



Market forces in Italian academia today (and yesterday)

Chiara Zanardello¹

Received: 22 April 2022 / Accepted: 21 October 2022 / Published online: 14 November 2022
© The Author(s) 2022

Abstract

This paper investigates the operation of the academic market in Italy, mapping current scholars' location choices. I build a new dataset of current professors, associating each scholar with a composite indicator of their quality. The analysis includes the quality of the university and the features of the city where the institution is located. I estimate the strength of different factors: gravity (distance), agglomeration (scholars are attracted to higher quality universities), selection (better scholars travel longer distances), and sorting (the better the scholar, the more the quality of universities is weighted). I find that all of these factors have an effect, and do not vary according to scholars' gender. I find a greater expected utility for scholars in choosing private universities over public ones, through a consistent nesting procedure. Comparing these forces to historical trends in Italian academia, the sorting effect delineates a new momentum for the current academic market in Italy.

Keywords Human capital · Academic market · University · Scholar · Sorting · Italy

Mathematics Subject Classification 91-08 · 91-10 · 91-11

JEL Classification I23 · O15 · N34 · N33

I thank David de la Croix, Massimiliano Bratti, Daniele Checchi, and Fabrizio Iacone, who have contributed since the early draft of the project. I also thank all the scholars who replied to the survey. Simone Bertoli and Kenneth Train provided useful technical advice. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, under grant agreement No 883033 Did elite human capital trigger the rise of the West? Insights from a new database of European scholars. The first draft of this research (Zanardello, 2022) can be found as a working paper at <http://hdl.handle.net/2078.1/258573>.

✉ Chiara Zanardello
chiara.zanardello@uclouvain.be

¹ IRES/LIDAM, UCLouvain, Place Montesquieu 3, 1348 Louvain-la-Neuve, Belgium

Introduction

Knowledge and knowledge mobility have enormous social impact. The location choice of scholars, who are in the upper tail of the human capital distribution, has a measurable effect on knowledge formation, productivity and innovation. Bahar and Rapoport (2018) show that the location choice of high-skilled migrants draws the knowledge diffusion path, which has relevant effects on productivity. Location and birthplaces analyses can also disentangle economic competitiveness and complexity. When there is skill-complementarity, more diverse birthplaces result in more complex economic systems and higher economic growth (Alesina et al., 2016; Bahar et al., 2019; Docquier et al., 2020). The connection between human capital and modern economic growth is also confirmed in theoretical and empirical results. The essential role of education for adapting to changing contexts and for driving modernization was first claimed by Nelson and Phelps (1966). Lucas (1988) modeled the positive spillover effects of education, and more recently many empirical studies have presented clear evidence about the essential role of human capital and its institutions in improving current societies (Agasisti et al., 2019; Barra & Zotti, 2017, 2018; Barro, 1991, 2001; Beine et al., 2001; Cohen & Soto, 2007; Cottini et al., 2019; Hanushek & Woessmann, 2008).

Notwithstanding its fundamental role in socio-economic progress, investments in knowledge and educational institutions are not linear over time. There is a non-steady process of human capital accumulation, with periods of growth, decline, and recovery that may seed a new cycle of expansion (Artige et al., 2004). Italy is an interesting example of this fluctuating process: it dominated intellectual activities until the late Renaissance, but its contemporary cultural and educational system has weak incentives (Cecchi et al., 2021). Italian universities outperformed learning institutions in other countries from the 14th century to the first part of the 16th century (de la Croix et al., 2022), but only three universities are counted among the 200 top institutions in the 2022 QS World University Rankings¹.

I study the mobility of modern day Italian academics, and how they choose where to develop their career, to learn more about the modern university system of the peninsula. I take several factors into account: distance is increasing the cost of travelling, but it might be offset by the skills and knowledge acquired by the scholar and by the prestige of the university. I estimate professors' location choice as a function of distance and quality, given the location of universities. I map the current academic market with its professors' human capital and compare this to previous eras of Italian academia. By comparing how scholars are moving nowadays to how they moved in the past, I locate Italy in the fluctuating cycle, estimate the path of Italian academia, and map out future directions.

For my research I have built a new dataset of contemporary Italian professors in the economic field, capturing information about their origins and their individual quality. To collect information about the birthplaces of live persons I had to secure privacy authorizations, which could have hindered data collection. To overcome this missing data issue, I used a more accessible proxy: the location of professors' lowest level of education. Once

¹ Polytechnic of Milan—142nd, University of Bologna—166th and Sapienza University of Rome—171st. 2022 QS World University Rankings are at: <https://www.topuniversities.com/university-rankings/world-university-rankings/2022>

I had a value for the birth and/or education location, I used a Principal Component Analysis (PCA) to build a composite indicator of individual quality out of eight bibliometric indexes.

I use a Random Utility Model (RUM), and specifically, a multinomial logistic regression to compare scholars' utility of living in a region other than their birthplace (Bertoli & Moraga, 2013; Bertoli & Rapoport, 2015; Beine et al., 2016; Ortega & Peri, 2013). I limit the analysis to choices made within the academic world, given the impossibility to consider the choice faced by academics when they decide whether to become a professor or to follow other career paths. I use the approach developed by de la Croix et al. (2022), who study the European Academic Market from 1000 CE to 1800 CE, to compare past and present outcomes in academia. My main estimations rely on information about geographical distance, individual quality of current professors (*human capital* hereafter), and aggregate quality of Italian universities (*notability* hereafter). (1) *Agglomeration* investigates whether Italian scholars are attracted by universities with higher notability, (2) *positive sorting* tests whether scholars with higher human capital weigh the notability of universities higher than do professors with lower individual quality, through the interaction term between human capital and notability, and (3) *positive selection* questions whether scholars with higher human capital move further, utilising the interaction between human capital and distance. There is an extensive literature showing that better-educated individuals are the most mobile portion of the population, with their higher growth perspectives giving them stronger incentives to move (Beine et al., 2021; Docquier & Marfouk, 2006; Faggian & McCann, 2009; Handler, 2018; Schiller & Cordes, 2016; Zhao et al., 2021). Grogger and Hanson (2011) showed precisely that highly educated people are more likely to move (positive selection) and that these highly-specialised migrants choose destinations that compensate knowledge better (positive sorting).

My hypotheses assume the presence of strong complementarity in knowledge and skills (Easterly, 2001), which leads to positive assortative matching: working with better scholars (i.e., with higher human capital) would increase the marginal gain of each professor, and this increase is greater for better scholars than for academics with lower human capital (Kremer, 1993). High complementarity also explains why more notable universities, populated with better scholars, attract more and more high-quality professors which may initiates a virtuous cycle of human capital accumulation.

I estimate these effects for Italian academia, and find that the standard distance effect is negative and has a magnitude in line with migration literature (Beine et al., 2011). To study *agglomeration* I include attractive features of the city in which the university is located (size and wealth), in addition to notability. The latter is positive but not significant, signalling that the relevance Italian universities' quality can be improved by public policy to increase its attractive power for contemporaneous scholars. *Agglomeration* is instead driven by the disposable household income in the city (i.e., city wealth). However, the estimator for the size of the city is negative, implying a dispersion effect—although with a lower magnitude than *agglomeration*. This finding is crucial for understanding mobility patterns and policy directions: it is essential to attract high-skilled people to create a dynamic context and generate positive spillovers for society, which may lead the country into the virtuous part of the cycle (Grogger & Hanson, 2015; Kerr et al., 2016, 2017; Stephan & Levin, 2001). I also find evidence of positive selection and positive sorting. Indeed, *positive selection* (interaction between distance and individual quality) is a solid result, which confirms that the higher the individual quality, the stronger the incentives for the scholar to travel to progressively better destinations. *Positive sorting* (interaction between human capital and notability) has a weaker significance level than selection. The weakness of sorting is

due to the structure of Italian higher education, which may be still influenced by the traditional seniority-based system (Capano, 2008; MacLeod & Urquiola, 2021; Reborá & Turri, 2008). Reforms to increase the autonomy of universities, like the decentralization of the recruitment process of university professors,² may have led to the greater importance of local excellence for mobility decisions. However, such reforms are too recent to be in the current project, and so sorting only reaches a low level of significance. The seniority-based system may explain why Italian universities lost their leading position: there is evidence of a highly significant positive sorting only until 1526, which fades towards 1800. The sorting effect only regains power in the sample of present-day scholars, but the significance of current sorting is not as strong as at the birth of Italian universities. This is probably due to the very recent academic reforms. Either way, these results are key to understanding Italy's current position in the cycle, where it seems to be new momentum for Italian universities. In addition to the main regressions, I test for gender differences and find no significant outcomes. Men and women have similar patterns of mobility in Italy, but women represent only 30% of the sample. I do find important differences between public and private universities. A variant of the standard logit model shows a greater expected utility for scholars in choosing private universities over public ones. This bolsters the argument in favour of a more autonomous, excellence-driven academic apparatus.

My analysis contributes to the migration and knowledge-based mobility literature. To the best of my knowledge, much of this literature deals with more general samples of high-educated/high-skilled people (Beine et al., 2011; Docquier & Marfouk, 2006; Grogger & Hanson, 2011; Handler, 2018; Kerr et al., 2016, 2017). Only a few articles investigate the mobility of academics or scientists. Stephan and Levin (2001) find evidence of the extra vitality brought by foreign scientists (foreign-born and foreign-educated) to the U.S. in the fields of Science and Engineering (S &E). Grogger and Hanson (2015) study the mobility of foreign-born students in S &E after earning an American Ph.D. degree, claiming positive spillover effects for destination countries. The migration of German-affiliated researchers is addressed in Zhao et al. (2021), who find a net outflow of researchers from Germany. The current research keeps the focus on the academia and aims to integrate the knowledge-based mobility literature about the Italian university system. To the best of my knowledge, published papers on Italian scholars have studied the role of individual quality on selection processes (Checchi & Verzillo, 2014; Checchi et al., 2014) or its link with the competition and incentives generated within the Italian scientific sector (Checchi et al., 2021). There have also been some case studies connecting mobility and human capital of professors [see Abramo et al. (2022) for Italy, Ejeremo et al. (2020) for Sweden, and Aksnes et al. (2013) for Norway]. These researches focus on the effects of professors' mobility on their performance. This project investigates the same relationship, but the other way round: it considers human capital as a possible driver for researchers' mobility. Within the Italian university system, mainly student mobility has been analysed (Agasisti & Bianco, 2007; Bratti & Verzillo, 2019; Triventi & Trivellato, 2008), and no previous works have investigated the drivers of scholars' location choice in Italy.

² DPR n. 390/1998, law n. 210/1998, and law n. 240/2010, among the others.

Data sampling

Institutional context

Italy is home to the oldest University in Europe³ and has a long tradition of literates and scholars such as Giovanni Boccaccio, Leonardo da Vinci and Galileo Galilei, who belong in the upper tail of the human capital distribution. The Italian academic system has interesting peculiarities, which are worth mentioning before the empirical analysis. Italy's education system was centralized for a long time (Cottini et al., 2019), making it subject to the whims of the governing body. This increased the importance of hierarchy within academia, based on informal relationships between the most important chaired scholars and government ministers (Capano, 2008; Rebora & Turri, 2008). In the 20th century, this centralization of the system was intended to reduce the inequalities in the Italian education system (Cottini et al., 2019; Triventi & Trivellato, 2008) and there were some positive outcomes. Social mobility improved (Barone & Guetto, 2016) and performance among geographical areas converged (Baldissera & Cornali, 2020), but academia remains seniority-based (Rebora & Turri, 2008), not only in Italy but throughout Europe (MacLeod & Urquiola, 2021). In 1946, to improve the functionality of the system, the universities' autonomy principle (art. 33 paragraph 6) was defined in the Italian Constitution. This precept aimed to underline local excellence (Checchi & Verzillo, 2014), giving each university the autonomy to hire eligible professors. However, this Constitutional principle entered into force only at the end of the 90s, due to the lack of technical standards. The actual implementation of the reforms⁴ fragmented the Italian academic market, and it retained some elements of the seniority-based apparatus (Bertola & Sestito, 2011; Bini & Chiandotto, 2003; Cottini et al., 2019; Rebora & Turri, 2008). Among the other modifications, it is important to note that Berlinguer's decree (DPR n. 390/1998) shifted recruitment from a national to a local process.⁵ Some could argue that the decentralization of the recruitment process may have had a negative impact [i.e., more opportunistic behaviour and nepotism (Perotti, 2008)] on the scientific productivity of selected professors. Nevertheless, Battistin et al. (2014) show no significant changes in the quality and meritocracy of the university system after the decentralization in 1998. In 2010,⁶ the selection procedure was modified again and became a two-stage process. Nowadays, a scholar has to pass a national open competitive exam to be eligible, and then must win a local contest to be hired by a university (Checchi & Verzillo, 2014; Rossi, 2016).

In a system where seniority was the main driver of an academic career, quality and individual ability may be irrelevant. However, Checchi et al. (2021) found evidence of the opposite. They showed how the most productive scholars are those who responded best to an increase in the level of competition within the university sector, even in the presence of weak incentives.

³ University of Bologna, founded as a university in 1088.

⁴ Among the others, the most important ones are laws n. 382/1980, n. 168/1989, n. 210/1998 and the DPR n. 390/1998.

⁵ For a detailed explanation of the recruitment process in Italy see the following web page of the Ministry of Education, Universities and Research (only in Italian): <https://www.miur.gov.it/reclutamento-nelle-universita>.

⁶ Law n. 240/2010.

The literature about mobility in Italian academia is thin and mostly focuses on student mobility (Agasisti & Bianco, 2007; Bratti & Verzillo, 2019); I have not found any literature on drivers of professors' location choices. Insight into scholars' mobility within the country, and a characterization of the forces that attract them to an institution, can inform public policy.

Professors and universities

This research is based on a new dataset. The data collection started with RePEc's⁷ ranking of the "Top 25% Institutions and Economists in Italy". I decided to focus on scholars in the economic field to exploit the high quantity and quality of information provided by RePEc website. It uses the EDIRC database (Economics Departments, Institutes, and Research Centers in the World), which includes universities, public agencies, central banks, independent research centres, and associations [for more details see sect. 2.3 in Zimmermann (2013)]. Each institution gains from every author's affiliation RePEc collates, implying an advantage for more populous entities (sect. 6 in Zimmermann, 2013; Seiler & Wohlrabe 2011).

For the present work, only the universities in Table 10 (Appendix) will be taken into account. In Table 10 (Appendix) there are 16 public universities, one polytechnic (UNIVPM), and four privately founded universities (BOCCONI, CATT, FUB and LUISS). I consider as *privately funded* a university not fully funded by the State: BOCCONI and LUISS may be defined as fully private, while CATT and FUB as hybrid, receiving funds from both public and private institutions. I include all of them in the *privately funded* group given their different funding system with respect to *fully public* universities.

Each institution includes a list of members (registered in the RePEc Author Service) and I include these observations in the dataset.⁸ The people registered on the server have different roles inside academia. In this study, I only include professors—full, associate, adjunct and assistant—and research fellows (also postdoctoral).⁹ I include a few emeritus professors who are still teaching. Only scholars who are active in teaching are included in the sample: I call this a "*teaching disclaimer*" and it captures emeritus professors and academics taking part in visiting programs or national/international collaborations. Hence, a visiting professor is only included in the sample if she explicitly mentions her teaching activity at the host university. Scholars "on leave" were not considered part of the sample, given the absence of the *teaching disclaimer*. This rule excluded research centres like CEPR, IZA, CESifo, given the honorific nature of their appointments. Table 11 (Appendix) presents the precise taxonomy for the scholars included in the dataset, with quantities and percentages.

Once a scholar is identified, they are associated with their university. This process required a careful investigation for each academic. The Curriculum Vitae (CV) was the main source, but where it was out of date or incomplete I used LinkedIn¹⁰ and personal web pages (institutional and/or private). I used the most updated affiliation at the moment

⁷ Research Papers in Economics (RePEc) is a project collecting bibliographic data about papers in economics and similar fields, aiming to spread and enhance relative researches.

