia in 1971.

where Y_{it} is (log) value added per capita in province i, at time t; F_{it} is the total number of faculties in province i at time t; X_{it} is a vector of province level controls, i.e. the rate of population growth, the participation rate and the size of the industry sector in province i at time t (as measured by the share of active population in the industry sector).[28] δ_i is a province fixed effect, and $\mu_{r(i),t}$ is a region-by-year FE. Finally, ϵ_{it} is the error term.

Our main parameter of interest is β: this measures the marginal value of one faculty in terms of local value added per capita. Two endogeneity issues complicate the estimation of equation (3). First, reverse causality going from economic performance to HE supply: richer and/or more productive provinces may express a larger (or smaller) demand of HE. Thus, the opening of a new faculty may result from this demand. Second, omitted factors may motivate both an increase in value added per capita and the opening of a new faculty.

We try and address these concerns implementing the same instrumental variables estimator, which uses pre-unitarian supply of higher education in the province, interacted with state-level reforms $Z_{it} = IC_i * R'_t$ as an instrument for F_{it} (both lagged ten years) in equation (3). We already discussed extensively the exogeneity of initial conditions and the reforms. Being prior to the formation of the country, and specific to the several pre-unitarian States present on the Italian territory, the pre-unitarian HE supply is pre-determined, thus not affected by the demand for education that emerged after the unification (Squicciarini and Voigtländer, 2015). The additional identifying assumption we need for Z_i being a valid instrument for F_{it} in equation (3) is that, conditional on the large set of province and region-by year fixed effects, the interactions between state level reforms and provinces' HE initial conditions, do not have a direct effect on the local value added per capita.

Information on value added, population, and participation come from the Italian Census Data, which are collected roughly every 10 years. This means that we exploit a ten years' variation to estimate equation (3). During the 150 years' sample period, we observe provinces at most 16 times, i.e. $T = 16$ in our unbalanced panel. This changes the interpretation of the state-level reforms. In some cases R'_t captures exposure to "reform packages" implemented during the decade t i.e. $R'_t = 1$ if $decade > t$, 0 otherwise.[29]

Table 13 reports results from various specifications of model (3). We start by presenting the effect of local HE supply on economic performance, i.e. we consider the HE supply of

[28] In their analysis of the HE determinants of the industrial revolution in Prussia, Becker et al. (2011) consider the share of population active in the industry sector as the main outcome variable. We use it rather as a control, as our time period is posterior to the industrial revolution.

[29] More precisely the dummy $R_{21/30}$ includes only the L. 2102/1923. Similarly $R_{31/40}$ includes the L. 1592/1933 and $R_{61/70}$ includes the Law 910/1969. Conversely $R_{71/80}$ includes a package composed by both Laws 766/1973 and 382/1980. Similarly, $R_{81/90}$ includes L. 168/1989 and Laws 245-341/1990. Finally $R_{91/01}$ covers Laws 59-127/1997.

Table 13: Economic value of HE supply and externalities

	[1] OLS	[2] 2SLS FE	[3] 2SLS FE	[4] 2SLS FE	[5] 2SLS FE LIML	[6] 2SLS FE	[7] 2SLS FE LIML	[8] 2SLS FE	[9] 2SLS FE LIML
Faculties in i	0.0077***	0.0070*	0.0081*	0.0133***	0.0108***	0.0144**	0.0109**	0.0114***	0.0112***
	(0.0017)	(0.0041)	(0.0042)	(0.0046)	(0.0040)	(0.0055)	(0.0044)	(0.0022)	(0.0021)
Faculties in $-i$				0.0073**	0.0063**	0.0087*	0.0069**	0.0029**	0.0036***
				(0.0035)	(0.0028)	(0.0045)	(0.0032)	(0.0012)	(0.0013)
Observations	937	937	937	937	937	937	937	874	874
$-i$ within 90 Km	–	–	–	no	yes	no	yes	no	yes
Instruments	no	IC*R	IC*R	IC*R	IC*R	IC*R	IC*R	IC*R	IC*R, F_{t-20}
Reforms (R)	–	all	three	three	three	three	three	three	three
K-P rk Wald F-stat	–	12.659	11.801	3.011	8.872	3.011	8.872	49.439	69.857
K-P rk LM-stat	–	25.160	23.065	17.437	15.219	17.437	15.219	24.848	24.058
p-value KP	–	0.000	0.000	0.004	0.009	0.004	0.009	0.001	0.001
Hansen J-stat	–	13.742	4.601	5.286	5.890	5.142	5.895	7.710	6.821
p-value Hansen	–	0.017	0.100	0.259	0.208	0.273	0.207	0.260	0.338

