
Regional Studies

Routledge
Taylor & Francis Group

ISSN: 0034-3404 (Print) 1360-0591 (Online) Journal homepage: https://www.tandfonline.com/loi/cres20

# The ‘gravity’ of quality: research quality and the attractiveness of universities in Italy

Massimiliano Bratti & Stefano Verzillo

**To cite this article:** Massimiliano Bratti & Stefano Verzillo (2019) The ‘gravity’ of quality: research quality and the attractiveness of universities in Italy, Regional Studies, 53:10, 1385-1396, DOI: 10.1080/00343404.2019.1566701

**To link to this article:** https://doi.org/10.1080/00343404.2019.1566701

REGIONAL STUDIES
2019, VOL. 53, NO. 10, 1385–1396
https://doi.org/10.1080/00343404.2019.1566701



# The 'gravity' of quality: research quality and the attractiveness of universities in Italy

Massimiliano Bratti[a] and Stefano Verzillo[b]

**ABSTRACT**
This paper investigates whether or not research quality is significantly associated with a university's ability to attract students from other provinces in Italy. First-university enrolments of students over the period 2003–11 are regressed on universities' research-quality indicators computed from various bibliometric databases using fixed-effects gravity models. The estimates suggest that research performance is a significant predictor of student enrolment, with estimated elasticities between 0.013 and 0.059, depending on the indicator used.

**KEYWORDS**
research quality; student mobility; university; Italy

**JEL** I23
**HISTORY** Received 27 September 2017; in revised form 7 December 2018

## INTRODUCTION

Cross-country differences in the quality of higher education institutions (HEIs) are an important factor driving international student mobility (Abbott & Silles, 2016; Aslangbengui & Montecinos, 1998; Beine, Noël, & Ragot, 2014; Gordon & Jallade, 1996).

The quality of HEIs may play an even bigger role in the internal mobility of university students, given the absence of costs related to 'border effects' (e.g., language differences). Yet, evidence regarding the importance of the quality of HEIs for students' internal mobility is mixed.

This paper seeks to provide new evidence on this issue by analyzing the internal mobility of students in Italy. The Italian case is interesting given the high heterogeneity in the research performance of HEIs across regions and across the north–south divide. In recent years, southern Italian HEIs have been losing an increasing number of students. The drop in student enrolment numbers throughout the country was close to 20.4% (66,000 students) between 2004 and 2015, with a larger drop in the islands and southern regions (about 30.2% and 25.5%, respectively) than in the central and northern regions (Viesti, 2016). This haemorrhage of students from the south may partly be due to the lower quality of HEIs in this part of the country, which, along with the search for better employment opportunities (Dotti, Fratesi, Lenzi, & Percoco, 2013), pushes individuals to move towards the north. The present study assesses whether a lower research performance is an important driver of the 'brain drain' that southern Italy is currently experiencing.

Because of the difficulty of building a 'catch-all' indicator of university quality, the analysis keeps the main focus on research quality. In our empirical strategy, teaching quality is controlled for by including proxies such as the student–teacher ratio and the *Censis-Repubblica* score, a newspaper-based university ranking that takes into account universities' teaching outputs. Research quality is measured using multiple indicators computed from various bibliometric sources. Unlike most previous papers, which have often focused on only one or a few years, this study uses panel data spanning almost a decade. This enables one to control for universities' time-invariant unobservable characteristics through fixed effects models and to assess

**CONTACT**
[a] **(Corresponding author)** massimiliano.bratti@unimi.it
European Commission Joint Research Centre (JRC), Ispra, Italy; Department of Economics, Management and Quantitative Methods (DEMM), Università degli Studi di Milano, Milan, Italy; Institute of Labor Economics (IZA), Bonn, Germany; and Centro Studi Luca d'Agliano (LdA), Milan, Italy.
[b] stefano.verzillo@ec.europa.eu
European Commission Joint Research Centre (JRC), Ispra, Italy; and Department of Economics, Management and Quantitative Methods (DEMM), Università degli Studi di Milano, Milan, Italy.
Supplemental data for this article can be accessed at https://doi.org/10.1080/00343404.2019.1566701.

the importance of research quality in very different macroeconomic conditions.

Last but not least, following the most recent literature (e.g., Chevalier & Jia, 2016; Gibbons, Neumayer, & Perkins, 2015), this paper analyzes student enrolment flows using subject groups (i.e., fields of study) within HEIs as units of observation. This has some advantages over estimating gravity models at either the region or province level (e.g., Agasisti & Dal Bianco, 2007; Dotti et al., 2013) or pooling all subject groups together for the same HEI (e.g., Cattaneo, Malighetti, Meoli, & Paleari, 2017): averaging research-quality indicators may hide substantial heterogeneity in the quality of published research between HEIs located in the same geographical unit or between subject groups within the same HEI.

Our empirical strategy consists in the estimation of a gravity model in which province-level student inflows (i.e., first-time enrolments) by subject group from the Nomenclature of Territorial Units for Statistics level 3 (NUTS-3) regions (i.e., Italian provinces) of residence (origin) towards an HEI branch (defined at the HEI/destination-province level) are regressed on measures of the latter's research quality and a comprehensive set of fixed effects. The analysis in this paper is limited to first-time enrolments in first-level degrees. Single-cycle degrees (such as medicine or architecture) are excluded from the analysis because entry is generally selective and the econometric model described in Appendix A in the supplemental material may not be well suited to analyze students' choices.

Our analysis demonstrates that although higher research-quality HEIs attract more students, research quality is not the main driver of student enrolment choices in Italy. From a theoretical point of view, our results point to the fact that although studying in HEIs with higher research quality offers some benefits in the labour market, these benefits are not large enough to compensate for the mobility costs. Indeed, in Italy, student financial aid is among the lowest in Europe (European Commission, 2017) and mobile students often have to rent accommodation on the private market. As a consequence, even a large increase in research quality would not be able to overturn the effect of geographical distance on student choices. Our results confirm early findings that although university graduates from different alma maters command different wages in the labour market, Italian students exhibit low geographical mobility (Brunello & Cappellari, 2008).

Our findings are not necessarily generalizable to other countries, however, especially to those where there is a long tradition of research evaluation and where, thanks to better information available on the prestige and quality of alma maters (e.g., the Russell versus non-Russell groups of HEIs in the UK), employers are able to differentiate graduates' wages accordingly.

The paper is structured as follows. The next section introduces the conceptual framework, which represents the starting point for the specification of the gravity equation described in the third section. The fourth section describes the data and the research-quality indicators used in the empirical analysis. The empirical results are commented on in the fifth section. The final section summarizes the main findings and draws some conclusions.