⁸ The ranking is updated month by month; hence the names collected (and the status granted to them) can change with respect to the period of data collection, which is approximately December 2020–September 2021.

⁹ After determining the status of each member, I exclude doctoral students from the dataset.

¹⁰ Professional social network: linkedin.com

of consultation.¹¹ Affiliations to telematic universities were not taken into account and research centres were excluded. For those universities with multiple locations, I counted the main location, assuming that the majority of the scholars teaching in one location are also teaching in the other(s). This can generate some bias when locations are far away from each other as in the case of Catholic University, with four locations, in Milan (main building), Brescia, Piacenza, and Rome. I discuss the robustness check for this in the Appendix 7.

Some scholars are associated with more than one university, in Italy or abroad. Multiple affiliations comprise 7.06% of the sample, with a maximum of four affiliations. In the past, academics linked with multiple institutions were associated with high-quality scores (de la Croix et al., 2022), whereas nowadays it is more common to encounter multiple-affiliated scholars with low bibliometric indicators. Usually, these academics are younger and have a postdoc position in a university while teaching in another institution. Empirically, each affiliation of the same scholar is treated as if it was chosen by different individuals, leading to their overestimation with respect to unique-affiliated scholars (see section “[Robustness checks](#)”). In the following part of the paper, the former will be called repeated movers (RM), and the latter single movers (SM).

Treating multiple affiliations in this way, the initial sample counts 1440 observations. A cleaning process removed from the sample scholars who are no longer members of the Italian academy, Ph.D. students, non-teaching emeritus and visiting professors, and those who are on leave.¹² The cleaning process reduced the sample to 1077 names. This procedure identified 76 universities, of which 39 are foreign universities and 37 are Italian. From this set, universities with fewer than 20 scholars have been excluded, given their minor relevance for academics’ choice and to have balanced choices. In the Appendix 9, I also show the main results considering a threshold of 5 scholars per university. The different threshold implies less balanced choices but a more comprehensive dataset (Fig. 1). The resulting list is the set of choices each professor faces when maximising their location decision. With the threshold at 20 scholars, this choice set has 17 universities, all of which are Italian. The number of scholars in the database decreased to 936 observations, the percentage of multiple affiliations is now only 3.10% and the maximum number of associations decreased to three. From here onwards, this is the subset for analysis.¹³ Table 1 summarises the differences between the original dataset and the subset obtained after dropping universities with fewer than 20 scholars.

Data on locations

In my analysis, I study the distance a scholar is willing to travel to a given university to develop her career. I treat distance as an increasing cost for the individual. The further she is from her point of origin, the greater the distance and the higher the cost (Schwartz, 1976), also in terms of family attachment. I collect the birthplace for each observation, and treat this as an observable proxy of scholars’ usual life context. Using other locations, like residential locations, would involve some endogeneity issues, which are avoided by using

¹¹ Consultation period: December 2020–September 2021.

¹² One of the main difficulties was understanding the meaning of the various roles and titles indicated by each scholar. The final dataset was built to the best of available knowledge, however, minor errors may still be present.

¹³ This dataset is available upon request, and for privacy reasons the data will be made anonymous.

birthplaces (Barbieri et al., 2011). Other variables are not observable—where academics’ families live is non-observable, as is the location of their partner’s employment.

CVs and personal webpages were the main sources for affiliations, given that neither LinkedIn nor RePEc provide birthplace information.¹⁴ Only about 30% of the sample indicated their place of birth somewhere in their public profile. Although information about living persons is abundant and often easy to access, bureaucratic and privacy authorizations, which are essential to protect personal information, slow the data collection process. Instead I sent direct emails requesting this information, increasing by about 55% the number of birthplaces collected.¹⁵ This gives me a known birthplace for 87.07% of the academics (815 observations).

I included in the dataset the location of the institution where each scholar obtained her lowest, publicly-stated degree of education. I consider this another proxy for birthplace, given that the two are likely to coincide or be reasonably close. To deeply investigate this, I study a sub-sample of 789 scholars in my database, for whom I know both their birthplace and their place of education. I consider a scholar to be born and educated in the same cultural environment when the place of birth and education are not further than 60 kilometers.¹⁶ Table 2 presents the percentage of scholars who studied, respectively, in the same place where they were born, not further than 30 kilometers and not further than 60 kilometers. More than 60% of this sample was born and studied in the same socio-cultural context. This confirms that the place of education can be used as a proxy for the birthplace.¹⁷ This measure increased the dataset to 904 observations, reaching coverage of 96.58%. For the majority of the sample, the lowest level of education is the bachelor’s degree, but some academics mentioned also the high school. Only for a few observations, the lowest education level available was the master’s degree, while for five scholars only information about the Ph.D. is known. Given their small number (only 5 out of 936 observations), and given the location of their Ph.Ds: four of them received their doctorate from the same university (or a close one) that they teach at, and only one obtained their title abroad, an ad-hoc robustness check was not necessary. For most observations I found educational information in CVs or LinkedIn profiles, and if I could not find it online I requested it with a direct email.¹⁸ However, education information is missing for 32 observations (3.42% of the sample) and they will be excluded. I implement two different regressions: birthplaces and locations of the lowest level of education analysis (see section “[Main results](#)”).

I match decimal coordinates to location data,¹⁹ giving me a dataset with i observations associated with a geo-localized birth and/or education site, and k geo-localized universities.

Data on quality

For quality indicators I collect *aggregate* quality scores (notability) and *individual* bibliometric indexes (human capital). The former are at the university level, while the latter are associated with each scholar.

¹⁴ My thanks to RePEc administrators for their prompt answers.

¹⁵ I express my heartfelt thanks to those who answered in so a interested manner.

¹⁶ 60 kilometers seems a reasonable distance, feasible for commuting every day.

¹⁷ The correlation between distance from birthplaces and distance from educational places is at 60.36%.

¹⁸ Usually in the case of emeritus professors, or persons not in the Italian academia anymore. Thus, this procedure helped also to determine the status of these exceptions.

¹⁹ The websites used are: latitudelongitude.org, tuttitalia.it for Italy, while latlong.net for foreign cities.

Table 1 Comparison between datasets

	Original	Subset
Tot. observations	1440	936
N. Universities	76	17
Obs. after cleaning process	1077	936
Obs. with known birthplaces	936 (86.91%)	815 (87.07%)
Obs. with known education	1044 (96.94%)	904 (96.58%)
Obs. with not-known education	33	32
Multiple affiliations	7.06%	3.10%
Max n. affiliations	4	3

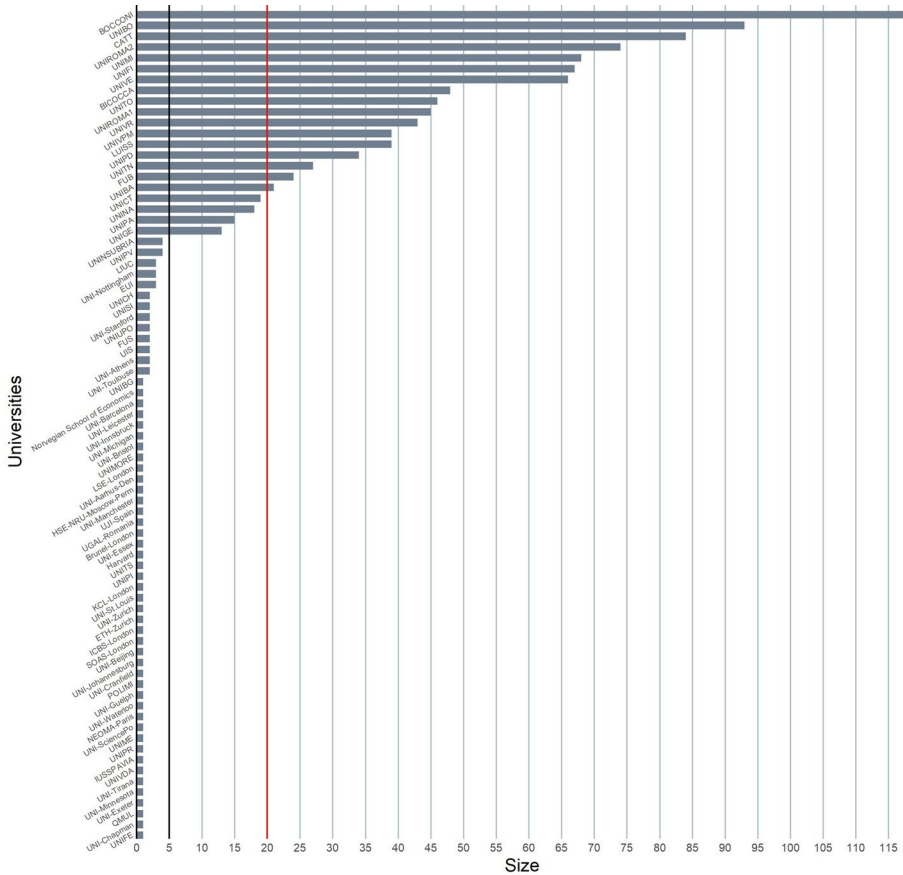


Fig. 1 Histogram of universities’ department size. The black line shows the threshold at 5 scholars and the red line shows the threshold at 20 scholars. The universities on the left are those excluded from the analysis at the respective threshold

Table 2 Descriptive statistics: scholars with both place of birth and education

Distance between place of birth and education	Percentage	N. obs out of 789
0 km	41.44	327
30 km	50.32	397
60 km	64.13	506

Notability indicators may suffer from endogeneity, because university scores are related to the quality and quantity of scholars. I address this by using past indicators. The average age for Italian academics is 48 years (Morana, 2020) and careers usually begin at around 30 years, so I look for quality indicators from 20 years earlier. The RePEc archives provide aggregate quality scores for top institutions, organised by country, going back as far as 2007. Prior to that, only simple/ordinal rankings are available, and there is no institutional score. I elected to consider scores from around 10 years ago, as of December 2010.²⁰ Scores for top universities were collected from country rankings. The scores in these rankings are weighted averages of the credit brought by each affiliated scholar: the highest portion (0.5) of affiliation is given to the scholar's main university and the remainder is a weighted average of the other appointments [for the specific formula see sect. 6 in Zimmermann (2013)]. This can generate some biases, for example decreasing the relevance of the main affiliation as more associations are added, as pointed out by Seiler and Wohlrabe (2011).

It was possible to assign a quality score to all 17 universities in the sample. RePEc uses reversed indexes in which lower scores indicate higher quality; I convert them to have a direct relation between indexes and quality. The notability ($\ln Q$) linked to each university ($k \in K$) can be visualized in Figs. 2 and 3.

To compute *human capital* there are many individual bibliometric indicators to choose from. RePEc has the top authors per country ranking (i.e. “Top 25% Institutions and Economists in Italy”). These human capital scores are the harmonic mean of various rankings based on different factors [sect. 5 in Zimmermann (2013)] and more than 800 scholars are ranked. I use the December 2020 ranking (see below for missing data).²¹

In the literature, academic quality is measured by indicators provided by *Web of Science* (*WoS*—with its three subject specific ISI citation databases; Yang & Meho 2006). The *WoS* social science indicator goes back to 1956.²² For a long time it has been one of the few multidisciplinary databases to assign authors' scores based on citations from an original set of sources (Jacso, 2005; Neuhaus & Daniel, 2008). The main issue with Web of Science measures is the relative coverage: only a fraction of sources are considered, although those that are considered (i.e., journal literature) are significant (Norris & Oppenheim, 2007). However, for economics and social science, this literature is not the main way that knowledge is disseminated (Neuhaus & Daniel, 2008).

²⁰ The research started in December 2020.

The ranking is available here: <https://ideas.repec.org/top/old/1012/>.

²¹ Given that the ranking is updated every month, the current online score could present some differences.

²² “Coverage in Web of Science goes back to 1945 for Science Citation Index, 1956 for Social Sciences Citation Index, and 1975 for Arts & Humanities Citation Index.” (Yang & Meho, 2006)

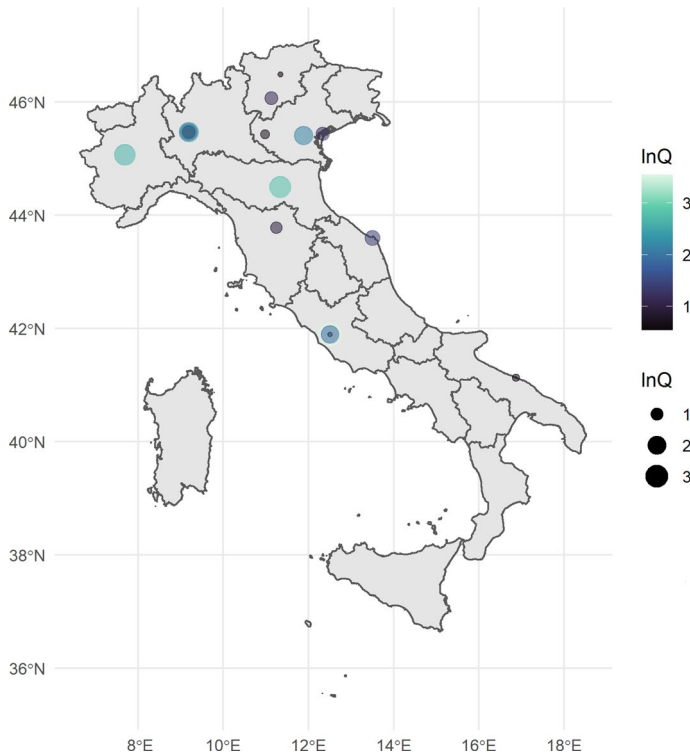


Fig. 2 Bubble Plot on Italy map. Showing notability indexes associated to each university: the higher the $\ln Q$, the bigger and the lighter the colour of the bubble. Each point represents the location of $k \in K$ institutions. Note: BICOCCA, BOCCONI and CATT are overwritten by UNIMI, which has the highest notability in Milan. UNIROMA1 and LUISS are overwritten by UNIROMA2, which has the highest notability in Rome

Quality-evaluation possibilities are now augmented with the automated databases *Scopus*, from Elsevier, and Google Scholar. The former covers a wider range of sources than *WoS*: it starts with an Elsevier database and it goes back to 1996 for social science²³ (Norris & Oppenheim, 2007; Jacso, 2005; Yang & Meho, 2006). *Google Scholar* is a free Google database that uses a wide range of sources, but does not identify clearly what those sources are. This gives it low reliability, which is added to weak, imprecise performance, as pointed out by Neuhaus and Daniel (2008). However, because it is free and has some of the widest coverage among bibliographic indicators, Google scholar still has value as a measure of quality (Neuhaus & Daniel, 2008).

I add to the comparison the *WorldCat identities* index. This database has measures for works (*Worldcat Works*) and library holdings (*Worldcat Library*) for each scholar (and organization) found in WorldCat.org and OCLC sources (OCLC Research, WorldCat identities).²⁴

²³ Scopus goes back at maximum to 1966. (Yang & Meho, 2006)

²⁴ <https://www.oclc.org/research/areas/data-science/identities.html>.