Notes: In all specifications the dependent variable is value added per capita in log terms. The total number of faculties in the province and in the neighborhood are lagged 10 years. All specifications include as controls, growth in local population, the participation rate and the size of the industry sector. All specifications include province FE and region by year dummies. In specifications [3]-[9] the three reform packages in vector R are $R_{21/30}$, $R_{31/36}$, and $R_{21/30}$ only. Standard errors clustered at the province level in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

the local province the only determinant of local economic prosperity. In columns [1] and [2] we report OLS and 2SLS results. These results point to positive returns from HE supply, of very similar magnitudes in OLS and 2SLS estimates. A concern with these estimates is a likely collinearity between our instruments and the country-by-region dummies, due to the ten years' span of the time dimension available for this set of estimates. First stage estimates in Table B-9 confirm this suspect, as in our vector of reform packages only the first three have actual explanatory power in the first stage. From columns [3] to [9], we accordingly report 2SLS FE estimates, using only the relevant instruments in the first stage estimate of column [2] i.e. $IC_i * R_{21/30}$, $IC_i * R_{31/36}$, and $IC_i * R_{61/70}$. This greatly mitigates the problem of collinearity.

Estimates in columns [1] to [3] do not take into account the fact that opening new faculties in the neighbourhood may actually produce direct externalities on local economic activity. In columns [4] to [9] we investigate this possibility, by including either an externality effect coming from all the neighbourhood (columns [4], [6], and [8]) or an externality effect coming from the neighbours within the spatial reach of 90 Km (columns [5], [6], [7]). In these estimates, the F-test statistic of the additional instruments is below the critical value of 10. Accordingly, in columns [5], [7], and [9] we estimate the model by limited information maximum likelihood (LIML), which is median-unbiased in over-identified models. Also, in column [9] we include as instruments the 20 years lagged values of HE supply in the local province and neighbourhood, which increases the power of instruments in the first stage. Estimated coefficients in columns [4]-[9] are larger in size, more robust and precisely estimated than in columns [1]-[3].

Overall, results in Table 13 confirm the positive local returns of HE supply. Taken at their face value, these estimates suggest that, on average, one faculty more in the local province raises value added there by more that 1%. We can interpret this value as the local opportunity cost of not opening one faculty in the province. As we discussed in Section 4, this is roughly the result of competition coming from e.g. 4 neighbouring provinces, each one opening 2 faculties. However, our estimates also suggest that 8 faculties in the neighbourhood produce direct positive externalities on local value added per capita by almost $(8 * 0.3) = 2.4\%$. This implies that direct positive externalities coming from faculties located in the neighbourhood, more than compensate the local costs of competition effects.

Two remarks are in order. First of all, our results are in line with those from existing studies that use a similar empirical approach (Valero and Van Reenen, 2016). This is non negligible, as we look at one country only (Italy), a more disaggregated regional unit (NUTS3 instead of NUTS2), and a different measure of HE supply (faculties instead of universities).

Second, our results do not allow pointing out the exact channel that determines the effect of HE supply on value added per capita. Our results can be determined by either human capital accumulation (Ciccone and Peri, 2006; Bratti and Leombruni, 2014) or long-term changes in the industry composition and technologies (see e.g. Ciccone and Peri, 2011), or both. Pointing out the exact channel is left as an interesting subject for further research.

6 Conclusions

In this paper, we use an own-built historical dataset on the formation of higher education institutions in Italy (History of Italian Universities - HIU) to investigate local competition in the supply of higher education.

To this purpose, we use HIU data to analyse whether the existence of a higher education institution in a neighbouring province hinder the local supply of higher education. Because of spatial correlation and reverse causality issues, we estimate local competition effects, by implementing an IV strategy which exploits exogenous variation associated with the initial conditions of the Italian higher education system interacted with the most comprehensive reforms of higher education that took place over Italian history.