## CONCEPTUAL FRAMEWORK AND RELATED LITERATURE

What is university 'quality' and why might students want to choose high-quality HEIs? In this section, these issues are discussed in light of the relevant literature.

First, it is necessary to define what is meant by the 'quality' of HEIs. Two different types of indicators of quality are considered in this paper, the first being *research-quality indicators*, namely bibliometric indicators. Although in Italy the higher education system is not characterized by a formal distinction between teaching and research universities, as a result of budgetary constraints, a debate has recently emerged regarding the concentration of research funds into a small number of 'excellent' universities, which would then assume the status of 'research universities'. In this context, our analysis sheds a light on the role of research quality as a potential driver of student enrolment choices. Second, in a robustness check, our analysis also considers university quality indicators produced by the statistical institute Censis in collaboration with the national newspaper *La Repubblica*. *La Repubblica* publishes a yearly ranking of HEIs by subject group (the so-called *Censis-Repubblica* University Guide – probably the most popular university ranking in Italy). The *Censis-Repubblica* ranking is based on a composite index that encompasses several components, namely student performance (such as the number of examinations passed, average marks and drop-out rates), teaching inputs (such as student–teacher ratios, infrastructure and student satisfaction questionnaires), research funding (mainly public), teacher quality (e.g., the quality of their research) and international cooperation (e.g., student and staff exchange and mobility programmes, international research projects). These rankings are published in both the web version of the newspaper and in more detailed form in a hardcopy version available at newsstands and bookshops.

There is little knowledge about whether research quality plays a significant role in the capacity of HEIs to attract students. In order for research quality to have an effect on student enrolment, prospective students must believe that it will positively affect their returns to education in the labour market. Regarding this, extant research for Italy has already demonstrated that the research quality of HEIs positively affects the future wages and employability of university graduates (Ciriaci & Muscio, 2014; Di Pietro & Cutillo, 2006).

There may be some concern that students, especially those enrolling at the undergraduate level, have limited access to information about research quality. However, in our empirical analysis, bibliometric indicators are used as *proxies* of research quality or research reputation, that is, our analysis does not rely on the assumption that students really know, for instance, the exact average impact factor or number of publications produced by the faculty of a particular HEI. For an effect of research quality to be observed, it is sufficient that students or their parents have enough formal or informal knowledge to rank universities and subject groups by the quality (or impact) of their research outputs. This may not be too far from reality, especially when students and their families choose between courses in different subject groups offered by the same HEI, or between different HEIs but within the same subject group and province, that is, the two types of variation exploited in our empirical models (see the third section).

Analyses of the effect of university rankings produced by newspapers or specialized intermediaries are instead more common in the literature. On the one hand, these indicators are easily accessible to the public and cover several aspects of universities' performance such as student satisfaction or employability. As such, they are likely to have an important impact on student choices. On the other hand, however, the methodology used to build these rankings is sometimes opaque or based on subjective (e.g., student) evaluations, and may be judged as not fully reliable.

For the United States, some studies have proxied university quality with their *prestige*. One study has found that the prestige of undergraduate programmes, measured by the Gourman ratings of colleges and universities and the Gross–Grambsch quality ratings of graduate facilities, explains only a small proportion of the variation in student inter-state migration (Abbott & Schmid, 1975). Somewhat contrasting are the results of another US study that proxies the quality of programmes with the selectivity of admissions, which show a positive effect on the attraction of non-resident students (Baryla & Dotterweich, 2001).

A positive association between university quality and regional student mobility is not ubiquitous outside the United States either. For instance, this relationship is not found by Sá, Florax, and Rietveld (2004), who use a composite quality index of educational programmes in the Netherlands based on a survey run by the weekly magazine *Elsevier*.[1]

Mixed results are also found in Italy, which is the focus of the current study. As for the results of the papers mentioned above, this may partly be due to the variety of proxies of university quality that scholars have used, which are likely to capture very different aspects of quality. These indicators include, *inter alia*, the proportion of faculty members who received national research funds (Dotti et al., 2013); composite indicators such as the *Censis-Repubblica*, which combine different aspects of university life such as student academic outcomes, teaching quality, research quality and international cooperation (Pigini & Staffolani, 2016); and the ranking obtained in a national research evaluation exercise (Ciriaci, 2014). Only the last two quality indicators were found to be positively associated with student choices.

## STUDENT GRAVITY EQUATIONS

The gravity equations described in this section are derived in Appendix A in the supplemental data online as the aggregation of students' individual enrolment choices.

The estimated gravity equation (equation A8 derived in Appendix A) reads as follows:

$$\begin{aligned} S_{jkt} = \exp[\ln(P_{jt}) + b_0 + b_1 \ln Q_{kt} + b_2 B_{ft} \\ + b_3 \ln STU\_TEACH_{hft} - a_0 - a_1 \ln DIST_{jb} \\ - a_2(\ln DIST_{jb} * SEA_b) - a_3 SEA_b - a_4 CONTIG_{jb} \\ - a_5 SAMEREG_{jb} - a_6 SAMEPROV_{jb} - a_7 C_{bt} \\ - a_8 C_{jt} - a_9 C_{ft} - \ln(\Omega_{jt})]h_{jkt} \end{aligned} \quad (1)$$

The dependent variable $S_{jkt}$ is the number of students coming from province $j$ enrolled in destination $k$ at time $t$. $j$ is the subscript for the origin province (i.e., the province students come from); $b$ is the branch in which students are enrolled. A 'branch' is defined by the combination HEI–province of destination. 'Destinations' ($k$) are defined as combinations of an HEI's branch and subject group (or field $f$). An HEI ($h$) may have branches in different provinces, and each branch may offer degree courses in several fields of study ($f$); $t$ is the time subscript.