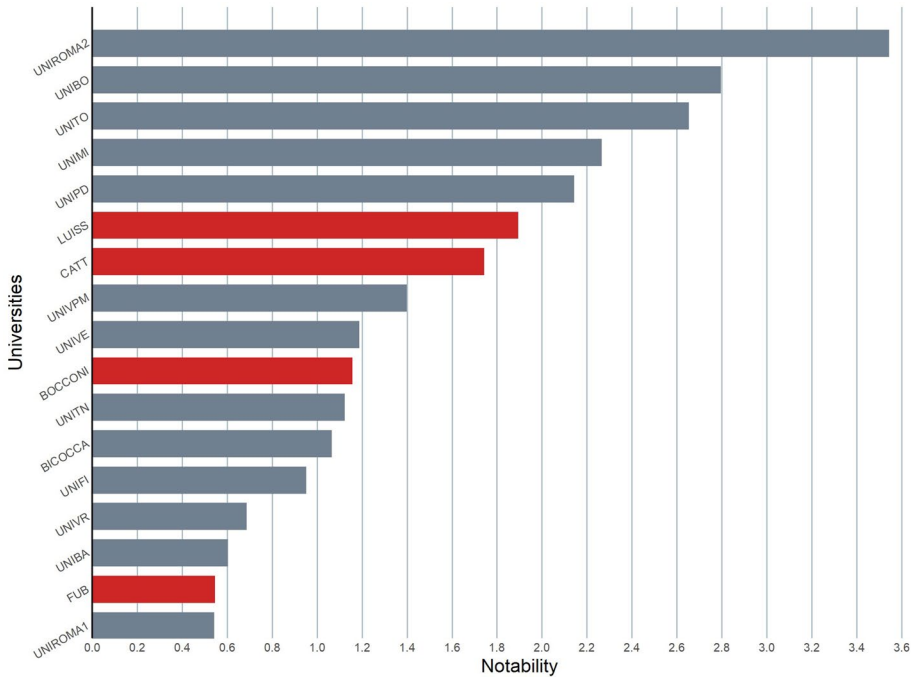


Fig. 3 Histogram of universities’ notability. Grey bars define public institutions, red bars private institutions. (Color figure online)

Fig. 4 Correlation Matrix Plot showing the correlations between the eight different bibliometric indicators included in the analysis

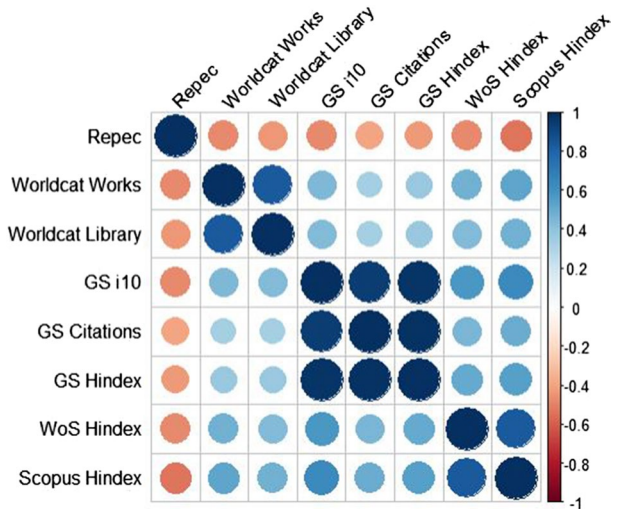


Table 3 Principal components table

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
St.dv.	2.2051	1.1768	0.9033	0.7604	0.4096	0.3861	0.1879	8e–02
Pr.Var.	0.6078	0.1731	0.1020	0.0723	0.0210	0.0186	0.0044	8e–04
Cum.Pr.	0.6078	0.7809	0.8829	0.9552	0.9762	0.9948	0.9992	1.0000

Showing the standard deviation (St.dv.), the proportion of variance (Pr.Var.) and the cumulative proportion (Cum.Pr.) for each principal component

Because no single indicator is perfect, I create a composite indicator of: RePEc score, Worldcat works and library holdings,²⁵ Google Scholar citations, H-index and i10-index,²⁶ WoS H-index,²⁷ and Scopus H-index.²⁸ To understand the information added by each indicator I use a *Principal Component Analysis* (PCA) to reduce the number of variables, without losing too much accuracy and information. Once the correlation between the variables is computed (Fig. 4) and their standardization is completed, the PCA compresses most of the information among the first principal components, which are new uncorrelated variables. For this research, I take the first component into consideration, because its standard deviation is greater than one and the cumulative information explained is sufficiently high (60.78% of the total—Table 3). Hence, considering the first component, the analysis gains simplicity while losing only a little portion of its accuracy. The following equation shows the factor loadings. The constant is the minimum of the first component and normalizes it to avoid negative human capital indexes. I use this linear combination of weights to represent the new individual quality index:

$$\begin{aligned}
 q_i = & + 4.47 - 0.30 \ln(\text{RePEc score}) + 0.32 \ln(\text{Worldcat works}) \\
 & + 0.30 \ln(\text{Worldcat library holdings}) + 0.37 \ln(\text{Google Scholar citations}) \\
 & + 0.39 \ln(\text{Google Scholar H - index}) + 0.41 \ln(\text{Google Scholar i10 - index}) \\
 & + 0.35 \ln(\text{WoS H - index}) + 0.37 \ln(\text{Scopus H - index})
 \end{aligned}$$

Methodology

Main hypotheses

In section “*Institutional context*”, I described some interesting features of Italian universities. One of the main aims of this paper is to understand how these features have changed over time. In order to achieve this objective I compare my results with (de la Croix et al., 2022). The authors tested the following hypotheses for the period between 1000 CE and

²⁵ <https://www.worldcat.org/identities/>.

²⁶ <https://scholar.google.com/>.

²⁷ https://app-webofknowledge-com.pros.lib.unimi.it/author/search?lang=en_US &SID=C45nUucTHD Od2z2VUPs.

²⁸ <https://www.scopus.com/freelookup/form/author.uri?zone=TopNavBar> &origin=NO%20ORIGIN%20DEFINED.

1800 CE, while I study them for contemporary times. The role of the location of higher education institutions (Agasisti et al., 2019; Audretsch, 1998; Barra & Zotti, 2017; Cottini et al., 2019; Drucker & Goldstein, 2007) is considered to be exogenous.

The current project assumes the presence of strong complementarity in knowledge and skills (Easterly, 2001, chapter 8). This property leads to positive assortative matching, where better scholars work together with other high-quality academics. The returns for working with better-skilled personalities are higher for better scholars than for their peers with lower human capital index (Kremer, 1993). Complementarity in knowledge and positive assortative matching may turn into possible virtuous cycles where notable universities, with better scholars, attract more and more high-quality human capital (Easterly, 2001; Kremer, 1993).

Hypothesis 1 Agglomeration: scholars are attracted by universities with higher notability.

I expect to find agglomeration (Grogger & Hanson, 2015; Kerr et al., 2016, 2017) although the distance covered by academics could appear shorter than in the past, with a lower magnitude of the coefficient. In Italy, the local appointment of professors may have increased the probability of finding local excellence (Checchi & Verzillo, 2014)—and the importance of networks and nepotism (Durante et al., 2011). With this hypothesis I test for agglomeration forces, such as notability of the university, and the attractiveness of the city in which the institution is located, measured by the size of the population (istat.it) and the local disposable income of private households (finanze.gov.it).

Hypothesis 2 Positive sorting: scholars with higher human capital weigh the notability of universities higher than scholars with lower human capital do.

I hypothesise that better scholars have better career prospects, and their expected gains are higher in high-quality environments (Docquier & Marfouk, 2006; Grogger & Hanson, 2015). Thus better professors would assign higher weight to the notability of the university.

Hypothesis 3 Positive selection: scholars with higher human capital move over greater distances than scholars with lower human capital.

The literature shows that better-educated people are more mobile (Beine et al., 2011, 2021; Grogger & Hanson, 2011; Schiller & Cordes, 2016), hence my hypothesis that better professors travel further.

The model

I use a Random Utility Model (RUM), a gravity model widely used in migration analysis (Beine et al., 2016; Bertoli & Moraga, 2013; Bertoli & Rapoport, 2015; Grogger & Hanson, 2011; Ortega & Peri, 2013). It determines the individual utility of living in a certain region and compares it to the expected utility from moving to alternative locations (Ramos, 2016).

I implement a standard multinomial logit model (Akcigit et al., 2016; Ortega & Peri, 2013), which is a specification of the RUM and requires perfect elasticity of demand in the academic market i.e., that there is a position available for every scholar. In Italian

academia, there is a two-step hiring procedure: scholars are filtered at the national level and then at the local level. The assumption of perfectly elastic demand implies that each professor who succeeds at the national level will succeed in finding a chair that she prefers at the local level. This is a reasonable assumption, because the reforms of the university system (in 1998 and in 2010) simplified bureaucratic processes and increased the opening of vacancies (Checchi & Verzillo, 2014; Rossi, 2016). However, in practice only professors with higher individual quality can freely choose the location of their career. To account for this I include the individual human capital score in the analysis. Keeping the perspective of partial equilibrium analyses, I introduce competition variables as demand-side factors, i.e., universities' notability, desirability of the city and individual human capital.

A multinomial logit model allows us to compute the probability that a university k , belonging to the set of choices K , is maximising a scholar i 's utility, with error terms independent and identically distributed (McFadden, 1974). Technical details are given in the Appendix 2.

Main results

In this section, I use the multinomial logit model described above to estimate the main regression of the research. First, I consider scholars for whom the place of birth is known (815 observations—87.07% of the sample). Second, as a robustness check, I use the site of their lowest level of education (904 observations—96.58% of the sample). I link each site with its geographic coordinates and each academic with a unique individual quality index, computed with a PCA (see section “Data on quality”). I discarded universities with fewer than 20 professors from the database, assuming that they have minor relevance in the total set of choices.²⁹ The university set counts 17 geo-localized institutions linked to their RePEc quality score (see section “Data on quality”). Because I work in logarithm terms, the estimation does not allow for zero indexes at aggregate or individual level. If a scholar does not have a positive score, I fill this gap with the lowest human capital index of the sample (794,82 for RePEc, 1 for all the other indicators). It is reasonable to assume that such a scholar does not publish as much as her peers with a positive score. However, it is possible that the sources used to compute bibliometric indicators do not accurately reflect her work, which is a known flaw in quality evaluations. I apply the same reasoning for universities with indexes at zero and link them with the lowest positive score of notability. Finally, I also take the logarithm of the measure of distance which raises the issue of zero distances, affecting scholars born in the same city where they teach. These academics bear the minimum cost of distance, which I assume to be the same as in de la Croix et al. (2022): 3,5 km, the walking distance from the Vatican city to the Colosseum, in the old city of Rome.

In the following part of the section, I describe the results of the main regression which considers scholars' locations of birth. I use the package called “mlogit”, written by Croissant (2020). I focus the evaluation on the sign and on the significance of the coefficients of distance, agglomeration, selection, and sorting effect. I control for unobserved characteristics of universities with fixed effects in each regression, except

²⁹ In the Appendix 9 show the main results considering a less stringent threshold of 5 scholars per university—the major difference is in the sorting effect, which falls just below the threshold of significance in the complete model of birthplaces analysis.

for when I introduce agglomeration effects. In this case, I include the variables which capture the observed characteristics of the city where the university is located (P_k and Y_k —see section “[The model](#)”) and represent the reputation of the institution (Q_k). Table 4 presents some descriptive statistics.

Table 5 shows the results of the multinomial logit estimations with known birth-places. The dataset counts 815 scholars (87.07% of the sample) who choose among 17 universities, resulting in 13855 possible dyadic matches.

The first column contains the basic gravity equation and highlights the negative sign of *distance* coefficients, $\ln d$. This means that the greater the distance between the birth-place and the location of the university, the higher the costs and the lower the probability of finding a dyadic match. Distance coefficients remain highly significant in every specification. The magnitude is consistent with the contemporary migration literature; for example, in “*Diasporas*” (by Beine et al., 2011) they also find distance coefficients of around 0.7 when migrants are not divided into low- and high-skilled categories. However, this coefficient is lower than in analyses of past periods (de la Croix et al., 2022).

I add a *selection* effect in the second column, defined by the interaction term between human capital and distance, $\ln q \ln d$. As expected the sign is positive, which means that scholars with higher human capital are less affected by distance than scholars with lower human capital. The high significance of the coefficient (at 1%) confirms the third hypothesis of *positive selection* in every specification of the model.

Column (3) shows the effect of *sorting*, through the interaction between individual human capital and university notability ($\ln q \ln Q$). The positive sign of the coefficient is evidence for *positive sorting*, as expected from the second hypothesis. Despite this, the significance of sorting appears weaker than selection. The sorting effect is non-significant when considered alone in column (3), but it becomes slightly significant (at 10%) in column (4) when I include selection. Sorting maintains the level of significance at 10% in the complete model [column (6)]. Finally, I compare log-likelihood (LL) values in order to compute a likelihood-ratio (LR) test: considering column (4) over column (1), the null hypothesis of no selection and no sorting is rejected at any conventional significance level (p value = 0.000).

To investigate *agglomeration*, I exclude university fixed effects from the regression [columns (5) and (6)], otherwise the effect of agglomeration variables cannot be identified (see section “[The model](#)”). Without fixed effects, I can study the relevance of the attractiveness of cities where universities are located. All three included variables are highly significant in column (5). The coefficient of the logarithm of population ($\ln P_k$) is negative, which precludes the presence of dispersion: the probability that a scholar chooses university k decreases as the city size increases. The coefficient of the logarithm of disposable income ($\ln Y_k$) is positive, which implies that the variable has a strong attractive force: the richer the city, the greater the likelihood a professor develops her career at that institution. The coefficient of the logarithm of university notability ($\ln Q$) is also significant at 1% and positive, which means that the better the university’s reputation, the higher the possibility that a scholar moves there. However, when I consider all the coefficients together [column (6)], notability loses its significance, while the other variables retain their signs, significance levels, and magnitudes—I confirm the second and third hypotheses. The first hypothesis about agglomeration also holds: although from $\ln P_k$ there is a tendency for dispersion (given its negative sign), it is more than compensated by the attractive force of city wealth ($\ln Y_k$). Nevertheless, these results show that agglomeration forces are driven by the income of the city and not by

Table 4 Descriptive statistics

Variables	Obs ^a	Mean	St.Dv.	Min	Max
Birthplace analysis					
In of distance	13855	5.454	1.310	1.253	9.198
In of human capital	13855	4.643	2.157	0.000	12.068
In of notability	13855	1.544	0.858	0.539	3.541
In of population	13855	12.990	1.713	7.858	14.879
In of income	13855	10.16	0.138	9.94	10.39
Lowest level of education analysis					
In of distance	15368	5.221	1.569	1.253	9.198
In of human capital	15368	4.494	2.203	0.000	12.068
In of notability	15368	1.544	0.858	0.539	3.541
In of population	15368	12.990	1.713	7.858	14.879
In of income	15368	10.16	0.138	9.94	10.39

^aObs counts the number of possible dyadic matches in each analysis

Table 5 Multinomial logit regressions: standard logit model—birthplaces analysis, threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
In d	-0.709*** (0.029)	-0.914*** (0.069)	-0.709*** (0.029)	-0.916*** (0.069)	- 0.710*** (0.028)	- 0.912*** (0.069)
Selection						
In $q \ln d$		0.043*** (0.013)		0.043*** (0.013)		0.042*** (0.013)
Sorting						
In $q \ln Q$			0.031 (0.019)	0.032* (0.019)		0.034* (0.020)
Agglomeration						
In P_k					- 0.122*** (0.029)	- 0.127*** (0.029)
In Y_k					2.076*** (0.355)	2.179*** (0.356)
In Q					0.171*** (0.044)	0.017 (0.101)
k FE	YES	YES	YES	YES	NO	NO
Obs	815	815	815	815	815	815
R^2	0.157	0.160	0.158	0.161		
LL	- 1867.206	- 1861.567	- 1865.936	- 1860.156	- 1916.167	- 1909.365

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

the university’s reputation ($\ln Q$). This result reveals room for public policies to improve the relevance of Italian universities’ quality in attracting human capital.

Robustness checks

In this section, I substitute the data on scholars' birthplaces with data on their lowest level of education, considered a proxy, which covers 96.58% of the sample. Table 12 (Appendix) presents the results of multinomial logit estimations when I study this proxy. Now the dataset counts 15368 dyadic matches, which associate 904 observations with 17 universities.

Only the *distance* and *agglomeration* coefficients remain significant. The sign of the former is still negative and each specification confirms the magnitude of about 0.7, although it slightly decreases compared to the birthplaces analysis. From the models without fixed effects [columns (5) and (6)], agglomeration variables ($\ln P_k$, $\ln Y_k$, $\ln Q$) confirm again the first hypothesis, with the same signs as in the birthplaces analysis.

