

Daniele Checchi  
Tullio Jappelli  
Antonio Uricchio *Editors*

# Teaching, Research and Academic Careers

An Analysis of the Interrelations  
and Impacts

OPEN ACCESS

 Springer

# Teaching, Research and Academic Careers

Daniele Checchi • Tullio Jappelli • Antonio  
Uricchio  
Editors

# Teaching, Research and Academic Careers

An Analysis of the Interrelations and Impacts

 Springer

*Editors*

Daniele Checchi  
Department of Economics  
University of Milan  
Milan, Italy

Tullio Jappelli  
Department of Economics and Statistics  
University of Naples Federico II  
Napoli, Italy

Antonio Uricchio  
ANVUR and University of Bari  
Bari, Italy



This work was supported by ITALIAN NATIONAL AGENCY FOR THE EVALUATION OF UNIVERSITIES AND RESEARCH INSTITUTES (ANVUR).

ISBN 978-3-031-07437-0      ISBN 978-3-031-07438-7 (eBook)  
<https://doi.org/10.1007/978-3-031-07438-7>

© The Editor(s) (if applicable) and The Author(s) 2022. This book is an open access publication. **Open Access** This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license 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.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Contents

## Part I Introduction

<b>Introduction</b> .....	3
Daniele Checchi, Tullio Jappelli, and Antonio Uricchio	

<b>Governance Reforms in Comparative Perspective and Their Path in the Italian Case</b> .....	15
Giliberto Capano	

## Part II Teaching and Students' Careers in Italy

<b>Do Financial Conditions Play a Role in University Dropout? New Evidence from Administrative Data</b> .....	39
Dalit Contini and Roberto Zotti	

<b>Drop-Out Decisions in a Cohort of Italian Universities</b> .....	71
Gianfranco Atzeni, Luca G Deidda, Marco Delogu, and Dimitri Paolini	

## Part III Recruiting and Academic Careers

<b>Is Entering Italian Academia Getting Harder?</b> .....	107
Daniele Checchi and Tindaro Cicero	

<b>Academic Careers and Fertility Decisions</b> .....	135
Maria De Paola, Roberto Nisticò, and Vincenzo Scoppa	

## Part IV Conformism in Research

<b>Social Network Tools for the Evaluation of Individual and Group Scientific Performance</b> .....	165
Domenico De Stefano, Luka Kronegger, Valerio Leone Sciabolazza, Maria Prosperina Vitale, and Susanna Zaccarin	

<b>Topic-Driven Detection and Analysis of Scholarly Data</b> .....	191
Alfio Ferrara, Corinna Ghirelli, Stefano Montanelli, Eugenio Petrovich, Silvia Salini, and Stefano Verzillo	
<b>Part V Research Quality and Impact on Teaching</b>	
<b>The Relationship Between Teaching and Research in the Italian University System</b> .....	225
Maria Rosaria Carillo, Alessandro Sapio, and Tiziana Venittelli	
<b>Degree-Level Determinants of University Student Performance</b> .....	267
Massimiliano Bratti, Giovanni Barbato, Daniele Biancardi, Chiara Conti, and Matteo Turri	
<b>Teaching Efficiency of Italian Universities: A Conditional Frontier Analysis</b> .....	319
Camilla Mastromarco, Pierluigi Toma, and Cinzia Daraio	

**Part I**  
**Introduction**

# Introduction



**Daniele Checchi, Tullio Jappelli, and Antonio Uricchio**

Universities are showing increasing interest in measuring research quality, teaching quality and the relationship between them. Research quality affects individual academic careers and has become important for the efficient allocation of public funding which in many countries and especially countries in Europe is the main component of university financial resources. Teaching quality affects students' careers, and higher quality teaching can reduce dropout rates, improve student performance, and facilitate graduates' transition to the labor market. The quality of the research and teaching in universities has mutual effects, since good quality and effective teaching is often related to good research performance. Mobility can emphasize these dynamics; the best students and the best teachers may be concentrated in a few universities creating potential quality gaps among public universities.