We find evidence that the expansion of HE institutions during Italian history, particularly post-WWII (OECD, 1999, 2008) was deeply shaped by higher education reforms and pre-unitarian supply. The main set of estimates reveals non-negligible local competition effects: on average a province which has 8 faculties in the neighbourhood reduces its local supply by more than 1 faculty. We also perform separate analyses and found that such local substitutability between faculties is mostly concentrated within the same field of study (humanities, stem, social sciences), within the same NUTS2 region, and in a spatial reach of 90 Km.

These findings have important implications for the higher education policy of developed countries. They suggest that local substitutability of higher education supply may act as a remarkable "discipline effect" to HE expansion from below. Starting from the 1960s, regional HE supply considerably increased in OECD countries (OECD, 1999). Based on our estimates, the cost associated with the ensuing HE competition may be non negligible in economic terms; however, it tends to be overcompensated by direct positive externalities coming from the neighbourhood. Overall, our findings suggest that national OECD governments should be willing to regulate, to favour an efficient provision of higher education services. The expansion of local higher education providers should be restrained, especially in smaller provinces, surrounded by metropolitan areas, which are subject to the strongest

competition effects. Differentiating HE supply (e.g. by field of study) can be a useful device to mitigate competition intensity, and feature an efficient design of local provision of higher education.

Acknowledgements

We would like to thank the Department of Economics and Finance, Università Cattolica del Sacro Cuore (UCSC) for selecting this project for the *Research Assistance Program*. We are particularly grateful to Ginevra Gallassi, Roberta Terranova, Giuseppe Grasso, and George Jacob for precious assistance at various stages of our research. We would also like to thank the personnel from the library "G. Billanovich, UCSC" for their valuable help in finding out materials used to build the HIU dataset and the Istituto Tagliacarne for the data on province value added. We finally wish to thank colleagues who attended presentations at the internal seminar of UCSC; the 2017 LEER Workshop in Leuven; 2017 IWAAE Workshop in Catanzaro; 2017 APET Conference in Paris; 2017 AIEL Conference in Cosenza, and the invited RUE seminar at the School of Economics of the University of Barcelona for very helpful discussions. The usual disclaimer applies.

References

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.

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.

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.

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.

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, G. (2005). The condition of the working class in england, 1209-2004. *Journal of Political Economy*, 113(6):1307–1340.

Davies, R. and Vadlamannati, K. C. (2013). A race to the bottom in labor standards? an empirical investigation. *Journal of Development Economics*, 103(C):1–14.

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 Mauro, T. (1970). *Storia linguistica dell'Italia unita*. Laterza, Bari.

Ertur, C. and Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, 22(6):1033–1062.

Eurydice (1999). Two decades of reform in higher education in europe: 1980 onwards. Eurydice studies, Eurydice.

Eurydice (2008). Higher education governance in europe: policies, structures, funding and academic staff. Eurydice thematic studies.

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

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

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

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.

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.

Hoxby, C. and Turner, S. (2014). What high-achieving low-income students know about college. *American Economic Review, American Economic Association*, 105(5):514–517.

Madsen, J. B. and Murtin, F. (2017). British economic growth since 1270: the role of education. *Journal of Economic Growth*, 22(3):229–272.

OECD (1999). *The Response of Higher Education Institutions to Regional Needs*. OECD Paris.

OECD (2008). *OECD Review of Tertiary Education. Final Report, OECD, Paris*. OECD Paris.

Palombelli, G. (2012). *L'evoluzione delle circoscrizioni provinciali dall'Unit d'Italia ad oggi*. I. Nicotra (eds.) Le Province dall'Unit alla Repubblica delle Autonomie. Donzelli editore, Roma.

Parchet, R. (2014). Are local tax rates strategic complements or strategic substitutes? IdEP Economic Papers 1407, USI Universit della Svizzera italiana.

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.

Pini, A. I. (2000). *Le "nationes" studentesche nel modello universitario bolognese nel medio evo.* Romano, Brizzi (a cura di), pp. 21-30. Bologna, CLUEB.

Schtz, 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.