The coefficients of *selection effect* ($\ln q \ln d$) are still positive, but not significant anymore. I find similar evidence for *sorting* ($\ln q \ln Q$), which has positive signs but not significant coefficients. These results prove that the second and the third hypotheses are confirmed only when I take into account the actual location of birth; indeed the LR test between (4) and (1) now fails to reject the null hypothesis of no effects (p value = 0.144). On the other hand, for the standard effect of distance and also for the agglomeration effect, the results remain in line with the birthplace investigation. In this case the analysis focuses on features of the universities (reputation/quality) and cities (population size and income level), aspects that do not vary compared to the previous analysis. The change of dataset affects distance and the individual level of quality, which appear in selection and sorting effects.

Given the results of both regressions, I consider the birthplaces analysis more relevant for the project. I use this as the benchmark model in the following part of the paper, where I develop further analyses.

For space constraints, I elaborate three additional investigations in the Appendix. Firstly, in Appendix 3, I correct the human capital index by scholars' age: younger professors with similar bibliometric indicators of senior ones should receive more credit in the computation of their human capital index. Once I introduce age into the analysis, Table 16 confirms almost all the benchmark results, but the sorting effect is just below the threshold of significance. These findings indicate that the human capital index employed in the main regression was already able to capture age specificities of Italian scholars. Secondly, I check the overestimation of repeat movers with different strategies in section D. Finally, in section E, I test for gender differences in the effects found in the benchmark model, but no significant discrepancies between male and female professors are found, although women are about one-third of my sample (30.24%).

Private/public universities

As mentioned in the description of the sample (section “[Professors and universities](#)”), four of the universities originally considered are private: Bocconi University, Catholic University, Free University of Bozen and LUISS University. Private universities have more hiring autonomy and discretion around remuneration (Trivellato et al., 2016; Agasisti & Ricca, 2016), making them more attractive to better scholars. To understand how private institutions influence the benchmark estimation, I run additional regressions in the Appendix (one estimation excludes all of them, the others exclude Bocconi and Catholic University one at the time, see Appendix 7).

In this section, I develop a nested logit model to investigate further. I divide the set of universities per status s : private and public. The nested logit still denies the correlation of error terms between the two groups (private and public), but there is the possibility of error terms dependency within a nest (McFadden, 1978; Train, 2003). With this method, it is possible to test whether one type of university implies systematically higher utility. Technical details and general results are found in the Appendix 6, here I focus on possible differences between private and public institutions.

As the nested logit model is consistent with random utility maximization (details in Appendix), I can define the expected gain each scholar obtains from choosing either a private or a public university. Given the lack of nest-specific variables, this utility is only given by the product $\lambda_s I_{is}$ (explained in Appendix 6), which varies for every scholar. Among the 815 professors considered, 127 have greater expected utility (EU) from teaching in public universities than in private ones, while 688 realize higher expected gains by affiliating to private institutions. I compare these two groups (Table 6) and the mean of the individual quality for those who prefer public universities is lower (4.02) than for those who prefer private institutions (4.76). *Sorting* effect is evident: better professors prefer more favourable environments. Private institutions, with more available resources, create better contexts to attract more relevant human capital.

Comparison between the present and the past

It is interesting to compare features of the contemporaneous academic market in Italy with those of the past. I run the same logistic regression as before but I use a sample of professors who worked in Italy from 1000 to 1800,³⁰ the whole period considered by de la Croix et al. (2022). Agglomeration variables are not fully comparable; I cannot test the first hypothesis with the updated dataset of de la Croix et al. (2022) because the authors consider the level of city democracy instead of the average disposable income of the households ($\ln Y_k$).

In Table 7, column (1) summarises the other findings using present professors (i.e., professors who currently work in Italy), while column (2) involves past scholars (i.e., professors who worked in Italy between 1000 and 1800).

In both cases, I confirm standard results for *distance* of gravity models: the greater the distance, the lower the probability a scholar chooses to travel that route. From Table 7, the difference in magnitude between these coefficients is evident but both are still in line with the literature, which provides greater magnitude for past periods than for current times. Furthermore, the distance in column (1) is the Euclidean distance, while in de la Croix et al. (2022) it is the cost distance. However, the Euclidean distance increases linearly with the cost distance, which limits the relevance of this computational difference. The magnitude of *selection* effects halves in current times with respect to the past, due to changes in individual quality measures—the human quality indexes are both the result of a *PCA* but they consider different bibliometric indicators.³¹ In the Appendix, Table 14 compares the

³⁰ Thank you to professor David de la Croix who provided this sample to the project.

³¹ Present indicators: RePEc score, Worldcat works and library holdings, Google Scholar citations, H-index and i10-index, WoS H-index, Scopus H-index.

Past indicators: number of characters of the longest Wikipedia page, number of Wikipedia pages in different languages, Worldcat works, library holdings, and publication languages.

Table 6 Descriptive statistics: groups of scholars preferring public or private universities

Variables	Obs	Mean	St.Dv.	Min	Max
EU Public > EU Private					
In of human capital	127	4.0245	1.8589	0.2715	7.8753
EU Private > EU Public					
In of human capital	688	4.757	2.1916	0.000	12.068

Table 7 Multinomial logit regressions: standard logit model, birthplaces analysis—comparison of results from the present and from the past without agglomeration variables

	(1) PRESENT ^a	(2) PAST ^b
Distance		
$\ln d$	− 0.916*** (0.069)	− 1.457*** (0.013)
Selection		
$\ln q \ln d$	0.043*** (0.013)	0.080*** (0.005)
Sorting		
$\ln q \ln Q$	0.032* (0.019)	0.012*** (0.002)
k FE	YES	YES
Obs	815	12,280
R^2	0.161	0.399
LL	− 1860.156	− 18,627.730

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

^aPresent professors, currently working in Italy

^bPast professors, working in Italy from 1000 to 1800

total effect of distance now and in the past for different levels of human capital: the effect is almost the same for top scholars in both columns, which confirms the comparability of the results. There is another important difference when I consider *sorting*. To compute the notability of the university, shown in the second column, de la Croix et al. (2022) aggregate the 5 highest human capital indexes associated with scholars active in that institution during the preceding 25 years (for technical details see de la Croix and Stelter 2021). In the first column, I link the notability index to the RePEc score of each university as of 10 years ago (see section “Data on quality”). Finally, the significance of sorting coefficients in Table 7 is weaker for current times than for the past, when the same effect had a high relevance.

The time horizon shown in the second column of Table 7 is too broad to freely compare it with the shorter time-span of column (1). Instead I exploit the division in periods developed by de la Croix et al. (2022) to seize more directly possible changes and fluctuations of the cycle that occurred in past centuries. In de la Croix et al. (2022) there are eight time segments with a different number of observations available for each period. I follow this two-by-two partition and I group together: the 2nd and 3rd period (1348–1449/1450–1526), the 4th and 5th (1527–1617/1618–1685), and the 6th and 7th (1686–1733/1734–1800). I exclude the first two periods from 1000 to 1347, because there are too few observations,

Table 8 Multinomial logit regressions: standard logit model, birthplaces analysis—comparison of results from the present and from the past without agglomeration variables

	(1) (1348–1526)	(2) (1527–1685)	(3) (1686–1800)	(4) (PRESENT)
Distance				
$\ln d$	- 1.360*** (0.022)	- 1.609*** (0.022)	- 1.598*** (0.032)	- 0.916*** (0.069)
Selection				
$\ln q \ln d$	0.090*** (0.010)	0.097*** (0.009)	0.056*** (0.013)	0.043*** (0.013)
Sorting				
$\ln q \ln Q$	0.014*** (0.005)	- 0.0002 (0.004)	- 0.010 (0.006)	0.032* (0.019)
k FE	YES	YES	YES	YES
Obs.	4472	4643	2314	815
R^2	0.346	0.416	0.435	0.161
LL	- 6040.509	- 6923.913	- 3473.547	- 1860.156

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

which leads to a negligible empirical relevance and less comparable results with respect to the other segments.

Table 8 shows the results. *Distance* is negative and highly significant in every specification. Its magnitude reflects the corresponding literature and historical period: it is higher in columns (2) and (3) than in column (1). This shift corresponds to the rise of national states and the increase in barriers and customs duties, leading to higher transportation costs. These burdens have decreased only in recent times with technological improvements and transport innovations. *Selection* effect is always present, with its positive sign and high relevance. Its magnitude drastically decreases in column (3) and lowers even more when the model involves current scholars.³² *Sorting* appears the most fluctuating effect across the time horizon, as expected. It is positive and highly significant in the first time range, when most of the major universities are already established (i.e., Bologna, Rome (Sapienza), Florence [Studium generale]) and are among the European top five institutions (de la Croix et al., 2022). These features allow me to position the 1348–1526 Italian academic market in the upward part of the aforementioned fluctuating cycle. However, sorting totally disappears in the second period I consider. Its sign is negative in both columns (2) and (3), but it is not significant in either column. These results might be due to the characteristics of the Italian academic world: its decline starts after the sixteenth century, a time of strict censorship of revolutionary concepts by the Catholic Church (Blasutto & de la Croix, 2021). Notable scholars were strongly attracted to the high quality of the first universities, but the sorting effect was diluted with the flow of time and with other universities entering the academic market. This decline in the sorting effect locates the 1527–1800 Italian university system in the downward portion of the cycle. Sorting regains its positive sign only when the model considers current scholars. In column (4), positive sorting is slightly significant, which may signal a new momentum for current Italian universities. With the local

³² When comparing past periods with the present, both human capital indexes are the results of a PCA but they involve different bibliometrics indicators.

recruitment of professors and the greater autonomy of each university, quality should gain attention and importance. However, the current analysis cannot detect these reforms with confidence; they are too recent and the influence of the previous seniority-based apparatus persists. This explains the weak sorting effect in the sample of contemporaneous Italian scholars. The same structural explanation applies to the low significance level of sorting in the past [column (2) and (3)]: the strong control of the powers in charge (e.g., Catholic Church) limited the relevance of university quality while favouring more denominational sorting, which relies on membership and networks rather than on meritocracy (MacLeod & Urquiola, 2021).

To further emphasize the relevance of *positive sorting* in the functioning of the academic market, I estimate scholars' choice probabilities and compute simulated outcomes with and without the sorting effect. In the Appendix 8 shows the estimated probabilities for three selected scholars, with different levels of human capital, but born in the same city. Constraining the sorting effect to zero drastically reduces the predicted probability to teach at the best universities. The effect is stronger for better scholars than for professors with lower human capital indexes: for the best scholar (A) the probability to choose the best university (UNIROMA2) halves without the sorting effect. This variation in the choice probabilities is visible for the first six institutions and lowers as I move down in the institutions' ranking. These findings (Table 25—Appendix) demonstrate the importance of positive sorting in fostering high-quality university contexts: the effect is much larger when better professors match with better universities.

To clarify the importance of positive sorting in enhancing quality in academia, I also estimated the total academic output with and without the sorting effect. Appendix 8 presents the production function I use. I assume the elasticity of substitution between professors' skills, ρ , to be finite. This assumption is crucial because it demonstrates complementarity between professors. As ρ falls, the gains from matching better scholars in the best institutions rise, improving the total output. Table 9 presents the results for two levels of ρ : the left part assumes low complementarity ($\rho = 3$), while the right part higher complementarity ($\rho = 2.6$). I estimate the total output by using both the benchmark [Table 5 column (6)] and the nested models [Table 20 column (6), in the Appendix].

When I compare the academic output with and without sorting in Table 9, the effect is already clear: when I do not constrain sorting to be zero, the gains are always higher, and this variation is greater with the nested logit model. To fully capture the importance of positive sorting, I compute academic output with different level of complementarity between scholars. When complementarity is higher, there is a significant increase in gains, with output that nearly doubles when the elasticity of substitution decreases by 0.4. This further underlines the relevance of positive sorting in creating a high-quality academic market. This also supports the background concept of this research: strong complementarity between scholars and positive assortative matching may lead to virtuous cycles by increasing the attraction of relevant human capital towards the best environments (Easterly, 2001; Kremer, 1993).

Conclusions

Using a new sample of contemporaneous scholars, this research confirms and discloses important features of the Italian academic market. Gravity highlights a recurrent effect widely explained in migration literature. Agglomeration forces of Italian universities

Table 9 Academic market output—role of sorting

	Low complementarity		High complementarity	
	$(\rho = 3)$		$(\rho = 2.6)$	
	Benchmark(6)	Nested(6)	Benchmark(6)	Nested(6)
With sorting:	26.27	26.58	42.93	43.62
Without sorting:	25.98	26.25	42.32	42.91
Δ sorting	0.28	0.33	0.60	0.71

All the values are divided by 1000

are driven by the average disposable income of the city where the institution is located and not by universities’ notability. This shows room for policy improvements: the quality of institutions is a strong factor for attracting relevant human capital and must be better exploited. Selection effect is also remarkably strong in the benchmark model, which implies that contemporaneous professors travel longer distances when they have greater human capital indexes. Sorting is weaker in this specification, but still significant and positive, which means that notability is more valuable for scholars with a higher individual quality index. Although it is less clear than the others, this last effect might direct the position of Italy in the human capital accumulation cycle. The difference in current and past sorting represents an important initial step for Italian academia: implementing reforms may enhance Italian universities’ notability. Policies to improve the quality of high-education institutions would stimulate excellence and in turn, would increase the attractiveness of Italian universities. This would trigger a virtuous circle for the whole economy—improving the sorting effect will feed the system with more resources, attract more remarkable scholars and increase the likelihood of innovations and economic enhancements, as demonstrated by the predicted academic market output. The United States, which has the top universities and research centers, has reaped the benefits of these positive spillovers. Since the early 1900s, sorting has been much stronger in America than it has been in Europe, where centralized systems favored equal growth of high-education institutions while simultaneously preventing the most promising ones from completely exploiting their potential (MacLeod & Urquiola, 2021). The recent reforms in the Italian system might be seen as a watershed moment: the positive achievements reached in terms of equality under a centralized system (Baldissera & Cornali, 2020; Barone & Guetto, 2016) can be bolstered by growing investments in excellence.

Future research can relax the assumption that demand in academia is totally elastic. This would necessitate the use of alternative gravity models, which do not impose the same stringent constraints as the multinomial logit. Gravity models that allow the consideration of both sides of the market can achieve a more complex general equilibrium analysis. Finally, the notability measure can be improved when using the conventional multinomial logit model. This could be accomplished by creating an index similar to de la Croix et al. (2022) to mitigate (if not eliminate) endogeneity issues with RePEc indicators.