As for the variables on the right-hand-side of equation (1), $\ln(P_{jt})$ is the logarithm of the population 'at risk' of enrolling in higher education (i.e., the stock of high school graduates); $\ln Q_{kt}$ is the logarithm of our proxies of research quality; $\ln STU\_TEACH_{hft}$ is the logarithm of the student–teacher ratio, used as a proxy of crowding and (lower) teaching quality;[2] $\ln DIST_{jb}$ is the logarithm of geodesic distance between the centroids of the origin province and of the province where the university branch is located (this variable has the same value for all destinations, i.e., branch–subject group combinations, within a branch);[3] $\ln DIST_{jb} * SEA_b$ is an interaction term between distance and an indicator ($SEA_b$) for either the province of origin or the destination branch being located on an island (Sicily or Sardinia), which captures the higher travel costs related to sea travel; $CONTIG_{jb}$ is a contiguity indicator between the province of origin and that of destination; $SAMEREG_{jb}$ is a dichotomous indicator for origin and destination provinces being located in the same region; $SAMEPROV_{jb}$ is a dichotomous indicator for the province of origin and that of the university branch being the same. $C_{bt}$, $C_{jt}$ and $C_{ft}$ capture time-varying cost components with variation at the level of the HEI's branch, the students' province of origin and the subject group, respectively. Variables such as the average level of unemployment, unskilled wages and living costs in the province of origin are captured by $C_{jt}$, while the average cost of living in the province where $b$ is located is captured by $C_{bt}$. Finally, $C_{ft}$ captures factors such as differences in average enrolment fees across subject groups at the national level and higher selectivity in student admissions for specific subject groups.[4]

$\ln(\Omega_{jt})$ is the logarithm of the 'multilateral resistance term' $\Omega_{jt}$ (Anderson & van Wincoop, 2003), which measures the average accessibility barriers to province $j$ compared with all other provinces; $B_{ft}$ is the average benefits from enrolling in field $f$ in year $t$ at the national level; and $h_{jkt}$ is an error term with $E(h_{jkt}) = 1$. Model (1) is estimated using the Poisson pseudo-maximum likelihood (PPML) estimator (Santos Silva & Tenreyro, 2006).

As measures of the variables $B_{ft}$, $C_{bt}$, $C_{jt}$ and $C_{ft}$ are not available, they are proxied with fixed effects ($D_{.t}$). Therefore, after noting that $\ln(P_{jt})$ and $\ln(\Omega_{jt})$ exhibit variation at the origin province by time level, equation (1) becomes:

$$\begin{aligned} S_{jkt} = \exp [&\gamma_0 + b_1 \ln Q_{kt} + b_3 \ln STU\_TEACH_{hft} \\ &- a_1 \ln DIST_{jb} - a_2 (\ln DIST_{jb} * SEA_b) - a_3 SEA_b \\ &- a_4 CONTIG_{jb} - a_5 SAMEREG_{jb} \\ &- a_6 SAMEPROV_{jb} + D_{bt} + D_{jt} + D_{ft}]\, h_{jkt} \end{aligned} \quad (2)$$

(subscripts: $j$ = origin province; $k$ = destination; $b$ = destination branch; $h$ = HEI; $f$ = subject group; and $t$ = time) where $\gamma_0 = b_0 - a_0$.

Our coefficient of interest is $b_1$, which measures the 'economic return to university research quality', for example, the expected increase in wages that university students will reap in the labour market after graduation thanks to attending HEIs of better quality. Several research-quality indicators are considered (described in the following section), some of which measure more quantitative dimensions of research output, and others that are more qualitative. By including time-variant university branch fixed effects ($D_{bt}$), only between-field variation in research quality within the same university branch is exploited in each year. The fixed effects $D_{ft}$ capture, *inter alia*, average differences in productivity and the impact of research between subject groups.[5]

Since equation (1) is likely to identify mainly the effect of research quality in HEIs that offer courses in multiple fields within the same branch, an alternative specification exploiting within-subject group variation across HEI branches located in the same destination province is also estimated:

$$\begin{aligned} S_{jkt} = \exp (&\gamma_0 + b_1 \ln Q_{kt} + b_3 \ln STU\_TEACH_{hft} \\ &- a_1 \ln DIST_{jb} - a_2 \ln DIST_{jb} * SEA_b \\ &- a_3 SEA_b - a_4 CONTIG_{jb} - a_5 SAMEREG_{jb} \\ &- a_6 SAMEPROV_{jb} + D_{pt} + D_{jt} + D_{ft})\, h_{jkt} \end{aligned} \quad (3)$$

(subscripts: $j$ = origin province; $p$ = destination province; $k$ = destination; $h$ = HEI; $f$ = subject group; and $t$ = time), where the destination branch–year fixed effects in equation (2) have been replaced with destination province–year fixed effects ($D_{pt}$).

Pull factors associated with differences in the attractiveness of destination provinces (e.g., the state of the labour market, available amenities and infrastructure) are controlled for through branch–year or province–year fixed effects in the first and second specifications, respectively. This ensures that research quality only captures a feature of HEIs and does not reflect province-level unobservable characteristics.

In the PPML estimates, coefficients of variables in logarithm scale can be interpreted as elasticities, and coefficients of variables in levels as percentage changes in the dependent variable.

## DATA

The variables' names, definitions and sources are described in Table 1. Further details on the data are given in Appendix B in the supplemental data online, which also reports the sample descriptive statistics (see Table B1 online).

This section discusses the research-quality indicators in detail.

### Description of the research-quality indicators

The analysis of university attractiveness proposed in this paper is based on the estimation of PPML gravity models using observations corresponding to the cells defined by students' province of origin (proxied by student residence at the time of enrolment) and the province, field of study and HEI of the degree course in which they enrolled, where the pairing of province–HEI defines a university branch. Consequently, the model estimation requires the computation of indicators of research quality for each university branch and scientific field.

Unfortunately, information is not available to link the flows of students attracted by each branch–subject group (i.e., destination) with the exact pool of researchers and professors providing them with teaching activities. Moreover, although potential students may have some information on the scientific reputation of teachers working in an HEI and belonging to a given subject group, it is perhaps implausible to assume that they know the scientific quality of the particular academics teaching the degree course in which they plan to enrol. Hence, student flows are assigned the average research quality of all professors and researchers belonging to the same HEI and the same 'broad scientific subject groups' of the degree course in which they enrolled. 'Broad scientific subject groups' are defined as the best lexicographical match between the subject group classification of university courses by the Italian Ministry of Education, University and Research (MIUR) for teaching purposes and the scientific research areas according to which the academic research staff are hired ('scientific disciplinary areas' as defined by the National University Council – CUN).[6] The correspondence table between teaching-related subject groups (MIUR) and scientific-related subject groups is provided in Table C1 in Appendix C in the supplemental data online.

Taking advantage of the above-mentioned bibliometric sources, the following seven research-quality indicators were computed for each researcher in field $f$ affiliated with institution $h$:[7]

- *Average number of Institute for Scientific Information (ISI) publications*$_{fht}$: the weighted average of the total number of publications recorded in the ISI Web of Science (WoS) in year $t$.
- *Average number of ISI citations*$_{fht}$: the weighted average of the total number of citations in ISI WoS publications in year $t$.