Unioncamere (2011). Italia 150. le radici del futuro. il sistema camerale tra imprese e istituzioni a 150 anni dall'unita' d'italia. Retecamere scrl - edizioni camere di commercio ditalia, Unioncamere.

Valero, A. and Van Reenen, J. (2016). The economic impact of universities: Evidence from across the globe. NBER Working Papers 22501, National Bureau of Economic Research, Inc.

Woesmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2):117–170.

Appendix A Data Appendix

Appendix A.1 HIU Register data

The register dataset on the History of Italian Universities (HIU) contains detailed and complete information on institutions providing higher education in Italy, disaggregated at the faculty level over the period 1861-2010. The register includes the following information:

- University name.
- Faculty name.
- 15 faculty field identifiers (Agricultural studies, Architecture, Chemical and Pharmaceutical studies, Economics and Statistics, Physical Education, Geo-Biological studies, Law, Engineering, Educational studies, Litery studies and Philosophy, Foreign Languages, Medicine, Political Sciences, Psycology, and Mathematical Sciences), which we aggregated into 7 teaching areas (Socio-Economic area, Physical Education, Law, Engineering and Architecture, Medicine, Sciences, and Humanities.) and 3 macro-areas of science (Social Sciences, STEM, Science, Technology, Engineering and Mathematics - STEM, and Humanities), according to the classification used by the Italian National Statistical Office.
- Year of establishment of the faculty. This is recorded as the year when the faculty is formally established as a provider of a higher education degree. Alongside with this basic information, we recorded other potentially important ancillary dates i.e. whether the faculty was built upon a pre-existing major of studies (e.g. belonging to an existing faculty), the year the faculty was formally recognised as a provider of University education, the date(s) when the faculty became part of a different university. Details over institutional developments that motivate these alternative definitions are available in the on-line appendix.
- Address of the university.
- Address of the faculty.

The final version of the register includes 582 faculties and (in) 78 universities registered on the Italian territory at some point between 1861 and 2010.

About the 99% of faculties in our sample deliver standard BSc education. There are 5 faculties specialised in post-graduate education only (i.e. belonging to Universita' Normale di

Pisa, Universita' Europea di Roma, Universita' di Scienze Gastronomiche, and IMT Lucca), and 3 faculties that enroll foreign students only (i.e. belonging to Universita' per Stranieri di Siena, Perugia, and Reggio Calabria).

While they track very precisely the creation of new HE institutions during Italian history, as well as their change in status and governance, our data record only seven cases of effective faculty closures. Four engineering faculties were closed due the re-organisation of the Politechnic School of Milan, which in 2000 closed down its campuses in Como and Lecco and opened brand new faculties in the Milan area. The faculty of Chemical studies was shut down by the Ca Foscari University of Venice, in 1990 as well as the faculties of Environmental Sciences and Mathematical Sciences in Urbino in 2006. This may underestimate the actual closure HE education institutions. As a matter of fact, our sources do not allow to map closures as precisely as start-ups. However, none of our original sources mentions significant waves of closure of HE institutions during Italian history. (see on-line appendix and Brizzi and Romano (2007) Vol.3 for details.).

Our data do not record changes in the exact address of each university and each faculty at each point in time i.e. do not record changes of address over time. The university address is identified by the address of the university dean, and the faculty address is the address of the faculty dean. Both are collected as for 2010.

Appendix A.2 Neighborhood and distance matrices

We constructed two alternative local interaction matrices. The first one (1) is a simple contiguity matrix available from ISTAT. The second one (2) is a distance matrix, which contains information on (a) linear distance, (b) travel distance, and (c) travel time. We computed distances using the Google Maps API Geometry Library, and Google Maps Distance Matrix API, respectively.

1 Contiguity matrix. Each province i is matched to (the HIU indicators of) its neighbors, which we define as provinces j that share a border with province i. Notice that due to the process of provinces creation discussed above, the neighbors of each province i in most cases change over time. In practice, consider the example of province k, which is a neighbor of province i at time t. Imagine that at time $t+1$ a new province z is established, which covers the geographical territory of k that shares a border with i. Our data register this changes overtime so that until time t province k is recorded as being neighbor of province i, while starting from time $t+1$ k is no longer adiacent to i, while province z appears in the data, which shares a border with i.