Appendix 1: Additional tables

Table 10 University sample

University name	Location	Status
University of Bologna (UNIBO)	Bologna	Public
Catholic University of the Sacred Heart (CATT)	Milan	Private
University of Verona (UNIVR)	Verona	Public
University of Catania (UNICT)	Catania	Public
University of Milan (UNIMI)	Milan	Public
University of Rome—Tor Vergata (UNIROMA2)	Rome	Public
University of Florence (UNIFI)	Florence	Public
University of Venice (UNIVE)	Venice	Public
Polytechnic University of the Marches (UNIVPM)	Ancona	Public
Sapienza University of Rome (UNIROMA1)	Rome	Public
University of Turin (UNITO)	Turin	Public
University of Trento (UNITN)	Trento	Public
University of Naples Federico II (UNINA)	Naples	Public
University of Padua (UNIPD)	Padua	Public
Bocconi University (BOCCONI)	Milan	Private
University of Genoa (UNIGE)	Genoa	Public
University of Palermo (UNIPA)	Palermo	Public
Free University of Bozen (FUB)	Bolzano	Private
University of Bari (UNIBA)	Bari	Public
University of Milan-Bicocca (BICOCCA)	Milan	Public
LuiSS University in Rome (LUISS)	Rome	Private

Table 11 Taxonomy of scholars

Categories	Quantity	Percentage (%)
Full professors	420	39
Associate professors	303	28.13
Assistant professors	104	9.66
Adjunct professors	51	4.74
Research fellows	116	10.77
Post-doctoral fellows	30	2.79
Emeritus professors	8	0.74
Visiting professors	45	4.18
Total	1077	100

Table 12 Multinomial logit regressions: standard logit model—lowest level of education analysis—threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	- 0.679*** (0.022)	- 0.729*** (0.049)	- 0.679*** (0.022)	- 0.734*** (0.049)	- 0.680*** (0.021)	- 0.731*** (0.049)
Selection						
$\ln q \ln d$		0.011 (0.010)		0.012 (0.010)		0.011 (0.010)
Sorting						
$\ln q \ln Q$			0.029 (0.019)	0.030 (0.019)		0.029 (0.019)
Agglomeration						
$\ln P_k$					- 0.094*** (0.028)	- 0.096*** (0.029)
$\ln Y_k$					0.635* (0.374)	0.685* (0.375)
$\ln Q$					0.137*** (0.043)	0.012 (0.094)
k FE	YES	YES	YES	YES	NO	NO
Obs	904	904	904	904	904	904
R^2	0.224	0.224	0.225	0.225		
LL	- 1911.623	- 1910.982	- 1910.463	- 1909.688	- 1947.317	- 1945.631

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 13 Multinomial logit regressions: nested logit model—lowest level of education analysis—threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	- 0.689*** (0.114)	- 0.730*** (0.125)	- 0.700*** (0.114)	- 0.749*** (0.127)	- 0.835*** (0.084)	- 0.915*** (0.103)
Selection						
$\ln q \ln d$		0.008 (0.009)		0.010 (0.010)		0.015 (0.011)
Sorting						
$\ln q \ln Q$			0.031 (0.021)	0.032 (0.021)		0.038 (0.025)
Agglomeration						
$\ln P_k$					- 0.106*** (0.038)	- 0.110*** (0.039)
$\ln Y_k$					1.073** (0.491)	1.168** (0.504)
$\ln Q$					0.147*** (0.055)	- 0.018 (0.122)
$\lambda_{private}$	1.423*** (0.307)	1.413*** (0.306)	1.457*** (0.315)	1.450*** (0.315)	1.350*** (0.205)	1.332*** (0.213)
λ_{public}	0.981*** (0.168)	0.985*** (0.169)	0.996*** (0.169)	1.003*** (0.171)	1.247*** (0.124)	1.264*** (0.128)
k FE	YES	YES	YES	YES	NO	NO
Obs	904	904	904	904	904	904
R^2	0.226	0.226	0.227	0.227		
LL	- 1907.040	- 1906.659	- 1905.737	- 1905.246	- 1944.246	- 1942.259

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 14 Effect of distance in the present and in the past without agglomeration variables

	(1) PRESENT ^a	(2) PAST ^b
q_{min} *	- 0.916	- 1.457
q_{75} **	- 0.653	- 1.405
q_{max} ***	- 0.395	- 0.447
Obs	815	12,280

*Minimum q : -4.468 for present professors, 0 for past professors.

**75th quantile of q : 1.611 for present professors, 0.359 for past professors.

***Maximum q : 7.600 for present professors, 12.604 for past professors

^aPresent professors, currently working in Italy

^bPast professors, working in Italy from 1000 to 1800

Appendix 2: Standard logit model—technical details

A multinomial logit model allows us to compute the probability that a university k , belonging to the set of choices K , is maximising a scholar i 's utility (McFadden, 1974). The first step is to define the utility function for each i scholar. It is defined by a deterministic component $V_{ik} = \beta x_{ik}$, capturing average benefits and costs of each location choice, and by a random component ϵ_{ik} orthogonal to βx_{ik} , which describes unobservable factors that may influence the utility. The utility function can be written as follows:

$$U_{ik} = V_{ik} + \epsilon_{ik} = \beta x_{ik} + \epsilon_{ik} \tag{1}$$

The standard logit model relies on the assumption of independent individual choices, which requires ϵ_{ik} be independent and identically distributed (Extreme Value distributed—type I). Under this assumption, the main equation of the multinomial logit model defines the probability of choosing a university k , which depends on the specificities of that institution compared to the specificities of the remaining set of available choices:

$$p_{ik} \equiv Prob[U_{ik} = \max_{k' \in K} U_{ik'}] = \frac{\exp(\beta x_{ik})}{\sum_{k' \in K} \exp(\beta x_{ik'})} \tag{2}$$

Another important assumption of the logit model is the Independence of Irrelevant Alternatives (IIA). With independent and identically distributed error terms, the IIA assumption implies that the choice between two specific alternatives should depend only on their own features, without any influence from a third feature (McFadden, 1974). This means that the choice between two universities depends only on the two institutions considered and not on other alternatives. In Sections “Robustness checks” and “Private/public universities” in the main text, I relax this assumption.

In the next step, I explicit the deterministic component, which captures the difference between average benefits and average costs of choosing $k \in K$. The benefits are an increasing function of the university’s notability Q_k (as defined in the previous Section), and of the attractiveness of the city. Hence, I include the variables P_k and Y_k , representing respectively cities’ total population (capturing the size of the city) and households’ disposable income³³ (capturing the wealth status) of the city in which university k is located. In addition, I include an interaction term ($q_i Q_k$) to capture the fact that better scholars (with a high individual quality index q_i) gain more from a welcoming environment (i.e., a university with high Q_k). Therefore, the benefits equation is:

$$B_{ik} = a_0 + a_1 Q_k + a_2 P_k + a_3 Y_k + a_4 q_i Q_k \tag{3}$$

where $\forall a \in \{a_0, a_1, a_2, a_3, a_4\}$ greater than zero.

The costs are mainly influenced by distance. The greater the distance from the birth-place (or education-place) the higher is the burden of travel, and hence the higher the costs. However, the better an academic (i.e., the higher the q_i), the more she has to gain in certain university environment Q_k , implying a cost reduction by the interaction term $q_i Q_k$. In addition, as shown by the literature (Beine et al., 2021; Grogger & Hanson, 2011; Schiller & Cordes, 2016), a better scholar should be willing to move longer distances. Human capital

³³ Italy (2018, average disposable income IRPEF, city-level): <https://www.finanze.gov.it/it/>.

negatively affects distance, so I include the interaction term between distance and individual quality ($d_{ik}q_i$), which may decrease the costs. Therefore, the costs equation is:

$$C_{ik} = b_0 - b_4q_iQ_i + b_5d_{ik} - b_6d_{ik}q_i \quad (4)$$

where $\forall b \in \{b_0, b_4, b_5, b_6\}$ greater than zero.

Having defined the equations for the benefits and the costs, I specify the net benefit for each dyadic match (i.e., the association of scholar i with university k) and explicit the deterministic component of the utility function. This is done by subtracting (4) to (3), with the addition of a fixed effect γ_k . The subscript k suggests that fixed effects refer to time-invariant, non-measurable universities' characteristics which may influence their ability to attract human capital.³⁴ These fixed effects almost perfectly identify the agglomeration variables (Q_k, P_k, Y_k) included in the model, given that both are destination-specific and time-invariant. For this reason, I exclude fixed effects when agglomeration forces are considered (more on this later). The final expression is:

$$\beta x_{ik} \equiv V_{ik} \equiv B_{ik} - C_{ik} = \beta_0 + \beta_1Q_k + \beta_2P_k + \beta_3Y_k + \beta_4q_iQ_k + \beta_5d_{ik} + \beta_6d_{ik}q_i + \gamma_k \quad (5)$$

where β is a vector, whose parameters are common to each scholar. Specifically, the constant β_0 in eq. (5) is the difference between the two constants in (3) and (4). To define the *agglomeration effect*, I look at $\beta_j \equiv a_j$ with $j = \{1, 2, 3\}$, which represent notability of universities and attractiveness of cities. A positive sign indicates the presence of agglomeration and a negative sign evidences dispersion. I measure the *sorting effect* with $\beta_4 \equiv a_4 + b_4$. A positive coefficient means that the higher the individual quality, the smaller the cost (or the higher the gain) to travel to better universities. The coefficient $\beta_5 \equiv -b_5$ captures the expected effect of distance, considered as a cost, as previously mentioned. $\beta_6 \equiv b_6$ underlines the *selection effect*: when there is a positive sign, better scholars move further.

Equation (5) includes only destination-specific regressors. Human capital (q_i) is always interacted with university-specific variables (i.e., notability and distance), because it influences all dyadic matches in a symmetric manner.

Appendix 3: Age analysis

In the benchmark model, I do not consider the age of scholars to compute the human capital index. I can define this as the *current* human capital indicator, computed as today. However, younger professors with the same level of human capital index as senior ones ought to receive more credit. I include professors' age in the dataset, to compute *age-expected* human capital index. CVs are the main source for this data: most of the scholars disclose their year of birth. However, for 28.85% of the dataset this information is missing and for these cases, I look to the final year of their Ph.D. and assume that scholars at the end of their doctorate are 30 years old. This leaves 29 observations for which I do not have any age reference, neither the year of birth nor the last year of their Ph.D. I exclude them from the analysis, bringing the number of dyadic matches to 13362, which corresponds to 786 academics. Figure 5 presents the age distribution of scholars where the average age is 50 years and a few months.

³⁴ Fixed effects are used to overcome omitted variables biases, they control for unobserved variables which do not change over time. If there is a change over time, they can be inefficient, with large standard errors. (Williams, 2018)

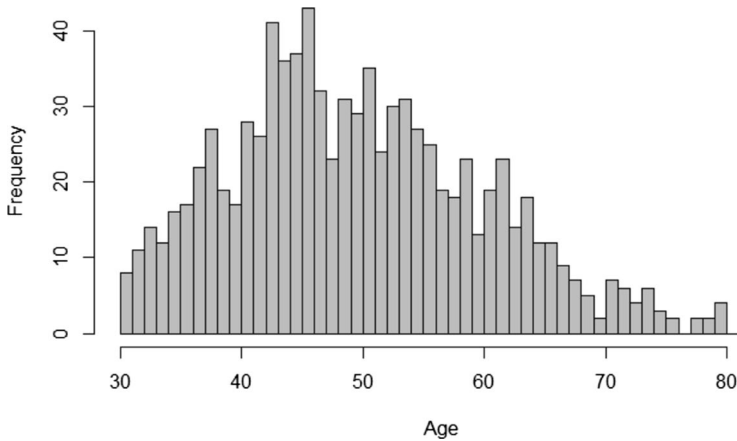


Fig. 5 Histogram showing the age distribution of scholars: mean 50.0751, median 49, variance 108.505 (std. dev. 10.4166), skewness 0.4387 (right-skewed distribution), kurtosis - 0.2510 (platykurtic distribution), min 30, max 83

I expect age to be a crucial factor in explaining individual human capital: as age increases, the probability of publishing highly-cited researches increases, augmenting in turn the individual quality index. I run an ad-hoc regression, which confirms this expectation at the 1% significance level (Table 15). I estimate the *age-expected* human capital index at 40 years old for each scholar using the coefficients of the age regression. I re-run the benchmark model with this new indicator and show the results in Table 16.

Distance ($\ln d$), *agglomeration* ($\ln P_k$, $\ln Y_k$, $\ln Q$) and *selection* ($\ln q \ln d$) coefficients have the same sign as in the benchmark model and are still highly significant—agglomeration is driven by city wealth also in the age-adjusted model. *Sorting* ($\ln q \ln Q$) coefficients are positive but not significant, they fall slightly below the threshold of significance. The second hypothesis does not hold anymore: I cannot claim that scholars with a higher human capital index weigh more the quality of universities than those with lower individual quality. The low significance level of both the sorting effect and notability in the agglomeration effect stresses the importance of public policies to improve the quality of the university environment in Italy.

In conclusion, when I introduce age, the model confirms almost all the benchmark results, but the sorting effect is just below the threshold of significance. These findings indicate that the *current* human capital index employed in the main regression was already able to capture age specificities of Italian scholars.

Table 15 Ordinary Least Square: age regression on human capital

	$\ln q$
AGE	0.485*** (0.015)
AGE ²	− 0.004*** (0.0001)
Constant	− 13.460*** (0.382)
Observations	13,632
R ²	0.119
Adjusted R ²	0.119
Residual Std. Error	2.026 (df = 13359)
F Statistic	900.433*** (df = 2; 13359)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 16 Multinomial logit regressions: standard logit model, birthplaces age-adjusted analysis—threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	− 0.703*** (0.029)	− 1.002*** (0.080)	− 0.703*** (0.029)	− 1.001*** (0.080)	− 0.704*** (0.029)	− 1.004*** (0.080)
Selection						
$\ln q \ln d$		0.058*** (0.014)		0.058*** (0.014)		0.058*** (0.014)
Sorting						
$\ln q \ln Q$			0.030 (0.021)	0.031 (0.021)		0.032 (0.022)
Agglomeration						
$\ln P_k$					− 0.117*** (0.029)	− 0.123*** (0.029)
$\ln Y_k$					2.093* (0.360)	2.219* (0.361)
$\ln Q$					0.174*** (0.045)	0.012 (0.118)
k FE	YES	YES	YES	YES	NO	NO
Obs	786	786	786	786	786	786
R ²	0.154	0.158	0.155	0.159		
LL	− 1805.994	− 1797.129	− 1804.951	− 1796.062	− 1853.601	− 1843.621

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix 4: Multiple affiliations analyses

After I eliminate universities with fewer than 20 scholars, multiple affiliations count for 3.10% of the sample. Given how they enter the dataset (see section “Professors and universities”—main text), the choices of these scholars³⁵ are overweighted with respect to those of single movers.³⁶ However, there should not be an over-representation of better scholars against worse ones as in de la Croix et al. (2022), because my dataset associates repeat movers with a low-quality score, with an average human capital of 4.12 against 4.66 of single movers.³⁷ To confirm that repeat movers do not influence benchmark results, I exclude them from the dataset in columns (3) and (4) of Table 18 and then I associate them randomly with one of their affiliations in columns (5) and (6). The main variation with respect to the benchmark model is in the notability coefficient, which becomes negative but remains not significant in both modifications. Significance levels remain almost as in the benchmark model, but in both variations [column (3),(4) and (6)] the sorting effect is now significant at 5% and gains some relevance.

So far, I assumed independent career choices, as required by the IIA assumption in standard logistic models. This is violated when individuals choose more than one alternative at the same time, which is the case when scholars are affiliated to more than one university. I develop a mixed logit model to test for correlated preferences. This version of logistic regression allows to consider the presence of heterogeneous agents. It is similar to the standard model but more flexible: the coefficients are scholar-specific and the utility function includes an additional term which permits correlated choices and the relaxation of the IIA assumption (Ye et al., 2020; Train, 2009). Hence, the mixed logit model modifies scholar’s utility function as follows:

$$U_{ik} = \beta_i x_{ik} + \eta_i x_{ik} + \epsilon_{ik} \tag{6}$$

Where the first term is the general deterministic component, which represents the utility of scholar i who chooses university k . The other two terms capture the unobservable part of the function: η_i is an individual deviation and ϵ_{ik} is a random term as before. I assume these two error terms to be normally distributed.

The mixed logit model, with the η term violating the IIA assumption, requires the integration of the conditional probability by using the joint probability density function, $f(\beta_i|\theta)$; where θ summarises the first and the second moment of the distribution. The vector of β coefficients is assumed to be independent and normally distributed and it is of length N . To obtain the unconditional probability of professor i choosing university k , the following formula applies:

$$\begin{aligned} P_{ik} &= E(P_{ik}|\beta_i) = \int_{\beta} (P_{ik}|\beta_i) f(\beta_i|\theta) d\beta \\ &= \int_{\beta_1} \int_{\beta_2} \dots \int_{\beta_N} (P_{ik}|\beta_i) f(\beta_i|\theta) d\beta_1 d\beta_2 \dots d\beta_N \end{aligned} \tag{7}$$

³⁵ NB: I call these scholars ‘repeated movers’, which means that they are associated with more than one university.

³⁶ NB: by ‘single mover’ I mean a scholar associated with only one university.

³⁷ To compute the mean of $\ln q$, I consider 26 observations for repeat movers and 789 observations for single movers.