Italy provides an interesting international case study. Rates of tertiary education enrolment in Italy are relatively low and completion rates are even lower, while tuition fees are relatively high compared to those in other European countries. The 2021 edition of *OECD Education at a Glance* (Table C5.1) reports that in the academic year 2019–2020, average tuition fees in Italian public universities were \$2013 for an undergraduate degree and \$2252 for a master's degree compared to \$148 and \$233 respectively for undergraduate courses in Germany and France and

---

D. Checchi (✉)  
University of Milan and INPS, Milano, MI, Italy  
e-mail: [daniele.checchi@gmail.com](mailto:daniele.checchi@gmail.com)

T. Jappelli  
Department of Economics and Statistics, University of Naples Federico II, Napoli, Italy  
e-mail: [tullio.jappelli@unina.it](mailto:tullio.jappelli@unina.it)

A. Uricchio  
ANVUR and University of Bari, Bari, Italy  
e-mail: [antonio.uricchio@anvur.it](mailto:antonio.uricchio@anvur.it)



free tertiary education in all the Nordic countries. The Italian university system is mostly a public system but is characterized by one of the lowest public funding rates in Europe. At the same time, according to several indicators research output (measured as number of journal articles) and research quality (in terms of number of citations) are comparable to most other countries at a similar level of economic development.

In Italy government allocates funds to public universities based on their teaching and research performance, the two main missions of academic institutions. About two-thirds of this funding is proportional to the number of students enrolled in the university (weighted according to disciplinary field), with the remainder allocated based on research output weighted by research quality. This funding mechanism abstracts from possible complementarity or substitutability between the two missions. The interaction between teaching and research and the implications for researcher incentives are the focus of the contributions to this volume.

The book brings together contributions from a range of economists, statisticians, and social scientists involved in an ANVUR-sponsored project.<sup>1</sup> The various chapters analyze different dimensions of research and teaching quality and their interaction, using sound statistical methods allowing comparison with other European countries. The aim is to address the question of whether the evaluation of universities and university departments should focus on both quantitative indicators (such as number of published papers, or number of graduates) and other dimensions of teaching and research, since academic careers, teaching, and students' achievements are strictly intertwined.

The evaluation of teaching and research is addressed also from a gender perspective to try to understand where and when gender discrimination occurs. There is considerable evidence that the glass ceiling is prominent in the Italian academia: women have higher enrolment rates and lower dropout rates relative to men, are represented almost equally at entry to an academic career but despite comparable research productivity are gradually side-lined among higher ranks.

## 1 Which University Model?

The Italian university system adopts the principle of a regulated autonomy (see Chapter "Governance Reforms in Comparative Perspective and Their Path in the Italian Case" by Capano). The Government, the Parliament, and the Ministry of University and Research decide on the allocation of funding and the rules governing its allocation. There are also detailed rules related to the content of the courses that universities can offer (undergraduate, master's, doctoral), access to the states of an academic career (PhD, postdoc, and assistant, associate, and full professor),

---

<sup>1</sup> ANVUR is the Italian agency responsible for the evaluation of teaching and research in academic institutions. See [www.ANVUR.it](http://www.ANVUR.it).

university governance, and department organization. However, universities in Italy can follow different paths.

For example, they might decide to offer courses in most or all disciplines at the undergraduate or graduate level, or choose to specialize in particular subjects and/or levels, and concentrate the resources in particular areas or departments. They might choose to allocate funds internally and reward those researchers and departments that are more productive, or use the funds generated by these departments to invest in the weakest departments. In terms of governance, they might open their boards to external stakeholders, or confine them to mostly incumbent professors. Finally, in terms of organization, they might decide to focus on specialized departments covering specific research areas, or include more and larger departments focused on heterogeneous research areas. Therefore, whether to specialize and give more weight to research than to teaching, or invest in those research areas likely to attract more funding and focus on reputation and international visibility, or offer master's level and doctoral level courses in only a few fields become relevant issues.

A problem particular to Italian academic institutions is their high drop-out rates especially during the first study cycle. This is exacerbated by the fact that for the average student the duration of the course of study can exceed the authorized length by one year or more, depending on the subject. At the national level, only about two-thirds of the student enrollment graduate, and less than half manage to graduate in three years. Drop-out rates and slow careers are particularly prevalent among students from relatively low-income classes. In Chapters “Do Financial Conditions Play a Role in University Dropout? New Evidence from Administrative Data” and “Drop-out Decisions in a Cohort of Italian Universities” Contini et al. and Atzeni et al. show that there is insufficient public support for students in the form of subsidies, services, and scholarships. However, problem goes deeper with the very small fraction of secondary school students from lower-income classes who enroll in tertiary education. Since Italy has one of the lowest rates of college graduates in Europe, one of the objectives of Italy's university system should be to improve the country's human capital<sup>2</sup> by increasing enrolment after secondary school and reducing dropouts during the course of study.