2a Linear distance matrix. Each province i is matched to (the HIU indicators of) all other provinces $j \neq i$, and for each (i, j) couple, the linear distance is recorded between i and j. This is the distance in Km calculated "on-air", by drawing, and measuring the lenght of a straight line between the capitals of province i and province j (capoluoghi di provincia). Notice that, as long as the capitals of provinces do not change over time, the distance between provinces remains constant over our sample. In the case of provinces with multiple "capoluoghi", the largest capoluogo is considered as capital. However there are only two cases in Italy (Barletta-Andria-Trani, and Pesaro-Urbino).

2b Travel distance matrix. Each province i is matched to (the HIU indicators of) all other provinces $j \neq i$, and for each (i, j) couple, the travel distance is recorded between i and j. This is the lowest distance in Km that one needs to travel from/to i to/from province j, with whatever transport available (plane, car, train). Notice that these distances are recorded in 2016, so they take a cross-sectional picture of travel connections between italian provinces in that specific year i.e. they do not take into account the process of infrastructures building that has occurred during italian history. The same convention as in the case of linear distances applies to Barletta-Andria-Trani, and Pesaro-Urbino.

2c Travel time matrix. Each province i is matched to (the HIU indicators of) all other provinces $j \neq i$, and for each (i, j) couple, the travel time is recorded between i and j. This is the lowest time in minutes, that one needs to travel from/to i to/from province j, with whatever transport available (plane, car, train). Also travel time is recorded in 2016, so it does not take into account the process of infrastructures building and innovation that has occurred during italian history. The usual convention as in the case of distances applies to Barletta-Andria-Trani, and Pesaro-Urbino.

Appendix A.3 Additional province level variables

We collected historical indicators on province level economic and social characteristics from Unioncamere (2011). Historical series were completed using the direct Census sources.

Total population: Population of residents in the province, all ages (Source: Unioncamere, 2011).

Total population in the 0-14 cohort: Population of residents aged 0-14 in the province (Source: Italian Census data, 1951-2010).

Provincial value added per capita (worker): i.e. total provincial value added divided the total population (active population) of the province (Source: Unioncamere, 2011).

Original historical series expressed in nominal terms (in lira, current values). Real figures were obtained by applying the VA deflator at constant 1911 prices, available from ISTAT.

Provincial participation rates: Active population, as a percentage of total population (Source: Unioncamere, 2011).

Share of active population in agriculture industry services: number of workers in agriculture industry services as a share of total workers (Source: Unioncamere, 2011).

Census data covers the period 1861-2010. The collection years are 1861, 1871, 1891, 1901, 1911, 1921, 1931, 1936, 1951, 1961, 1971, 1981, 1991, 2001, 2010. Census data are not available for 1881, so we retrieved them by linear interpolation.

Appendix B Additional Tables

Table B-1: Number of Faculties in Italy by field of study: 1870 and 2010

Faculty	1870	2010
1. HUMANITIES	16	127
Education	1	34
Languages	1	24
Literature	12	54
Psychology	2	15
2. SCIENTIFIC AND MEDICAL ST.	51	186
Agriculture	9	37
Chemistry&Pharmacy	18	31
Geology&Biology	1	3
Scientific studies	15	48
Architecture	3	23
Engineering	5	44
3. SOCIAL SCIENCES	42	210
Medical studies	19	39
Economics&Statistics	1	68
Law	21	55
Socio-political studies	1	48