Table 17 Multinomial logit regressions: mixed logit model—birthplaces analysis—threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	- 1.141*** (0.086)	- 1.509*** (0.178)	- 1.139*** (0.086)	- 1.501*** (0.178)	- 1.075*** (0.073)	- 1.343*** (0.153)
Selection						
$\ln q \ln d$		0.078** (0.031)		0.077** (0.031)		0.059** (0.028)
Sorting						
$\ln q \ln Q$			0.028 (0.022)	0.028 (0.022)		0.033 (0.026)
Agglomeration						
$\ln P_k$					- 0.152*** (0.044)	- 0.159*** (0.043)
$\ln Y_k$					3.228*** (0.504)	3.305*** (0.516)
$\ln Q$					0.160*** (0.054)	- 0.0004 (0.129)
k FE	YES	YES	YES	YES	NO	NO
Obs	815	815	815	815	815	815
R^2	0.175	0.177	0.175	0.177		
LL	- 1828.727	- 1824.090	- 1827.756	- 1823.157	- 1877.079	- 1871.267

The mixed logit involves the six s.d. associated to each coefficient only the s.d. linked to $\ln d$ and $\ln Q$ are significantly different from zero

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

where

$$(P_{ik} | \beta_i) = \frac{\exp(\beta_i x_{ik})}{\sum_{k' \in K} \exp(\beta_i x_{ik'})} \tag{8}$$

is the conditional probability.

Simulations are used to draw the parameters from the β distribution: the unconditional probability is the average of the conditional probabilities computed for each scholar (Train, 2009; Ye et al., 2020; León & Miguel, 2013). Table 17 presents all the specifications of the mixed logit regression, columns (4) and (6) are those reported in Table 18.

With respect to the benchmark model, the magnitude varies for every coefficient. Signs and significance levels of *distance* coefficients confirm the gravity literature and magnitudes increase by almost 0.4. *Agglomeration* is still driven by the city wealth which compensate the negative sign of the population coefficient, and notability does not play any role in attracting scholars, as before. *Selection* effect confirms the third hypothesis with its significance levels at 5%. *Sorting* always appears weaker than the other effects: it is positive, but not significant. The LR test between columns (7) and (1) rejects the benchmark version, but I should notice that there are six additional parameters estimated with the mixed logit (not reported in Table 18). Nevertheless, the mixed logit is weaker than the benchmark since it involves simulations and not a maximization. Moreover, the assumption

Table 18 Repeat Movers’ robustness checks and Mixed logit—birthplace analysis threshold at 20

	Benchmark		Removing RM		RM linked to 1.uni.		Mixed Logit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance								
$\ln d$	-0.916*** (0.069)	-0.912*** (0.069)	-0.965*** (0.071)	-0.960*** (0.071)	-0.955*** (0.071)	-0.950*** (0.071)	-1.501*** (0.178)	-1.343*** (0.153)
Agglomeration								
$\ln P_k$		-0.127*** (0.029)		-0.129*** (0.029)		-0.135*** (0.029)		-0.159*** (0.043)
$\ln Y_k$		2.179*** (0.356)		2.094*** (0.371)		2.143*** (0.365)		3.305*** (0.516)
$\ln Q$		0.017 (0.101)		-0.049 (0.105)		-0.013 (0.103)		-0.0004 (0.129)
Selection								
$\ln q \ln d$	0.043*** (0.013)	0.042*** (0.013)	0.051*** (0.013)	0.050*** (0.013)	0.050*** (0.013)	0.049*** (0.013)	0.077** (0.031)	0.059** (0.028)
Sorting								
$\ln q \ln Q$	0.032* (0.019)	0.034* (0.020)	0.042** (0.020)	0.044** (0.020)	0.038* (0.019)	0.041** (0.020)	0.028 (0.022)	0.033 (0.026)
k FE	YES	NO	YES	NO	YES	NO	YES	NO
Obs	815	815	763	763	789	789	815	815
R^2	0.161		0.169		0.166		0.177	
LL	-1860.156	-1909.365	-1724.414	-1774.466	-1786.364	-1838.901	-1823.157	-1871.267

The mixed logit involves the six s.d. associated to each coefficient, only the s.d. linked to $\ln d$ and $\ln Q$ is significantly different from zero

RM repeat movers

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

on parameters’ distribution is essential to obtain these results, which may change when considering another assumption. The original model remains the benchmark.

Appendix 5: Gender analysis

In the dataset, women are about one-third of the sample (30.24%). In this Section, I test for gender differences in the effects found in the benchmark model.

I estimate the same regression with the addition of an interaction term: each categorical variable interacts with a gender dummy (1 if male, 0 if female). As shown in Table 19, none of the interaction terms with the gender dummy ($\ln dxM$, $\ln P_kxM$, $\ln Y_kxM$, $\ln QxM$, $\ln q \ln dxM$, $\ln q \ln QxM$) are significant, revealing no evidence of gender differences. When I include the selection effect in the models [columns (2), (4), and (6)], negative *distance* coefficients show lower magnitudes when only women are considered. When I do not include the selection effect in the model [columns (1), (3), and (5)], distance coefficients have greater magnitudes when I consider only women, with positive men’s distance coefficients. Agglomeration coefficients of population and city wealth are lowered if I analyse male scholars, notability is reinforced but it is not significant. Selection and sorting are

almost always reinforced when the male portion of the sample is involved, apart in column (2) for selection, and in column (6) for sorting. Most of the coefficients for women are in line with the benchmark model, which confirms that there are no significant gender differences.

Table 19 Multinomial logit regressions: standard logit model, birthplaces analysis—gender differences threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	- 0.709*** (0.029)	- 0.916*** (0.069)	- 0.709*** (0.029)	- 0.917*** (0.069)	- 0.710*** (0.028)	- 0.911*** (0.069)
$\ln dxM$	0.036 (0.042)	- 0.046 (0.103)	0.037 (0.043)	- 0.043 (0.103)	0.036 (0.043)	- 0.033 (0.100)
Selection						
$\ln q \ln d$		0.043*** (0.013)		0.044*** (0.013)		0.042*** (0.013)
$\ln q \ln dxM$		- 0.016 (0.019)		0.016 (0.019)		0.014 (0.019)
Sorting						
$\ln q \ln Q$			0.031 (0.019)	0.032* (0.019)		0.034* (0.020)
$\ln q \ln QxM$			0.002 (0.010)	0.002 (0.010)		- 0.017 (0.023)
Agglomeration						
$\ln P_k$					- 0.123*** (0.029)	- 0.127*** (0.029)
$\ln P_k xM$					0.028 (0.031)	0.026 (0.031)
$\ln Y_k$					2.082*** (0.355)	2.181*** (0.356)
$\ln Y_k xM$					- 0.568 (0.387)	- 0.525 (0.387)
$\ln Q$					0.171*** (0.044)	0.017 (0.101)
$\ln QxM$					0.026 (0.053)	0.108 (0.122)
k FE	YES	YES	YES	YES	NO	NO
Obs.	815	815	815	815	815	815
R^2	0.158	0.160	0.158	0.161		
LL	- 1866.834	- 1860.827	- 1865.548	- 1859.418	- 1914.320	- 1907.022

“ xM ” represents the relative effect when only Male scholars are considered

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix 6: Nested logit model—technical details

The nested logit model enables more appropriate comparison of the two groups of universities and permits the relaxation of the IIA assumption. This method still denies correlation of error terms between the two sectors (private and public), but now there is the possibility of error terms dependency within a nest (McFadden, 1978; Train, 2003). Hence, the IIA assumption holds within a nest, where the unobserved portions of utility still have the same mean, while the assumption does not hold between nests, where means of the error terms can now differ (Train, 2003; Heiss, 2002).

In general, the nested logit model permits grouping of alternatives in nests with similar characteristics, with a certain degree of correlation λ_s . I divide the university set K per status s : private and public. Thus, the utility function of scholar i is decomposed into two parts, plus a random component $\epsilon_{ik'}$. The first portion H_{is} depends only on the nest s , and the other portion $M_{ik'}$ depends on a specific alternative k' within nest s (Train, 2003). The new utility function is defined as follows:

$$U_{ik'} = H_{is} + M_{ik'} + \epsilon_{ik'} \tag{9}$$

for $k' \in K_s$.

Starting from this decomposed utility function, it is possible to describe the probability of choosing $k \in K_s$ as the product between two probabilities: the conditional probability of choosing k given that the choice of nest K_s has been made (i.e., a standard logit model between alternatives in nest K_s) and the marginal probability of choosing universities in nest K_s (i.e., a standard logit model between nests). The probability of the final choice k for scholar i is the product of two standard logit models:

$$p_{ik} = P_{ik|K_s} P_{iK_s} \tag{10}$$

where

$$P_{ik|K_s} = \frac{\exp(M_{ik'}/\lambda_s)}{\sum_{k' \in K_s} \exp(M_{ik'}/\lambda_s)} \tag{11}$$

and where

$$P_{iK_s} = \frac{\exp(H_{is} + \lambda_s I_{is})}{\sum_{l=1}^S \exp(H_{il} + \lambda_l I_{il})} \tag{12}$$

with

$$I_{is} = \ln \sum_{k' \in K_s} \exp(M_{ik'}/\lambda_s) \tag{13}$$

The quantity I_{is} is called *inclusive value* or *log-sum term*. It is essential for connecting information in the upper model (marginal probability) with information in the lower model (conditional probability) and it is defined by the logarithm of the lower model denominator—Eq. (11) (Train, 2003). λ_s is called *log-sum coefficient* or *dissimilarity parameter* and it reveals informations about the degree of error terms correlation: the higher the λ_s , the higher the independence (or the lower the correlation) of the unobserved portion of utility. The standard multinomial logit model requires λ_s be equal to 1, which implies complete independent error terms (i.e., zero correlation of error terms) (Train, 2003; Heiss,

2002). λ_s captures the substitutability of alternatives: if there is more substitution within than between nests, then λ_s is lower than one, while if substitution is greater between rather than within nests, then λ_s is greater than one (Train et al., 1987).

Once I_{is} multiplies λ_s , their product $\lambda_s I_{is}$ represents the extra expected utility of scholar i from choosing the best university in nest K_s . This extra expected utility is added to H_{is} , which defines the expected utility of choosing whatever alternative is in the nest. H_{is} depends on nest-specific variables, which are not present in my analysis. Hence, only the product $\lambda_s I_{is}$ tells the difference in the expected utility of choosing a private or a public university: the higher the $\lambda_s I_{is}$, the higher the gain for the scholar (Train, 2003). For a nested logit model to be globally consistent with Random Utility Models, the density function must be non-negative; this condition is always met for dissimilarity parameters within the unit interval (Börsch-Supan, 1990; Kling & Herriges, 1995). When λ_s are larger than one (Train et al., 1987), the consistency condition may still hold locally, i.e. for some value of the explanatory variables (Börsch-Supan, 1990; Kling & Herriges, 1995).

Table 20 presents the results of the simultaneous nested logit model for the birthplace analysis (Table 13 in section A shows the nested logit model for the lowest level of education analysis).³⁸ All results still hold qualitatively when compared to the benchmark regression. One cannot compare magnitudes between the benchmark and the nested logit model directly, given the presence of the additional parameters λ_s , but it is possible to analyze meaningful ratios. With *selection* effect in column (6) of Table 20, when a scholar has a human capital index of 10 (i.e., top scholar), her distance costs decrease by more than 35% with respect to a scholar with a human capital indicator of 4 (i.e., average scholar). Once I compute this percentage using benchmark coefficients, I can claim that there is no relevant difference between the two models in terms of *selection*: the percentage of cost reduction is almost the same (−35.20% for the nested model,³⁹ −33.87% for the standard model⁴⁰). The *sorting* effect presents some difference, with more inequalities among scholars in the nested than in the standard specification. In column (6) of Table 20, when a scholar has a human capital index of 10, with the sorting effect her gains are 199%⁴¹ higher than the gains of a scholar with an individual quality indicator of 4. On the other hand, in the benchmark model (Table 5—main text), gains for better scholars are 133%⁴² higher than for scholars with lower human capital.⁴³ I also compare two opposite situations to compute the gain percentage variation: the gains of a better scholar (i.e., with a human

³⁸ A consistent nested logit model can be computed also sequentially, but this latter method is less efficient than the simultaneous approach currently employed (Heiss, 2002; Train, 2003).

³⁹ The cost for a scholar with a human capital index of 4 is: $-1.452 + 4 \cdot 0.069 = -1.176$ The cost for a scholar with a human capital index of 10 is: $-1.452 + 10 \cdot 0.069 = -0.762$ The cost reduction for a better scholar is: $(-0.762 + 1.176) / -1.176 = -0.3520 = -35.20\%$

⁴⁰ The cost for a scholar with a human capital index of 4 is: $-0.912 + 4 \cdot 0.042 = -0.744$ The cost for a scholar with a human capital index of 10 is: $-0.912 + 10 \cdot 0.042 = -0.492$ The cost reduction for a better scholar is: $(-0.492 + 0.744) / -0.744 = -0.3387 = -33.87\%$

⁴¹ The gain for a scholar with a human capital index of 4 is: $-0.055 + 4 \cdot 0.056 = 0.169$
The gain for a scholar with a human capital index of 10 is: $-0.055 + 10 \cdot 0.056 = 0.505$
The increase of gains for a better scholar is: $(0.505 - 0.169) / 0.169 = 1.988 = \mathbf{198.82\%}$

⁴² The gain for a scholar with a human capital index of 4 is: $0.017 + 4 \cdot 0.034 = 0.153$
The gain for a scholar with a human capital index of 10 is: $0.017 + 10 \cdot 0.034 = 0.357$
The gain increase for a better scholar is: $(0.357 - 0.153) / 0.153 = 1.333 = \mathbf{133.33\%}$

⁴³ Here, better scholars have a human capital index of 10, while scholars with a lower human capital index have an indicator of 4.

Table 20 Multinomial logit regressions: nested logit model—birthplaces analysis—threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	- 0.860*** (0.170)	- 1.137*** (0.238)	- 0.874*** (0.172)	- 1.170*** (0.245)	- 1.085*** (0.144)	- 1.452*** (0.218)
Selection						
$\ln q \ln d$		0.053*** (0.019)		0.055*** (0.019)		0.069*** (0.022)
Sorting						
$\ln q \ln Q$			0.041 (0.026)	0.046* (0.028)		0.056* (0.033)
Agglomeration						
$\ln P_k$					- 0.167*** (0.050)	- 0.180*** (0.053)
$\ln Y_k$					4.093*** (0.827)	4.413*** (0.893)
$\ln Q$					0.191*** (0.073)	- 0.055 (0.162)
$\lambda_{private}$	1.821*** (0.553)	1.871*** (0.552)	1.886*** (0.579)	1.968*** (0.590)	1.635*** (0.354)	1.730*** (0.378)
λ_{public}	1.185*** (0.240)	1.218*** (0.247)	1.204*** (0.243)	1.249*** (0.253)	1.582*** (0.210)	1.637*** (0.223)
k FE	YES	YES	YES	YES	NO	NO
Obs	815	815	815	815	815	815
R^2	0.159	0.161	0.159	0.162		
LL	- 1864.157	- 1858.387	- 1862.610	- 1856.624	- 1902.012	- 1894.613

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

capital index of 10) who teaches in a better university (i.e., with a notability index of 3) to the gains of a scholar with a lower individual quality (i.e., with a human capital index of 4) who teaches in a worse university (i.e., with a notability index of 1). In the nested logit model (Table 20), the gains for the better scholar are 796%⁴⁴ higher than those of her peer with a lower human capital indicator. When I compare these two opposite situations in the benchmark model, the gains for the better scholar who teaches in a better university are 600%⁴⁵ higher than for a scholar with a lower human capital index who teaches in a

⁴⁴ The gain for a scholar with a human capital index of 4 in a university with notability of 1 is: $-0.055 + 4 \cdot 0.056 = 0.169$

The gain for a scholar with a human capital index of 10 in a university with notability of 3 is: $-0.055 \cdot 3 + 10 \cdot 3 \cdot 0.056 = 1.515$

The gain increase for a better scholar in a better university is: $(1.515 - 0.169)/0.169 = 7.964 = 796.45\%$

⁴⁵ The gain for a scholar with a human capital index of 4 in a university with notability of 1 is: $0.017 + 4 \cdot 0.034 = 0.153$

The gain for a scholar with a human capital index of 10 in a university with notability of 3 is: $0.017 \cdot 3 + 10 \cdot 3 \cdot 0.034 = 1.071$

The gain increase for a better scholar in a better university is: $(1.071 - 0.153)/0.153 = 6 = 600\%$

university with lower notability. In general terms, the *sorting* effect is stronger in the nested than the standard logit model.