**Table 1.** Description and source of explanatory variables.

| Variables | Description | Data source |
|---|---|---|
| Student flows | Flow of students from a province towards a destination (i.e., higher education institutions’ (HEI) branch–subject group) for all first-level degrees | MIUR Statistical Office |
| Distance (km) | Geodesic distance (the shortest route between the two centroids) of each pair of provinces | Authors’ computation on the coordinates of provinces’ centroids provided by ISTAT |
| Sea | Dichotomous indicator for either the origin or the destination province being an island | Authors’ computation on the data set |
| Sea × Distance | Interaction effect between Distance and Sea | Authors’ computation on the data set |
| Province contiguity | Dummy variable taking the value 1 for each pair of provinces sharing at least 1 km of border | ISTAT |
| Flow within region | Dichotomous indicator for student inflow from within the region where the university branch is located | Authors’ computation on the data set |
| Flow within province | Dichotomous indicator for student inflow from within the province where the university branch is located | Authors’ computation on the data set |
| Student–teacher ratio | Average number of enrolled students per academic professor in the university–field–province of destination | MIUR Statistical Office |
| *Censis* score | HEI subject group score from the *Censis-Repubblica* University Guide | *Censis-Repubblica* |
| *Censis* score missing | Dichotomous indicator for missing *Censis-Repubblica* score | Authors’ computation on the data set |
| ISI publications | Average number of publications collected on Institute for Scientific Information (ISI) Web of Science (WoS) | ISI WoS (downloaded 2012) |
| ISI citations | Average number of citations of all publications collected on ISI WoS | ISI WoS (downloaded 2012) |
| ISI impact factor | Average impact factor of all publications collected from WoS | ISI WoS plus Journal Citation Reports (downloaded 2012) |
| GS publications | Average number of publications collected on Google Scholar (GS) | GS (downloaded 2015) |
| GS citations | Average number of citations of all publications collected from GS | GS (downloaded 2015) |
| Scopus publications | Average number of publications collected from Scopus | Scopus (downloaded 2016) |
| Scopus citations | Average number of citation of all publications collected from Scopus | Scopus (downloaded 2016) |

Notes: All variables refer to the period 2003–11. Distance, student–teacher ratio, *Censis* score and research-quality indicators are measured in logarithms. ISTAT, Italian National Statistical Institute; MIUR, Ministry of Education, University and Research.

- *Average impact factor of ISI publications*$_{fbt}$: the weighted average of the impact factor (Journal of Citation Reports) of ISI WoS publications in year *t*.
- *Average number of GS publications*$_{fbt}$: the weighted average of the total number of publications in Google Scholar (GS) in year *t*.
- *Average number of GS citation*$_{fbt}$: the weighted average of the number of citations in GS publications in year *t*.
- *Average number of Scopus publications*$_{fbt}$: the weighted average of the total number of publications in Scopus in year *t*.
- *Average number of Scopus citations*$_{fbt}$: the weighted average of the total number of citations in Scopus publications in year *t*.

When ‘broad scientific subject groups’ encompass more than one CUN scientific group, weighted averages are computed according to the number of enrolled first-level students belonging to each scientific group (CUN). Moreover, pure teaching branches, that is, branches that do not formally have research staff linked to them, are attributed the research quality of their headquarters.

Although indicators from the same source are highly correlated (see Table D1 in Appendix D in the supplemental data online), the same is not true of indicators from different sources, suggesting that they have some independent variation. The main advantage of using several bibliometric indicators to proxy for research quality is that they differ in their level of ‘inclusivity’ or coverage of research output. ISI is the least inclusive (i.e., the most selective), followed by Scopus and GS. Thus, within each subject group, research outputs recorded in ISI and Scopus can generally be considered of 'better quality' compared with those published in GS, as they have to meet a given standard for inclusion in those databases. This also entails differential coverage of subject groups, with GS generally having a better coverage of the arts and humanities (i.e., non-bibliometric disciplines as defined by the National Agency for the Evaluation of University and Research – ANVUR) and of articles and books not written in English, compared with ISI and Scopus. Despite the larger coverage, GS indicators are more likely to suffer from duplications, as this database also includes 'unpublished' work such as discussion papers or papers published in multiple working paper series.

## EMPIRICAL RESULTS

### Models with time-variant university 'branch' (i.e., HEI–province) fixed effects

This section reports the results of the model including time-variant university 'branch' fixed effects, that is, the gravity equation (1). In this model, identification comes from differences in research quality between the subject groups of degree courses provided by the same university branch in a given year.

The estimates in all columns of Table 2 use the same control variables, but different indicators of research quality (RQ, hereafter). Our models explain around 83% of the total variance in student inflows. Results generally show the expected signs for the control variables included.

**Table 2.** Model with time-variant university branch fixed effects.

| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Distance | –0.983*** | –0.984*** | –0.984*** | –0.984*** | –0.984*** | –0.984*** | –0.984*** |
| | (0.059) | (0.059) | (0.059) | (0.059) | (0.059) | (0.059) | (0.059) |
| Sea | 1.351** | 1.353** | 1.353** | 1.353** | 1.353** | 1.353** | 1.353** |
| | (0.578) | (0.578) | (0.578) | (0.578) | (0.578) | (0.578) | (0.578) |
| Sea × Distance | –0.353*** | –0.354*** | –0.354*** | –0.354*** | –0.354*** | –0.354*** | –0.354*** |
| | (0.101) | (0.101) | (0.101) | (0.101) | (0.101) | (0.101) | (0.101) |
| Province contiguity | 1.113*** | 1.116*** | 1.116*** | 1.116*** | 1.116*** | 1.116*** | 1.116*** |
| | (0.111) | (0.111) | (0.111) | (0.111) | (0.111) | (0.111) | (0.111) |
| Flow within region | 1.160*** | 1.160*** | 1.160*** | 1.160*** | 1.160*** | 1.160*** | 1.160*** |
| | (0.133) | (0.133) | (0.133) | (0.133) | (0.133) | (0.133) | (0.133) |
| Flow within province | –5.295*** | –5.302*** | –5.302*** | –5.302*** | –5.302*** | –5.302*** | –5.302*** |
| | (0.554) | (0.554) | (0.554) | (0.554) | (0.554) | (0.554) | (0.554) |
| Student–teacher ratio | 0.192*** | 0.193*** | 0.192*** | 0.194*** | 0.192*** | 0.192*** | 0.192*** |
| | (0.017) | (0.017) | (0.017) | (0.016) | (0.016) | (0.017) | (0.017) |
| ISI publications | 0.026*** | | | | | | |
| | (0.007) | | | | | | |
| ISI citations | | 0.013** | | | | | |
| | | (0.005) | | | | | |
| ISI impact factor | | | 0.017*** | | | | |
| | | | (0.005) | | | | |
| GS publications | | | | 0.038** | | | |
| | | | | (0.015) | | | |
| GS citations | | | | | 0.034*** | | |
| | | | | | (0.009) | | |
| Scopus publications | | | | | | 0.059*** | |
| | | | | | | (0.01) | |
| Scopus citations | | | | | | | 0.026*** |
| | | | | | | | (0.006) |
| Constant | –1.508*** | 2.487*** | 2.504*** | 2.584*** | 2.593*** | 2.712*** | 2.564*** |
| | (0.498) | (0.614) | (0.612) | (0.625) | (0.622) | (0.606) | (0.609) |
| Observations[a] | 757,617 | 768,521 | 768,521 | 768,521 | 768,521 | 768,521 | 768,521 |
| $R^2$ | 0.833 | 0.833 | 0.834 | 0.833 | 0.834 | 0.835 | 0.834 |