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

Table B-2: Omitted neighbourhood variables and 2SLS: full set of coefficients

	[1] baseline sample		[3] pre-unitarian provinces		[4] pre-unitarian prov. (lagged 10 y.)	
	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE
Panel A) average neighbour approach						
total no. faculties in $-i$	–0.48***	–1.14***	–0.09	–1.08**	–0.11	–0.94*
	(0.14)	(0.39)	(0.08)	(0.49)	(0.10)	(0.50)
no. of universities	1.48***	1.36**	1.82***	2.00***	1.62**	2.11***
	(0.55)	(0.57)	(0.64)	(0.63)	(0.61)	(0.55)
no. of A-level universities	1.06***	0.86**	0.87**	0.31	0.95***	0.59
	(0.34)	(0.34)	(0.34)	(0.41)	(0.31)	(0.39)
no. of private universities	2.38***	2.20***	1.93***	1.41**	2.28***	1.71**
	(0.57)	(0.56)	(0.53)	(0.57)	(0.71)	(0.65)
Panel B) pairwise approach						
total no. faculties in j	–0.08***	–0.16***	–0.04**	–0.13**	–0.05*	–0.12***
	(0.03)	(0.05)	(0.02)	(0.06)	(0.02)	(0.04)
no. of universities	2.17***	2.20***	2.34***	2.48***	2.17***	2.37***
	(0.52)	(0.49)	(0.56)	(0.49)	(0.54)	(0.45)
no. of A-level universities	0.65**	0.64***	0.48*	0.46**	0.56**	0.57**
	(0.27)	(0.24)	(0.26)	(0.21)	(0.26)	(0.22)
no. of private universities	1.95***	1.79***	1.47***	1.29***	1.82***	1.77***
	(0.54)	(0.50)	(0.45)	(0.38)	(0.59)	(0.54)

Notes:. Full set of coefficients refers to OLS FE and 2SLS 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 2SLS estimates, the instruments are interactions of initial conditions with higher education reforms, as reported in table B-7. Standard errors clustered at the province level. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table B-3: Competition across and within field of study: first stage

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

Notes:. First stage of IV FE estimates reported in Table 7. 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 B-4: HE supply Competition: local demand effects

	[1] OLS FE	[2] 2SLS FE	[3] Obs.
Panel A) baseline specification			34912
total no. faculties in j	–0.05*	–0.12***	
	(0.02)	(0.04)	
no. of universities	2.22***	2.42***	
	(0.54)	(0.45)	
no. of private universities	1.78***	1.72***	
	(0.59)	(0.54)	
no. of elite universities	0.55**	0.56**	
	(0.27)	(0.23)	
Panel B) drop metropolitan cities			22814
tot. no. faculties in j	–0.11**	–0.22***	
	(0.04)	(0.07)	
no. of universities	1.93***	2.14***	
	(0.50)	(0.40)	
no. of private universities	0.27	0.16	
	(0.52)	(0.43)	
no. of elite universities	0.77***	0.71***	
	(0.27)	(0.23)	
Panel C) control for population size			34912
total no. faculties in j	–0.04	–0.12***	
	(0.02)	(0.04)	
no. of universities	2.01***	2.24***	
	(0.60)	(0.50)	
no. of private universities	1.68***	1.74***	
	(0.60)	(0.55)	
no. of elite universities	0.47*	0.54**	
	(0.27)	(0.23)	
total population (log)	1.37*	0.75	
	(0.76)	(0.68)	
Panel D) control for share of active in industry sector			34912
total no. faculties in j	–0.05*	–0.12***	
	(0.02)	(0.04)	
no. of universities	2.17***	2.42***	
	(0.55)	(0.45)	
no. of private universities	1.82***	1.72***	
	(0.58)	(0.53)	
no. of elite universities	0.55**	0.55**	
	(0.27)	(0.23)	
share of active in the industry sector	–0.01	–0.01	
	(0.01)	(0.01)	
participation rate	0.01	0.01	
	(0.01)	(0.01)	
Panel E) control for share of 0-14 years old			11442
total no. faculties in j	–0.05*	–0.09*	
	(0.03)	(0.05)	
no. of universities	1.92***	2.17***	
	(0.65)	(0.59)	
no. of private universities	1.96***	1.78***	
	(0.47)	(0.43)	
0-15 cohort size	0.03	0.05	
	(0.08)	(0.08)	
Panel F) placebo: "alphabetical" neighbors			48337
total no. faculties in j	0.009	0.063	
	(0.020)	(0.061)	
no. of universities	1.750***	1.796***	
	(0.587)	(0.554)	
no. of elite universities	0.986***	1.055***	
	(0.284)	(0.261)	
no. of private universities	2.421***	2.224***	
	(0.626)	(0.600)	

Notes: Full set of coefficients refers to OLS FE and 2SLS FE estimates reported in Table 8. All specifications include provincial pair fixed effects, and region-by-year fixed effects. In 2SLS estimates, the instruments are interactions of initial conditions with higher education reforms. Standard errors clustered at the province level. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table B-5: Competition in HE supply: distance and time matrices