The nesting procedure seems to be justified: the null hypothesis of no nests is rejected through a log-likelihood ratio test (LL = -1909.4, p value = 0.000), and the correlation within nests is different from zero (Wald test = 21.636, p value = 0.000), but the null hypothesis of unique nest elasticity cannot be rejected (Wald test = 0.3037, p value = 0.5816; LL = -1894.8, p value = 0.507). This raises questions about the applicability of the grouping strategy I employ here, although dividing private and public universities appears reasonable. To clearly define the pertinence of this nested logit model, it is necessary to look at the additional parameters in the last two rows of the output, the λ_s . Firstly, all dissimilarity parameters exceed unity, which poses another question on the global consistency of this nested model with utility maximization. Indeed, Daly and Zachary (1978) and McFadden (1979) show that obtaining λ_s inside the unit interval is essential for the model to be globally consistent. Nevertheless, when this consistency is relaxed to hold only locally, the dissimilarity parameters can exceed one; as showed by Kling and Herriges (1995). Specifically, two conditions must be checked. (i) The non-negativity of the first-order partial derivatives of the choice probabilities is the first necessary condition, and it is described as follows:

$$\lambda_s \leq U_{1s}(v) \equiv \frac{1}{1 - Q_s(v)} \quad s = 1, S \quad (14)$$

with v_k defined as the utility delivered by each alternative and $v \equiv (v_1, \dots, v_K)$, and where $Q_s(v)$ is the upper model as in Eq. 12:⁴⁶

$$Q_s(v) = P_{iK_s} = \frac{\exp(\lambda_s I_{is})}{\sum_{l=1}^S \exp(\lambda_l I_{il})} \quad (15)$$

For this first necessary condition to hold, Q_s must be sufficiently large. (ii) The second condition questions the non-positivity constraint on the mixed second-order of the choice probabilities, as follows:

$$\lambda_s \leq U_{2s}(v) \equiv \frac{4}{3[1 - Q_s(v)] + \sqrt{[1 + 7Q_s(v)][1 - Q_s(v)]}} \quad (16)$$

To define these conditions, I compute Q_s from Eq. 14 and compare it with the λ_s , which are already in the output (Table 20). Kling and Herriges (1995) present different approaches to precisely test the consistency of nested models, I follow them and Table 21 shows my results.

One possible approach confronts $\hat{\lambda}_s$ with both $\hat{U}_{1s}(\bar{v})$ and $\hat{U}_{2s}(\bar{v})$, where \bar{v} denotes the mean of the indirect utility function. Already this approach seems to highlight the consistency of my nested model, by finding that $\hat{\lambda}_s \leq \hat{U}_{1s}(\bar{v})$ and $\hat{\lambda}_s \leq \hat{U}_{2s}(\bar{v})$. Notwithstanding, another approach investigates at which level of precision the estimated coefficients (the λ_s) are able to reject—or not—the local consistency. Hence, I develop a one-tailed test for each condition. For the first-order condition, I test the null hypothesis $H_{1O} : \lambda_s \leq U_{1s}(\bar{v})$ against the alternative $H_{1A} : \lambda_s > U_{1s}(\bar{v})$, and for the second-order condition I compare the

⁴⁶ The term H_{is} is not reported, because of the lack of nest-specific variables.

Table 21 Consistency tests of NLM with RUM

First-order conditions			
Nest	$\hat{\lambda}_s$	$\hat{U}_{1s}(\bar{v})$	<i>t</i> ratio
Private	1.7300	18.4201	– 44.1279
Public	1.6368	2.1878	– 2.4687
Second-order conditions			
Nest	$\hat{\lambda}_s$	$\hat{U}_{2s}(\bar{v})$	<i>t</i> ratio
Private	1.7300	3.8092	– 5.49714
Public	1.6368	1.3067	1.4798

null hypothesis $H_{1O} : \lambda_s \leq U_{2s}(\bar{v})$ to the alternative $H_{1A} : \lambda_s > U_{2s}(\bar{v})$. The last column of Table 21 reports the *t* ratios of each test statistic ($t_{1,2} \equiv [\hat{\lambda}_s - \hat{U}_{1,2s}(\bar{v})]/Std.Error$); negative coefficients immediately imply that the null hypothesis of local consistency cannot be rejected, which is almost always the case. Only the second-order consistency condition for public universities is rejected, as it was the case of the previous approach. Nevertheless, grouping private and public universities seems appropriate and the first-order consistency condition largely approves this nesting procedure. I consider these results consistent enough with utility maximisation models.

Appendix 7: Private/public universities—further insights

In this section, I develop other estimations to further investigate the role played by private universities.

Table 22 presents the results of the regressions I run when I exclude all four private universities from the sample. The total of dyadic matches is now 7774, with 598 observations and 13 universities.

Distance coefficients confirm previous findings, each of them is negative, highly significant and with a magnitude a little larger than the benchmark, but still in line with the literature. *Agglomeration* variables in column (5) confirm the first hypothesis by the positive sign and high significance level of the population size ($\ln P_k$) and university notability ($\ln Q$) coefficients. However, I cannot confirm the presence of agglomeration in the last column, where notability loses its significance. This result is driven by the decreasing in magnitude of the city wealth ($\ln Y_k$), which is not significant neither in the fifth nor in the sixth column—likely because the private universities excluded are located in rich cities: two of the four are in Milan, the city with the highest disposable income. *Selection* ($\ln q \ln d$) has a positive sign, its significance level is at 1% and its magnitude is similar to the benchmark model. *Sorting* coefficients are positive and highly significant as well, their magnitude almost double. Table 22 confirms *positive selection* and *positive sorting* and brings evidence for a reinforcement of the latter effect when only public universities are included in the model, but there is no *agglomeration*.

Following, I study two additional logistic regressions: one excludes only Catholic University while keeping Bocconi University, and another excludes Bocconi University while keeping Catholic University. With the former, I investigate the issue of secondary locations

Table 22 Multinomial logit regressions: standard logit model, birthplaces analysis—private universities excluded—threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	- 0.735*** (0.032)	- 0.945*** (0.079)	- 0.735*** (0.032)	- 0.943*** (0.079)	- 0.732*** (0.031)	- 0.931*** (0.077)
Selection						
$\ln q \ln d$		0.045*** (0.015)		0.044*** (0.015)		0.043*** (0.015)
Sorting						
$\ln q \ln Q$			0.070*** (0.022)	0.069*** (0.022)		0.072*** (0.022)
Agglomeration						
$\ln P_k$					- 0.082*** (0.031)	- 0.092*** (0.031)
$\ln Y_k$					0.216 (0.463)	0.352 (0.464)
$\ln Q$					0.209*** (0.046)	- 0.106 (0.107)
k FE	YES	YES	YES	YES	NO	NO
Obs.	598	598	598	598	598	598
R^2	0.208	0.211	0.211	0.214		
LL	- 1179.174	- 1174.645	- 1174.030	- 1169.584	- 1199.210	- 1189.387

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

to understand whether it alters previous results. With the latter, I examine the position of excellence recently reached by Bocconi in several rankings—it has lately become the best university in Italy for economics and related fields and this may be due to its well-known ability to attract high-ranked personalities.

Table 23 presents the results obtained by excluding Catholic University: almost all coefficients are significant. *Distance* and *agglomeration* remain as in the benchmark (Table 5 in the main text) for what concern signs and significance, the magnitude is similar as well, the income coefficient experiences the largest variation: a drop of 0.198 (from 2.179 to 1.981). *Positive selection* maintains its significance level of 1% in each specification, with a slight decrease in magnitude. *Positive sorting* is again weaker than selection but it gains significance in the third column when it is considered alone (with respect to zero significance level of the coefficient in the benchmark). This implies that there is no relevant bias due to the imprecise geographical coordinates associated with this university. Although I assume that every scholar teaches only in Milan, the benchmark model is not significantly influenced.

Excluding Bocconi from the set of choices allows me to find only significant coefficients which improves the solidity of the results (Table 24). *Distance* and *agglomeration* variables remain with the same significance as in the benchmark model. The magnitude of the notability coefficient decreases and the coefficient turns negative, while income's magnitude almost halves but still confirms agglomeration. I find *positive selection* as in

Table 23 Multinomial logit regressions: standard logit model, birthplaces analysis—Catholic University excluded

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	- 0.705*** (0.029)	- 0.891*** (0.071)	- 0.705*** (0.029)	- 0.893*** (0.070)	- 0.708*** (0.029)	- 0.890*** (0.070)
Selection						
$\ln q \ln d$		0.039*** (0.013)		0.039*** (0.013)		0.038*** (0.013)
Sorting						
$\ln q \ln Q$			0.034* (0.020)	0.035* (0.019)		0.036* (0.020)
Agglomeration						
$\ln P_k$					- 0.114*** (0.029)	- 0.119*** (0.029)
$\ln Y_k$					1.872*** (0.383)	1.981*** (0.384)
$\ln Q$					0.171*** (0.044)	0.004 (0.103)
k FE	YES	YES	YES	YES	NO	NO
Obs	741	741	741	741	741	741
R^2	0.169	0.171	0.170	0.172		
LL	- 1635.306	- 1630.858	- 1633.796	- 1629.271	- 1683.633	- 1677.802

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

the previous estimation without Catholic University. In this regression, there is strong evidence for *positive sorting* in each specification. This model reaches a high significance level, which means that private universities have different features not totally captured by the variables included in the benchmark: further investigations are necessary (i.e., applying different models, like the nested logit model—see section “[Private/public universities](#)” in the main text). Despite this, I can claim that all the expected features of the contemporaneous academic world described in section “[Main hypotheses](#)” of the main text still hold if I exclude Bocconi University.

Table 24 Multinomial logit regressions: standard logit model, birthplaces analysis—Bocconi University excluded, threshold at 20

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	- 0.708*** (0.029)	- 0.920*** (0.073)	- 0.709*** (0.029)	- 0.921*** (0.072)	- 0.710*** (0.029)	- 0.919*** (0.072)
Selection						
$\ln q \ln d$		0.045*** (0.014)		0.046*** (0.014)		0.045*** (0.014)
Sorting						
$\ln q \ln Q$			0.062*** (0.021)	0.063*** (0.021)		0.068*** (0.021)
Agglomeration						
$\ln P_k$					- 0.120*** (0.029)	- 0.128*** (0.029)
$\ln Y_k$					1.154*** (0.397)	1.276*** (0.399)
$\ln Q$					0.230*** (0.046)	- 0.070 (0.105)
k FE	YES	YES	YES	YES	NO	NO
Obs	714	714	714	714	714	714
R^2	0.174	0.176	0.176	0.179		
LL	- 1578.996	- 1573.514	- 1574.347	- 1568.781	- 1612.901	- 1602.348

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix 8: Role of sorting

To underline the relevance of *positive sorting* in the functioning of the academic market, I estimate scholars' choice probabilities and compute simulated outcomes with and without the sorting effect. I select three scholars (A, B, C) with different levels of human capital (low, medium, and high) but born in the same city (Florence). Table 25 displays the values of their predicted probabilities for each location in the choice set. The values on the left part of Table 25 refer to the complete model of the benchmark (Table 5 column (6), in the main text). The right portion of Table 25 includes the predicted probabilities when I set the sorting effect to zero in the same model. I list the institutions by notability to highlight the effect of sorting on the best universities.

Constraining the sorting effect to be zero drastically reduces the predicted probability to teach at the best universities. The effect is stronger for better scholars than for professors with lower human capital indexes: for the best scholar A the probability to choose the best university UNIROMA2 halves without the sorting effect. This variation in the choice probabilities is visible for the first six institutions and lowers as I move down in the institutions' ranking. These findings (Table 25) demonstrate the importance of positive sorting in fostering high-quality university contexts: the effect is much larger when better professors match with better universities.

Table 25 Predicted values of individual location choice probabilities—role of sorting

	Benchmark (6)			Benchmark (6)—NO sorting		
	A	B	C	A	B	C
Birthplace ln of HC	Florence 10.097 (%)	Florence 4.415 (%)	Florence 2.487 (%)	Florence 10.097 (%)	Florence 4.415 (%)	Florence 2.487 (%)
UNIROMA2	6.5	3.0	2.1	3.3	2.1	1.8
UNIBO	11.1	7.7	6.3	7.2	6.2	5.5
UNITO	4.1	2.1	1.5	2.8	1.7	1.4
UNIMI	7.3	4.3	3.3	5.7	3.7	3.0
UNIPD	6.3	4.0	3.2	5.1	3.5	3.0
LUISS	3.6	2.3	1.8	3.2	2.1	1.7
CATT	6.1	3.9	3.1	5.7	3.7	3.0
UNIVPM	3.8	2.8	2.4	4.0	2.8	2.3
UNIVE	5.6	4.2	3.5	6.3	4.3	3.6
BOCCONI	4.9	3.6	3.0	5.6	3.7	3.0
UNITN	3.4	2.4	2.0	3.9	2.5	2.0
BICOCCA	4.8	3.5	2.9	5.6	3.6	3.0
UNIFI	23.5	49.0	58.5	28.8	51.9	60.0
UNIVR	3.0	2.6	2.2	4.0	2.8	2.4
UNIBA	1.3	0.8	0.7	1.7	0.9	0.7
FUB	2.7	2.1	1.8	3.9	2.4	1.9
UNIROMA1	2.2	1.8	1.6	3.1	2.0	1.7

To clarify the importance of positive sorting in enhancing quality in academia, I also estimated the total academic output with and without the sorting effect. I interpret the academic output as a result of high-level teaching and high-impact research. Following de la Croix et al. (2022), I first aggregate the individual quality of every scholar i predicted to teach in a university k to compute institutions’ output (characterized as a CES production function), and I aggregate them to compute the total output of the model—with and without the sorting effect. I proceed as follows:

$$Y = \sum_k \left(\sum_i \hat{p}_{ik} q_i^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \tag{17}$$

where \hat{p}_{ik} is the predicted probability for a scholar i to choose university k , as shown in the three examples above. These values weigh scholar’s human capital q_i . The additional parameter ρ denotes the elasticity of substitution between individual quality of professors in producing institutions’ output. This parameter is crucial because, by assuming it to be finite, it demonstrates complementarity between professors. As ρ falls, the gains from matching better scholars in the best institutions rise, improving the total output. The results are in the main text, Table 9.

Appendix 9: Main regressions with the threshold at 5 scholars

I develop the two main regressions of the study with varied thresholds in order to further investigate the peculiarities of the current Italian academic market. Hence, rather than limiting myself to universities with more than 20 scholars, I now present the results considering institutions with more than 5 scholars. With this new threshold, I involve in the analysis also smaller universities—Genoa, Catania, Naples, and Palermo. On the one hand, this expands the scope of the research by allowing more universities in the southern part of Italy to participate (i.e., Catania, Naples, and Palermo). On the other hand, these small universities reduce the balance of the choice set, because they appear within the career possibilities of each scholar but they are rarely chosen.