Notes*:* The dependent variable is the number of students enrolled in a university branch, in a given subject group and coming from a given Italian province. All models also control for subject group–year, branch–year and origin province–year fixed effects. Distance, the student–teacher ratio and research-quality indicators are measured in logarithms. Standard errors clustered by province of origin are in parentheses.
[a]The number of observations may vary across columns since the STATA command ppml drops some observations causing convergence problems.
*, **, *** statistically significant at the 10%, 5% and 1% levels, respectively.

A negative and statistically significant relationship emerges between student inflow and the geodesic distance between a student's province of residence and the province of the degree course in which the student is enrolled, with elasticity very close to −1. The elasticity of student inflows with respect to distance decreases by about −0.35 for travels that involve sea travel.

As expected, province contiguity is positively associated with student inflow, with contiguous provinces enjoying a premium of about 1% in student inflow (1.11–1.12% depending on the research-quality indicator used). This confirms, conditional on distance, students' attitudes regarding moving from their province of origin to a destination in an adjacent province. Inflows of students within provinces in the same region are also more likely to occur than flows involving provinces located in different regions. In all the estimated models in Table 2, the coefficients of these control variables are highly stable across columns. The negative coefficient on within-province inflows suggests that, on average, HEI branches attract more students from outside the province than within the province where they are located (on average around 5.3% more).

Teaching quality is proxied by the (log of) the student–teacher ratio, as is common in the related literature (Agasisti & Dal Bianco, 2007; Ciriaci, 2014). The estimates show a statistically significant elasticity of 0.19. The result of a *positive* effect of the student–teacher ratio on student inflow can be explained by the mechanical positive correlation between degree courses in high demand (i.e., 'crowded' courses) and student inflow.

Turning now to the variables of interest, the RQ indicators, Table 2 demonstrates the high significance of research quality in explaining student flows.

All RQ indicators are positively associated with the attractiveness of university branches and are statistically significant at least at the 1% level. Student inflow shows a 0.026 elasticity to *research productivity*, proxied by the number of ISI publications (column 1). The elasticity is somewhat smaller when *research influence* is considered, using the average number of ISI citations and the average ISI impact factor, with elasticities of 0.013 and 0.017 in columns (2) and (3), respectively. Interestingly, the effects appear to be larger when GS's research-quality indicators are used, with an elasticity of student flow to scientific productivity and impact of 0.038 (column 4) and 0.034 (column 5), respectively. This may partly be explained by the higher accessibility of the scientific productions collected by GS, which has a much better coverage of articles and books written in Italian, and the higher inclusivity of GS for subject groups in the arts and humanities. Lastly, the elasticity of student enrolments is largest (0.059) when the average number of Scopus publications is considered (column 6). The elasticity of student inflow to the number of Scopus citations lies instead in between those estimated with WoS and GS (0.026).

Although the coefficients on RQ indicators are statistically significant and positive, distance plays a predominant role in predicting student inflow. Focusing, for instance, on the estimates in column (1) of Table 2, for an HEI located on the mainland, a 38% increase in RQ measured by the average number of ISI publications would be necessary to compensate for the negative effect of a 1% increase in geodesic distance on student inflow (from the mainland).[8]

## Models with time-variant destination-province fixed effects

Table 3 shows the estimates of the gravity equation exploiting the within-subject group variation between HEI branches located in the same province, that is, equation (3).

These estimates are reported for the sake of completeness, although compared with those in the previous section, they are less robust to the presence of unobserved HEI branch characteristics potentially correlated with the attractiveness of universities. They are nonetheless informative, as they show whether students tend to choose among HEI branches *in the same subject group and province* according to research quality.

The results are consistent with those in the previous section, although the elasticities are all larger in magnitude. The estimated elasticities are 0.131, 0.085 and 0.106 for the number of ISI publications, citations and impact factor, respectively; 0.16 and 0.146 for the number of GS publications and citations, respectively; and 0.145 and 0.083 for the number of Scopus publications and citations, respectively. The larger point estimates compared with those in Table 2 may partly reflect HEI-specific factors that are omitted from the regression. Interestingly, the coefficients of the control variables are very close to those in Table 2.

## Interactions between research quality and distance

This section investigates the potential interactions between distance and research quality. Students may attach more value to quality when choosing between HEIs of better quality does not entail a huge increase in enrolment costs. To put it differently, students might be more sensitive to quality when choosing between universities not too far from their residence. Alternatively, quality may be particularly salient when students have already decided to move, and for longer moves, for example, for students migrating from the south. These interactions are shown in Figure 1, which, in addition to the estimates of Table 2 for the full sample, also reports the coefficients (with their confidence intervals) on the log of research-quality indicators estimated in two subsamples, one including only intra-region and not intra-province (i.e., 'stayers') student enrolments, and the other only interregion student enrolments (i.e., from other regions). The estimated elasticities are generally larger for inter- than for intra-region student enrolments, with the exception of the GS indicators, but confidence intervals are often overlapping.