Panel A) travel distance (Km)								
	[1] within 90 Km		[2] between 90 and 180 Km		[3] between 180 and 270 Km		[5] between 270 and 360 Km	
	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE
total no. faculties in j	–0.046**	–0.136***	–0.019	0.021	–0.016	–0.024*	0.009	–0.002
	(0.018)	(0.044)	(0.014)	(0.014)	(0.012)	(0.013)	(0.011)	(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
K-P rk LM-stat		42.040		52.940		49.915		55.750
p-value		0.000		0.000		0.000		0.000
Hansen J-stat		9.525		8.164		6.271		6.867
p-value		0.217		0.318		0.508		0.443
Panel B) travel time (minutes)								
	[1] within 80 minutes		[2] between 80 and 160 min.		[3] between 160 and 240 min.		[5] between 240 and 360 min.	
	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE
total no. faculties in j	–0.090***	–0.162***	–0.014	–0.003	–0.003	–0.018	0.017	0.007
	(0.028)	(0.045)	(0.013)	(0.014)	(0.008)	(0.012)	(0.011)	(0.007)
Observations	29403	29403	102981	102981	115585	115585	84469	84469
K-P rk Wald F-stat		28.582		193.543		178.643		73.552
K-P rk LM-stat		38.753		53.891		52.401		53.702
p-value		0.000		0.000		0.000		0.000
Hansen J-stat		9.270		9.724		5.520		6.045
p-value		0.234		0.205		0.597		0.534

Notes:. Specification as in Table 6 column [4] with pre-unitarian provinces during the entire period and 10 years lags of regressor and controls. In 2SLS estimates, the total no. of faculties in province j is instrumented by the initial conditions (i.e. number of majors in j in 1861) interacted by a battery of dummies for higher education reforms in Italy. See Table B-7 below for details. 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 B-6: HE supply Competition: heterogeneity (first stages)

	[1] pre-war	[2] post-war	[3] North	[4] Centre-South	[5] Intra-regional	[6] Inter-regional
(IC_j)*(L. 2102/1923)	0.15***	–	0.21***	0.11**	0.13**	0.20***
	(0.03)	–	(0.04)	(0.05)	(0.05)	(0.07)
(IC_j)*(L. 1592/1933)	–	–	0.14***	0.20***	0.23***	0.06
	–	–	(0.03)	(0.06)	(0.03)	(0.07)
(IC_j)*(L. 910/1969)	–	0.19***	0.18***	0.28***	0.23***	0.22***
	–	(0.02)	(0.02)	(0.03)	(0.02)	(0.04)
(IC_j)*(L. 766/1973)	–	–0.00	–0.03	0.02	–0.05*	0.08*
	–	(0.03)	(0.03)	(0.06)	(0.03)	(0.04)
(IC_j)*(L. 382/1980)	–	0.19***	0.20***	0.18**	0.17***	0.17**
	–	(0.04)	(0.04)	(0.07)	(0.04)	(0.07)
(IC_j)*(L. 168/1989)	–	0.19***	0.16***	0.23**	0.23***	0.08
	–	(0.05)	(0.06)	(0.09)	(0.07)	(0.06)
(IC_j)*(L. 245/1990)	–	–0.02	–0.05**	0.02	–0.02	–0.05
	–	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)
(IC_j)*(L. 59/1997)	–	–0.00	0.03	–0.04	–0.01	0.01
	–	(0.03)	(0.03)	(0.03)	(0.02)	(0.06)
Observations	17283	14330	17785	17275	20234	14687

Notes: First stage of IV FE estimates reported in Table **??**. 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 B-7: Competition in HE supply: alternative definitions of pre-unitarian provinces: 1st stage