Tables 26 and 27 show the results of the birthplaces analysis and the lower level of education analysis, respectively. The latter is similar to the estimation in the main text, with *distance* and *agglomeration* maintaining their high significance levels and with weak evidence for selection effect and no sorting effect. I find the main difference of the new threshold in the birthplace analysis. In Table 26 *distance*, *selection* and *agglomeration* correspond to those of the project, while *sorting* lose almost all the significance. It is slightly significant in column (3) when it is considered alone, but it is not significant when the model includes selection and/or agglomeration. This result is in favour of a more balanced choice set: the results gain clearness and precision when the threshold increased to 20 scholars. Keeping smaller universities in the dataset detracts attention from the real

Table 26 Multinomial logit regressions: standard logit model—birthplaces analysis, threshold at 5

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	− 0.777*** (0.026)	− 1.000*** (0.063)	− 0.777*** (0.026)	− 0.994*** (0.063)	− 0.774*** (0.025)	− 0.984*** (0.062)
Selection						
$\ln q \ln d$		0.046*** (0.012)		0.045*** (0.012)		0.044*** (0.012)
Sorting						
$\ln q \ln Q$			0.032* (0.018)	0.028 (0.018)		0.029 (0.018)
Agglomeration						
$\ln P_k$					− 0.155*** (0.026)	− 0.159*** (0.026)
$\ln Y_k$					2.535*** (0.310)	2.619*** (0.311)
$\ln Q$					0.208*** (0.040)	0.073 (0.093)
k FE	YES	YES	YES	YES	NO	NO
Obs	876	876	876	876	876	876
R^2	0.200	0.203	0.200	0.203		
LL	− 2017.696	− 2009.554	− 2016.089	− 2008.352	− 2070.630	− 2061.524

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 27 Multinomial logit regressions: standard logit model—lowest level of education analysis—threshold at 5

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	− 0.726*** (0.020)	− 0.791*** (0.046)	− 0.726*** (0.020)	− 0.793*** (0.046)	− 0.723*** (0.020)	− 0.789*** (0.045)
Selection						
$\ln q \ln d$		0.014 (0.009)		0.015* (0.009)		0.015* (0.009)
Sorting						
$\ln q \ln Q$			0.026 (0.018)	0.027 (0.017)		0.026 (0.017)
Agglomeration						
$\ln P_k$					− 0.127*** (0.026)	− 0.129*** (0.026)
$\ln Y_k$					1.086*** (0.317)	1.135*** (0.318)
$\ln Q$					0.178*** (0.040)	0.065 (0.087)
k FE	YES	YES	YES	YES	NO	NO
Obs	969	969	969	969	969	969
R^2	0.259	0.259	0.259	0.260		
LL	− 2069.106	− 2067.837	− 2068.037	− 2066.652	− 2108.792	− 2106.510

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

scholars’ choices, although the inclusion of Southern universities would improve the reach of the project.

Funding This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, under Grant agreement No 883033 Did elite human capital trigger the rise of the West? Insights from a new database of European scholars.

Data availability The dataset used in this paper is available upon request, and for privacy reasons the data will be made anonymous.

Declarations

Conflict of interest I also declare that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abramo, G., D'Angelo, C. A., & Di Costa, F. (2022). The effect of academic mobility on research performance: The case of Italy. *Quantitative Science Studies*, 3(2), 345–362.
- Agasisti, T., Barra, C., & Zotti, R. (2019). Research, knowledge transfer, and innovation: The effect of Italian universities' efficiency on local economic development 2006–2012. *Journal of Regional Science*, 59(5), 819–849.
- Agasisti, T., & Bianco, A.D. (2007). Determinants of college student migration in Italy: Empirical evidence from a gravity approach. Available at SSRN 1063481 .
- Agasisti, T., & Ricca, L. (2016). Comparing the efficiency of Italian public and private universities (2007–2011): An empirical analysis. *Italian Economic Journal*, 2(1), 57–89.
- Akcigit, U., Baslandze, S., & Stantcheva, S. (2016). Taxation and the international mobility of inventors. *American Economic Review*, 106(10), 2930–81.
- Aksnes, D. W., Rørstad, K., Piro, F. N., & Sivertsen, G. (2013). Are mobile researchers more productive and cited than non-mobile researchers? A large-scale study of Norwegian scientists. *Research Evaluation*, 22(4), 215–223.
- Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2), 101–138.
- Artige, L., Camacho, C., & de La Croix, D. (2004). Wealth breeds decline: Reversals of leadership and consumption habits. *Journal of Economic Growth*, 9(4), 423–449.
- Audretsch, B. (1998). Agglomeration and the location of innovative activity. *Oxford Review of Economic Policy*, 14(2), 18–29.
- Bahar, D., & Rapoport, H. (2018). Migration, knowledge diffusion and the comparative advantage of nations. *The Economic Journal*, 128(612), F273–F305.
- Bahar, D., Rapoport, H., & Turati, R. (2019). Does birthplace diversity affect economic complexity? Cross-country evidence. *CESifo Working Paper Series 7950, CESifo* .
- Baldissera, A., & Cornali, F. (2020). Geography of human capital in Italy: A comparison between macro-regions. *Modern Italy*, 25(3), 261–277.
- Barbieri, G., Rossetti, C., & Sestito, P. (2011). The determinants of teacher mobility: Evidence using Italian teachers' transfer applications. *Economics of Education Review*, 30(6), 1430–1444.
- Barone, C., & Guetto, R. (2016). Verso una meritocrazia dell'istruzione? Inerzia e mutamento nei legami tra origini sociali, opportunità di studio e destini lavorativi in Italia (1920–2009). *Polis*, 30(1), 5–34.
- Barra, C., & Zotti, R. (2017). Investigating the human capital development-growth nexus: Does the efficiency of universities matter? *International Regional Science Review*, 40(6), 638–678.
- Barra, C., & Zotti, R. (2018). The contribution of university, private and public sector resources to Italian regional innovation system (in) efficiency. *The Journal of Technology Transfer*, 43(2), 432–457.
- Barro, R. J. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, 106(2), 407–443.
- Barro, R. J. (2001). Human capital and growth. *American Economic Review*, 91(2), 12–17.
- Battistin, E., Checchi, D., & Verzillo, S. (2014). 02. La riforma di decentramento dei concorsi universitari in Italia. *RIV Rassegna Italiana di Valutazione*: 50–79.
- Beine, M., Bertoli, S., & Fernández-Huertas Moraga, J. (2016). A practitioner's guide to gravity models of international migration. *The World Economy*, 39(4), 496–512.
- Beine, M., Bierlaire, M., & Docquier, F. (2021). New York, Abu Dhabi, London or stay at home? Using a cross-nested logit model to identify complex substitution patterns in Migration. *IZA Discussion Papers 14090* .
- Beine, M., Docquier, F., & Özden, Ç. (2011). Diasporas. *Journal of Development Economics*, 95(1), 30–41.
- Beine, M., Docquier, F., & Rapoport, H. (2001). Brain drain and economic growth: Theory and evidence. *Journal of Development Economics*, 64(1), 275–289.
- Bertola, G., & Sestito, P. (2011). A comparative perspective on Italy's human capital accumulation. *Bank of Italy Economic History Working Paper No. 6* .
- Bertoli, S., & Moraga, J. F. H. (2013). Multilateral resistance to migration. *Journal of Development Economics*, 102, 79–100.
- Bertoli, S., & Rapoport, H. (2015). Heaven's swing door: Endogenous skills, migration networks, and the effectiveness of quality-selective immigration policies. *The Scandinavian Journal of Economics*, 117(2), 565–591.
- Bini, M., & Chiandotto, B. (2003). La valutazione del sistema universitario italiano alla luce della riforma dei cicli e degli ordinamenti didattici. *Studi e note di economia*, 2, 29–61.

- Blasutto, F., & de la Croix, D. (2021). Catholic censorship and the demise of knowledge production in early modern Italy. *CEPR Discussion Paper No. DP16409*.
- Börsch-Supan, A. (1990). On the compatibility of nested logit models with utility maximization. *Journal of Econometrics*, 43(3), 373–388.
- Bratti, M., & Verzillo, S. (2019). The ‘gravity’ of quality: Research quality and the attractiveness of universities in Italy. *Regional Studies*, 53(10), 1385–1396.
- Capano, G. (2008). Looking for serendipity: The problematical reform of government within Italy’s universities. *Higher Education*, 55(4), 481–504.
- Cecchi, D., De Fraja, G., & Verzillo, S. (2014, July). And the winners are... An axiomatic approach to selection from a set. IZA Discussion Papers 8344, Institute of Labor Economics.
- Cecchi, D., De Fraja, G., & Verzillo, S. (2021). Incentives and careers in academia: Theory and empirical analysis. *The Review of Economics and Statistics*, 103(4), 786–802. https://doi.org/10.1162/rest_a_00916.
- Cecchi, D., & Verzillo, S. (2014). *Selecting university professors in Italy: Much ado about nothing?* Mimeo.
- Cohen, D., & Soto, M. (2007). Growth and human capital: Good data, good results. *Journal of Economic Growth*, 12(1), 51–76.
- Cottini, E., Ghinetti, P., & Moriconi, S. (2019). Higher education supply, neighbourhood effects and economic welfare. *CESifo Working Paper Series 7483*, CESifo.
- Croissant, Y. (2020). Estimation of random utility models in r: The mlogit package. *Journal of Statistical Software*, 95(11), 1–41. <https://doi.org/10.18637/jss.v095.i11>
- Daly, A., & Zachary, S. (1978). *Improved multiple choice models*.
- de la Croix, D., Docquier, F., Fabre, A., & Stelter, R. (2022). The academic market and the rise of universities in medieval and early modern europe (1000–1800). WorkingPaper 2022-04, LISER.
- de la Croix, D., & Stelter, R. (2021). Scholars and literati at the university of göttingen (1734–1800). *Reperitorium eruditorum totius Europae*, 4, 1–8.
- Docquier, F., & Marfouk, A. (2006). International migration by education attainment, 1990–2000. *International migration, remittances and the brain drain*, pp. 151–199.
- Docquier, F., Turati, R., Valette, J., & Vasilakis, C. (2020). Birthplace diversity and economic growth: Evidence from the US states in the Post-World War II period. *Journal of Economic Geography*, 20(2), 321–354.
- Drucker, J., & Goldstein, H. (2007). Assessing the regional economic development impacts of universities: A review of current approaches. *International Regional Science Review*, 30(1), 20–46.
- Durante, R., Labartino, G., & Perotti, R. (2011, November). Academic dynasties: Decentralization and familism in the Italian academia. NBER Working Papers 17572, National Bureau of Economic Research, Inc.
- Easterly, W. (2001). *The elusive quest for growth—Economists’ adventures and misadventures in the tropics*. The MIT Press.
- Ejermo, O., Fassio, C., & Källström, J. (2020). Does mobility across universities raise scientific productivity? *Oxford Bulletin of Economics and Statistics*, 82(3), 603–624.
- Faggian, A., & McCann, P. (2009). Universities, agglomerations and graduate human capital mobility. *Journal of Economic and Social Geography*, 100(2), 210–223.
- Grogger, J., & Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, 95(1), 42–57.
- Grogger, J., & Hanson, G. H. (2015). Attracting talent: Location choices of foreign-born PhDs in the United States. *Journal of Labor Economics*, 33(S1), S5–S38.
- Handler, H. (2018). Economic links between education and migration: An overview. *Flash Paper Series, Policy Crossover Center Vienna-Europe*, 4, 1–17.
- Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3), 607–68.
- Heiss, F. (2002). Structural choice analysis with nested logit models. *The Stata Journal*, 2(3), 227–252. <https://doi.org/10.1177/1536867X0200200301>.
- Jasco, P. (2005). As we may search-Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. *Current Science*, 89(9), 1537–1547.
- Kerr, S. P., Kerr, W., Özden, Ç., & Parsons, C. (2016). Global talent flows. *Journal of Economic Perspectives*, 30(4), 83–106.
- Kerr, S. P., Kerr, W., Özden, Ç., & Parsons, C. (2017). High-skilled migration and agglomeration. *Annual Review of Economics*, 9, 201–234.
- Kling, C. L., & Herriges, J. A. (1995). An empirical investigation of the consistency of nested logit models with utility maximization. *American Journal of Agricultural Economics*, 77(4), 875–884.

- Kremer, M. (1993). The O-ring theory of economic development. *The Quarterly Journal of Economics*, 108(3), 551–575.
- León, G., & Miguel, E. (2013). Transportation choices and the value of statistical life. Technical Report 19494, National Bureau of Economic Research.
- Lucas, R. E., Jr. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1), 3–42.
- MacLeod, W. B., & Urquiola, M. (2021). Why does the United States have the best research universities? Incentives, resources, and virtuous circles. *Journal of Economic Perspectives*, 35(1), 185–206.
- McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior* University of California at Berkeley, Chapter 4.
- McFadden, D. (1978). *Modeling the choice of residential location*. Transportation Research Record.
- McFadden, D. (1979). Quantitative methods for analysing travel behaviour of individuals: Some recent developments. In D. Hensher & P. Stopher (Eds.), *Behavioural travel modelling* (pp. 279–318). Croom Helm.
- Morana, M. (2020). Focus “Il personale docente e non docente nel sistema universitario italiano”—a.a 2019/2020. *Elaborazioni su banche dati MIUR, DGCASIS—Ufficio VI Gestione patrimonio informativo e statistica*.
- Nelson, R. R., & Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *The American Economic Review*, 56(1/2), 69–75.
- Neuhaus, C., & Daniel, H. D. (2008). Data sources for performing citation analysis: An overview. *Journal of Documentation*, 64, 193–210.
- Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences’ literature. *Journal of Informetrics*, 1(2), 161–169.
- Ortega, F., & Peri, G. (2013). The effect of income and immigration policies on international migration. *Migration Studies*, 1(1), 47–74.
- Perotti, R. (2008). *L’università truccata*. Einaudi.
- Ramos, R. (2016). Gravity models: A tool for migration analysis. *IZA World of Labor*, p. 239.
- Rebora, G., & Turri, M. (2008). La governance del sistema universitario in italia: 1989–2008. *Liuc Papers Serie Economia Aziendale*, p. 32 .
- Rossi, P. (2016). Stato giuridico, reclutamento ed evoluzione della docenza universitaria (1975–2015). 4(1).
- Schiller, D., & Cordes, A. (2016). Measuring researcher mobility: A comparison of different datasets and methods with an empirical application of micro-data for the United States and Germany. In *OECD Blue Sky Forum*.
- Schwartz, A. (1976). Migration, age, and education. *Journal of Political Economy*, 84(4, Part 1), 701–719.
- Seiler, C., & Wohlrabe, K. (2011). Ranking economists and economic institutions using RePEc: Some remarks. *Ifo Working Paper No. 96*.
- Stephan, P. E., & Levin, S. G. (2001). Exceptional contributions to US science by the foreign-born and foreign-educated. *Population Research and Policy Review*, 20(1), 59–79.
- Train, K. (2003). *Discrete choice method with simulation*. Cambridge University Press.
- Train, K. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press.
- Train, K. E., McFadden, D. L., & Ben-Akiva, M. (1987). The demand for local telephone service: A fully discrete model of residential calling patterns and service choices. *The RAND Journal of Economics*, 18(1), 109–123.
- Trivellato, P., Triventi, M., & Traini, C. (2016). *Private higher education in Italy, a global perspective on private higher education* (pp. 29–51). Elsevier.
- Triventi, M., & Trivellato, P. (2008). Le onde lunghe dell’università italiana. Partecipazione e risultati accademici degli studenti nel Novecento. *Polis*, 22(1), 85–118.
- Williams, R. (2018). Panel data 3: Conditional logit/fixed effects logit models. *University of Notre Dame*, pp. 1–5 .
- Yang, K., & Meho, L. I. (2006). Citation analysis: A comparison of Google Scholar, Scopus, and Web of Science. *Proceedings of the American Society for Information Science and Technology*, 43(1), 1–15.
- Ye, M., Chen, Y., Yang, G., Wang, B., & Hu, Q. (2020). Mixed logit models for travelers’ mode shifting considering bike-sharing. *Sustainability*, 12(5), 2081.
- Zanardello, C. (2022). Market forces in Italian Academia today (and yesterday). *LIDAM Discussion Paper - 2022/01*.
- Zhao, X., Aref, S., Zagheni, E., & Stecklov, G. (2021). International migration in academia and citation performance: An analysis of German-affiliated researchers by gender and discipline using scopus publications 1996–2020.
- Zimmermann, C. (2013). Academic rankings with RePEc. *Econometrics*, 1(3), 249–280.