Finally, Figure 1 also plots the coefficients estimated in the sample of students moving from southern Italy and the islands towards northern Italy HEIs (i.e., a gravity model for northern HEI destinations limited to southern Italian and island provinces of origin only) focusing only on very

**Table 3.** Model with destination province–year fixed effects.

| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Distance | –0.984*** | –0.984*** | –0.984*** | –0.984*** | –0.984*** | –0.984*** | –0.984*** |
| | (0.059) | (0.059) | (0.059) | (0.059) | (0.059) | (0.059) | (0.059) |
| Sea | 1.353** | 1.353** | 1.353** | 1.353** | 1.353** | 1.353** | 1.353** |
| | (0.578) | (0.578) | (0.578) | (0.578) | (0.578) | (0.578) | (0.578) |
| Sea × Distance | –0.354*** | –0.354*** | –0.354*** | –0.354*** | –0.354*** | –0.354*** | –0.354*** |
| | (0.101) | (0.101) | (0.101) | (0.101) | (0.101) | (0.101) | (0.101) |
| Province contiguity | 1.116*** | 1.116*** | 1.116*** | 1.116*** | 1.116*** | 1.116*** | 1.116*** |
| | (0.111) | (0.111) | (0.111) | (0.111) | (0.111) | (0.111) | (0.111) |
| Flow within region | 1.160*** | 1.160*** | 1.160*** | 1.160*** | 1.160*** | 1.160*** | 1.160*** |
| | (0.133) | (0.133) | (0.133) | (0.133) | (0.133) | (0.133) | (0.133) |
| Flow within province | –5.302*** | –5.302*** | –5.302*** | –5.302*** | –5.302*** | –5.302*** | –5.302*** |
| | (0.554) | (0.554) | (0.554) | (0.554) | (0.554) | (0.554) | (0.554) |
| Student–teacher ratio (log) | 0.133*** | 0.131*** | 0.132*** | 0.125*** | 0.127*** | 0.142*** | 0.137*** |
| | (0.016) | (0.016) | (0.016) | (0.019) | (0.017) | (0.013) | (0.013) |
| ISI publications | 0.131*** | | | | | | |
| | (0.025) | | | | | | |
| ISI citations | | 0.085*** | | | | | |
| | | (0.016) | | | | | |
| ISI impact factor | | | 0.106*** | | | | |
| | | | (0.021) | | | | |
| GS publications | | | | 0.160*** | | | |
| | | | | (0.045) | | | |
| GS citations | | | | | 0.146*** | | |
| | | | | | (0.031) | | |
| Scopus publications | | | | | | 0.145*** | |
| | | | | | | (0.030) | |
| Scopus citations | | | | | | | 0.083*** |
| | | | | | | | (0.018) |
| Constant | 0.611 | –1.832*** | –1.907*** | –1.782*** | 0.18 | –2.646 | –2.727*** |
| | (0.426) | (0.507) | (0.508) | (0.414) | (0.518) | (0.600) | (0.599) |
| Observations | 768,521 | 768,521 | 768,521 | 768,521 | 768,521 | 768,521 | 768,521 |
| $R^2$ | 0.732 | 0.734 | 0.739 | 0.706 | 0.724 | 0.746 | 0.739 |

Note: The dependent variable is the number of students enrolled in a university branch, in a given subject group and coming from a given Italian province. All models also control for subject group–year, destination province–year and origin province–year fixed effects. Distance, the student–teacher ratio and research-quality indicators are measured in logarithms. Standard errors clustered by province of origin are in parentheses.
*, **, ***Statistically significant at the 10%, 5% and 1% statistical levels, respectively.

long-distance movers. All the RQ indicators except the GS indicators show positive and significant elasticities. To give one example of the magnitude of the effects, a 1% increase in the number of ISI publications (citations) for northern Italy's HEIs would lead to an average 0.099% (0.087%) increase in student enrolments from southern Italy and the islands.

The estimates using GS indicators exhibit a negative and significant elasticity for the number of publications and an insignificant elasticity for the number of citations. They are, however, the least reliable RQ indicators because of the presence of unpublished work and potential duplications of publications.

An additional model including a dichotomous indicator for 'no-commuting' distance (i.e., a distance greater than 50 km) is reported in Table E1 in Appendix E in the supplemental data online. Although this specification has the advantage of allowing for a direct test of the relative importance of research quality for non-commuting students vis-à-vis commuting students, compared with the sample-split estimates in Figure 1, it has the weakness of treating all non-commuting students in the same way, irrespective of distance. The results are broadly consistent with those in Table 2, with some estimates suggesting a significantly smaller effect of quality for non-commuting students, namely those using the GS quality indicators (similarly to what is found in Figure 1). All in all, this section highlights that research quality spurs an HEI's capacity to attract students from both nearby and distant provinces.