	[1]	[2]	[3]	[4]
(IC_j)*(L. 2102/1923)	0.16***	0.17***	0.19***	0.13***
	(0.03)	(0.03)	(0.03)	(0.04)
(IC_j)*(L. 1592/1933)	0.09***	0.10***	0.14***	0.11***
	(0.03)	(0.02)	(0.02)	(0.03)
(IC_j)*(L. 910/1969)	0.21***	0.21***	0.24***	0.20***
	(0.02)	(0.02)	(0.02)	(0.02)
(IC_j)*(L. 766/1973)	0.05***	0.06***	0.05***	0.05***
	(0.02)	(0.02)	(0.02)	(0.02)
(IC_j)*(L. 382/1980)	0.11**	0.04*	0.04***	0.06**
	(0.04)	(0.02)	(0.02)	(0.02)
(IC_j)*(L. 168/1989)	0.06*	0.04**	0.04***	0.04***
	(0.03)	(0.02)	(0.01)	(0.01)
(IC_j)*(L. 245-341/1990)	0.17***	0.16***	0.20***	0.22***
	(0.05)	(0.04)	(0.04)	(0.05)
(IC_j)*(L. 59/1997)	0.16*	0.17***	0.20***	0.20***
	(0.08)	(0.05)	(0.02)	(0.03)
R sq.	0.94	0.95	0.96	0.97
Observations	42813	42813	68410	35328

Notes: First stage of IV FE estimates reported in Table 11. 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 B-8: Sensitivity analysis: alternative empirical strategies

	[1] 2SLS FE	[2] 2SLS FE	[3] 2SLS FE	[4] 2SLS FE	[5] 2SLS FE	[6] 2SLS FE
no. faculties in 2^{nd} degree neighbour	–0.01*					
	(0.00)					
(IC_j)*(L. 2102/1923)	0.19***		0.27***	0.15***	0.26***	0.13***
	(0.05)		(0.04)	(0.04)	(0.04)	(0.04)
(IC_j)*(L. 1592/1933)	0.14***		0.16***	0.16***	0.17***	0.11***
	(0.02)		(0.02)	(0.02)	(0.02)	(0.03)
(IC_j)*(L. 910/1969)	0.22***	0.10***	0.23***	0.24***	0.19***	0.20***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
(IC_j)*(L. 766/1973)	0.06***	0.07***	0.08***	0.06***	0.01	0.05***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
(IC_j)*(L. 382/1980)	–0.03**		0.00	0.01	0.17***	0.06**
	(0.01)		(0.01)	(0.02)	(0.04)	(0.02)
(IC_j)*(L. 168/1989)	0.02*		0.01	0.03***	0.17***	0.04***
	(0.01)		(0.01)	(0.01)	(0.05)	(0.01)
(IC_j)*(L. 245/1990)	0.10***		0.11***	0.13***		0.22***
	(0.02)		(0.03)	(0.03)		(0.05)
(IC_j)*(L. 59/1997)	0.08***		0.10***	0.14***		0.20***
	(0.03)		(0.03)	(0.03)		(0.03)
R sq.	0.93	0.99	0.94	0.96	0.93	0.97
Observations	109817	2653	34904	35060	32210	35328

Notes:. First stage results of regressions reported in Table 12. All specifications include provincial pair fixed effects, region-by-year fixed effects, and the usual set of provincial controls. Details are Table 12. Standard errors clustered by province are reported in parentheses. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table B-9: Value of HE supply: First stage

	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE	OLS FE	2SLS FE
$IC_i * R_{21/30}$	0.2819***	0.2816***	0.2988***	0.2986***	0.2988***	0.2986***	0.0593	0.0610
	(0.0652)	(0.0650)	(0.0732)	(0.0719)	(0.0732)	(0.0719)	(0.0499)	(0.0503)
$IC_i * R_{31/36}$	0.1842***	0.1844***	0.2035***	0.1931***	0.2035***	0.1931***	–0.0971	–0.1001
	(0.0548)	(0.0546)	(0.0545)	(0.0537)	(0.0545)	(0.0537)	(0.0956)	(0.0924)
$IC_i * R_{61/70}$	0.1555***	0.2758***	0.2742***	0.2861***	0.2742***	0.2861***	–0.0116	–0.0036
	(0.0427)	(0.0921)	(0.0997)	(0.0976)	(0.0997)	(0.0976)	(0.0772)	(0.0757)
$IC_i * R_{71/80}$	0.0277							
	(0.0425)							
$IC_i * R_{81/90}$	0.2047							
	(0.1340)							
$IC_i * R_{91/00}$	–0.0096							
	(0.0393)							
F_{t-20}							1.1100***	1.1095***
							(0.1348)	(0.1337)
Observations	937	937	937	937	937	937	874	874

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