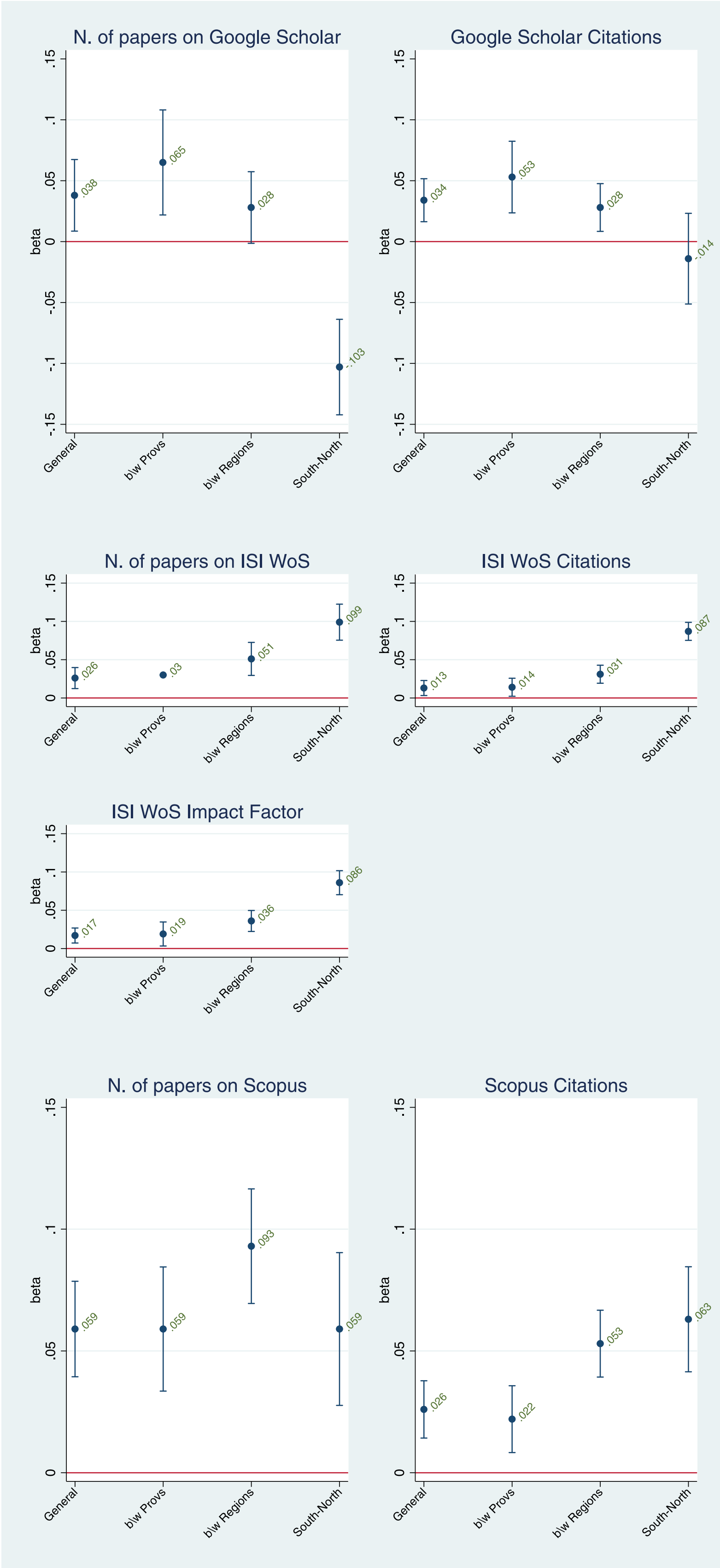


**Figure 1.** Elasticities of student enrolments to research-quality indicators by sample split.
Note: Reported are point estimates and 95% confidence intervals of the elasticities of student enrolments to research-quality indicators on the general (General), inter-province but intra-region (b/w Provs), interregion (b/w Regions), and south and islands to the north (South–North) student flow samples. ISI WoS = ISI Web of Science. Confidence intervals are not shown when the econometric software fails to compute standard errors.

### Research quality versus newspaper league tables

As previously mentioned in the introduction, students have full access to rankings of HEIs published every year by newspapers, such as the *Censis-Repubblica* University Guide. In order to evaluate whether RQ indicators predict student flows over and above the *Censis-Repubblica* (C-R) score, the baseline gravity model in Table 2 is re-estimated including the C-R score. Since not all HEI subject groups were ranked in each year, a dummy for missing C-R score is also included in the regression. Table E2 in Appendix E in the supplemental data online shows that the coefficients on the RQ indicators remain unaffected while the C-R score, in spite of its wide accessibility to students and coverage of several aspects of university quality (see the second section), does not significantly predict student inflow.

## CONCLUSIONS

This study employs gravity models to investigate the impact of research quality on students' internal mobility using Italian data covering almost a decade (2003–11). In a first for Italy, using panel data and framing the analysis at the subject group level allows one to leverage on between-subject group differences in research quality within university branches (and in another specification, within-subject group differences across university branches in the same province), controlling, *inter alia*, for potential unobservable variables at the HEI branch level (e.g., an alma mater's reputation). Using a large set of fixed effects, the gravity model's estimates provide evidence that a stronger research performance is associated with larger inflows of university students. According to our preferred estimates (Table 2) – controlling for time-variant university branches fixed effects – the elasticity of student inflow to research quality varies between 0.013 and 0.059, depending on the research-quality indicator used (ISI WoS, GS or Scopus).

Further analysis confirms that research-quality indicators are generally positively associated with student enrolment both from nearby and distant provinces, and that research-quality indicators predict student enrolment over and above university rankings published by newspapers (namely the *Censis-Repubblica* University Guide).

In spite of its statistical significance, research quality still appears to play a minor role in students' HEI subject group choices compared with geographical distance, suggesting that when deciding where to enrol, Italian students weight relocation and living costs more than the potentially higher returns of enrolling in a research-intensive institution. This is both good and bad news. On the one hand, it suggests that a worsening of the research gap between northern and southern universities in Italy per se is unlikely to be the main driver of south Italy's secondary school graduates' enrolling in northern HEIs. On the other hand, it means that an HEI's efforts towards improving research quality would not lead to important gains in student enrolment.

This paper leaves some open issues that could be addressed in future research. First, it suggests that in Italy economic returns to research quality are low. This could be further investigated by estimating university graduates' wage equations including research quality among the explanatory variables and using individual-level data. Second, it may be interesting to assess whether information asymmetries in the labour market are a potential explanation for the low economic returns to research quality. On this specific aspect, scholars could investigate whether the returns to research quality increased after the two recent Italian Research Evaluation Exercises (*Valutazione della Qualità della Ricerca* – VQR), which should have removed most of the pre-existing information asymmetries.

## ACKNOWLEDGEMENTS

This version of the paper benefited from comments received at the University of Cagliari, the 6th International Workshop on Applied Economics of Education (IWAEE, Catanzaro), the XXXIst National Conference of Labour Economics (Cagliari), the European Regional Science Association/Regional Science Association International (ERSA/RSAI) 2015 Barcelona Workshop on Regional and Urban Economics: Spatial Perspectives of Human Capital (Barcelona) and the conference *Il Sistema universitario: è già 'Buona università?* (Naples). Comments by the Associate Editor (Ugo Fratesi), three anonymous reviewers and Marco Bertoni, Steve Gibbons and Stefano Usai are especially acknowledged. The information and views set out in this paper are those of the authors and do not reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use that may be made of the information contained therein. The usual disclaimers apply.

## DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

## FUNDING

Funding from ARTI Puglia is gratefully acknowledged (Research Project: Studi ed analisi in specifiche aree tematiche d'interesse dell'osservatorio Regionale dei Sistemi di Istruzione e Formazione della Regione Puglia – Lotto 2 – 'Il sistema universitario italiano: un'analisi regionale').

## NOTES

1. Other papers analyze the effect of league tables on student enrolment by HEI, subject group and time, but do not consider student origin, that is, they do not estimate gravity models. For the UK, see, for instance, Chevalier and Jia (2016) and Gibbons et al. (2015).

2. This is computed as the ratio between the number of students and the number of permanent academic staff (i.e., assistant, associate and full professors). This student–teacher ratio is likely to be overestimated because temporary personnel (e.g., research assistants or doctoral students) can also do some teaching, and especially because it neglects external professors. Unfortunately, data on the latter are not available.

3. Geodesic distance is strongly correlated with travel distance but, unlike the latter, it does not depend on the endogenous choice of means of transportation. Geodesic distance is often used in student gravity models (Dotti, Fratesi, Lenzi, & Percoco, 2013, 2014). Although geodesic distance is likely to underestimate travel costs, the measurement error is unlikely to be correlated with research-quality indicators, whose coefficients are the main focus of the current paper.

4. Unfortunately, in our data, average fees cannot be computed by either field of study or branch, and vary only at the HEI level.

5. In a recent paper, Cattaneo et al. (2017) investigate the effect of spatial competition among universities. In our models, spatial competition factors are captured by the multilateral resistance term, proxied by destination-province fixed effects. Moreover, competitors’ proximity indexes similar to those included by Cattaneo et al. are subsumed in our models in branch–year fixed effects.

6. In the Italian higher education system, each researcher is allocated to one of 370 scientific sectors (*settori scientifico disciplinari*), which are mainly relevant for career progression because the hiring and promotion procedures are carried out within these sectors. Degrees are instead classified in ‘degree classes’ (*classi di laurea*) according to their prevailing teaching content. The lexicographic match implies that a student enrolled in a course in the economics degree class primarily cares about the scientific quality of the HEIs in economics, and not, for instance, about the quality of courses in other subject groups that are not the ‘core’ of their degree, such as law or sociology. Prospective students in economics will be very unlikely to have or gather detailed information on the quality of all teachers of the degree course, and especially on those teaching other subject groups. Should this assumption be very far from reality, the consequence will be a low predictive power of our econometric models. On the assumptions made by econometricians about student behaviour, see the seminal paper by Manski (1993).

7. For a detailed description of the procedure followed to build these indicators, see Bratti and Verzillo (2017, appx C).

8. This is computed by setting $b_1 \ln Q_{kt} = a_1 \ln DIST_{jb}$ in equation (1), obtaining $\ln Q_{kt} = \left(\frac{a_1}{b_1}\right) * \ln DIST_{jb}$.

## ORCID

*Massimiliano Bratti* http://orcid.org/0000-0002-4565-6260
*Stefano Verzillo* http://orcid.org/0000-0002-1895-8554

## REFERENCES

Abbott, W. R. F., & Schmid, C. F. (1975). University prestige and first-time undergraduate migration in the United States. *Sociology of Education*, *48*(2), 168–185. doi:10.2307/2112474

Abbott, A., & Silles, M. (2016). Determinants of international student migration. *World Economy*, *39*(5), 621–635. doi:10.1111/twec.12319

Agasisti, T., & Dal Bianco, A. (2007). *Determinants of college student migration in Italy: Empirical evidence from a gravity approach*. http://dx.doi.org/10.2139/ssrn.1063481

Anderson, J. E., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, *93*(1), 170–192. doi:10.1257/000282803321455214

Aslangbengui, N., & Montecinos, V. (1998). Foreign students in US doctoral programs. *Journal of Economic Perspectives*, *12*(3), 171–182. doi:10.1257/jep.12.3.171

Baryla, E., & Dotterweich, D. (2001). Student migration: Do significant factors vary by region? *Education Economics*, *9*(3), 269–280. doi:10.1080/09645290110086135

Beine, M., Noël, R., & Ragot, L. (2014). Determinants of the international mobility of students. *Economics of Education Review*, *41*(C), 40–54. doi:10.1016/j.econedurev.2014.03.003

Bratti, M., & Verzillo, S. (2017). *The ‘gravity’ of quality: Research quality and universities’ attractiveness in Italy* (IZA DP No. 11026). Bonn: Institute of Labor Economics (IZA).

Brunello, G., & Cappellari, L. (2008). The labour market effects of alma mater: Evidence from Italy. *Economics of Education Review*, *27*(5), 564–574. doi:10.1016/j.econedurev.2007.05.004

Cattaneo, M., Malighetti, P., Meoli, M., & Paleari, S. (2017). University spatial competition for students: The Italian case. *Regional Studies*, *51*(5), 750–764. doi:10.1080/00343404.2015.1135240

Chevalier, A., & Jia, X. (2016). Subject-specific league tables and students’ application decisions. *Manchester School*, *84*(5), 600–620. doi:10.1111/manc.12124

Ciriaci, D. (2014). Does university quality influence the interregional mobility of students and graduates? The case of Italy. *Regional Studies*, *48*(10), 1592–1608. doi:10.1080/00343404.2013.821569

Ciriaci, D., & Muscio, A. (2014). University choice, research quality and graduates’ employability: Evidence from Italian national survey data. *European Educational Research Journal*, *13*(2), 199–219. doi:10.2304/eerj.2014.13.2.199

Di Pietro, G., & Cutillo, A. (2006). University quality and labour market outcomes in Italy. *Labour (Committee on Canadian Labour History)*, *20*(1), 37–62. doi:10.1111/j.1467-9914.2006.00333.x

Dotti, N. F., Fratesi, U., Lenzi, C., & Percoco, M. (2014). Local labour market conditions and the spatial mobility of science and technology university students: Evidence from Italy. *Review of Regional Research*, *34*, 119–113. doi:10.1007/s10037-014-0088-y

Dotti, N. F., Fratesi, U., Lenzi, C., & Percoco, M. (2013). Local labour markets and the interregional mobility of Italian university students. *Spatial Economic Analysis*, *8*(4), 443–468. doi:10.1080/17421772.2013.833342

European Commission/EACEA/Eurodyce. (2017). *National student fee and support systems in European higher education – 2017/2018. Eurodyce facts and figures*. Luxembourg: Publication Office of the European Union.

Gibbons, S., Neumayer, E., & Perkins, R. (2015). Student satisfaction, league tables and university applications: Evidence from Britain. *Economics of Education Review*, *48*(C), 148–164. doi:10.1016/j.econedurev.2015.07.002

Gordon, J., & Jallade, J. (1996). ‘Spontaneous’ student mobility in the European Union: A statistical survey. *European Journal of Education*, *31*(2), 133–151.

Manski, C. F. (1993). Adolescent econometricians: How do youth infer the returns to schooling? In *Studies of supply and demand* *in higher education* (pp. 43–60). Cambridge, MA: National Bureau of Economic Research (NBER).

Pigini, C., & Staffolani, S. (2016). Beyond participation: Do the cost and quality of higher education shape the enrolment composition? The case of Italy. *Higher Education*, *71*(1), 119–142. doi:10.1007/s10734-015-9892-8

Sá, C., Florax, R., & Rietveld, P. (2004). Determinants of the regional demand for higher education in The Netherlands: A gravity model approach. *Regional Studies*, *38*(4), 375–392. doi:10.1080/03434002000213905

Santos Silva, J. M. C., & Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, *88*(4), 641–658. doi:10.1162/rest.88.4.641

Viesti, G. (Ed). (2016). *Università in declino: Un Rapporto sugli atenei italiani da Nord a Sud*. Rome: Fondazione Res./Donzelli.