From a policy perspective there are at least three options. One would be to focus on creating a few national university “champions” with the remaining universities offering mainly undergraduate education. A second option would be to pursue a specialization model which would mean that each university would aim for international parity in the areas of its comparative advantages. The third and more traditional option would be to try to maintain a more balanced tertiary education system involving all universities offering master's level and doctoral level courses in all fields. These options have different implications for academic careers which do not distinguish between research and teaching positions and apply the same standards to all at both entry level and promotion.

---

<sup>2</sup> According to Eurostat, in Italy tertiary education attainment in the population aged 25–54 was 22.6% in 2020, while it was 32.7% in Germany, 44.3% in France, and 43.3% in Spain.

The choice between specialization and a universal model depends on the relation between teaching quality and research quality. In Chapter “The relationship Between Teaching and Research in the Italian University System” Carillo et al. show the complementarity between research and teaching; that is they find that good researchers are also good teachers which would imply that specialization is a sub-optimal solution.

Past contributions by Bratti and coauthors suggest that students “vote with their feet”, and that the best students enroll in higher-ranked universities. In addition, paucity of high-quality study courses results in student mobility to other countries, and in domestic universities attracting only in lower quality national and international students which has long term consequence for human capital formation and growth. Novel results by Bratti et al. in Chapter “Degree-Level Determinants of University Student Performance” suggest that higher education institutions play an important role in ensuring the academic success of their students. Indeed, several degree-level characteristics significantly predict students’ progression and satisfaction with their university education.

Since research and teaching go hand-to-hand, it is important to offer economic incentives and career prospects for young researchers and teachers in particular. Checchi and Cicero (Chapter “Is Entering Italian Academia Getting Harder?”), and De Paola et al. (Chapter “Academic Careers and Fertility Decisions”) show that in the recent past the ability to do this has been limited severely by budget cuts.

## 2 Incentives

In the case of incentives while in principle the regulatory framework allows funding to be channeled to the best universities, university departments, and individual researchers, in practice this often does not happen.

At the university and department levels, incentives are designed based on the evaluation of their research output. Ferrara et al. (Chapter “Topic-Driven Detection and Analysis of Scholarly Data”), and De Stefano et al. (Chapter “Social Network Tools to Evaluate Individual and Group Scientific Performance”) points out mainstreaming and adapting to the rules of the game can occur after the results of three evaluation exercises, and particularly in fields involving research conducted by large research teams rather than individuals.

Each university might implement local incentives for periodic salary increase for the most productive researchers, and might also mobilize internal and external funds to incentivize research, teaching, external finance, and other activities which would allow increased compensation for the most active researchers. However, in practice, these types of rewards are not relevant for differentiating among academic salaries within universities, and most academics boost their income by engaging in consultancy and/or professional activities (lawyers, clinic doctors, architects, etc.).

### 3 Evaluating University Performance

Several of the chapters in this volume are methodological contributions. Mastro-marco et al. in Chapter “Teaching Efficiency of Italian Universities: A Conditional Frontier Analysis” suggest that resources must be distributed efficiently across fields of study and geographical areas which requires measurement of the extent to which current allocations are efficient. Their analysis considers only the distribution of human resources (teachers) but could be extended to integrate other inputs such as infrastructures and staff. Ferrara et al. (Chapter “Topic-Driven Detection and Analysis of Scholarly Data”) propose a framework that could be used to benchmark the research produced by particular universities or particular themes against international research. Also, ANVUR, which sponsored the research for this book, has published several indicators of the impact of teaching on students’ academic careers.

### 4 A Tour of the Book

The book includes five sections and eleven chapters. This introduction by the editors and a chapter by Capano on “Governance reforms in comparative perspective” comprise the first Chapter. Capano discusses models of university governance in a European framework, and whether the Italian model of steering at a distance is consistent with university autonomy. Lack of guidance about prioritization of research, teaching, and knowledge transfer limits the ability of individual institutions to identify clear strategies to improve their performance. Analysis of the reforms the Italian higher education system implemented since 1990 should help the reader to contextualize the dynamics of the institutional and policy arrangements within which research, teaching, and an academic profession have developed.

Section 2 discusses evidence based on administrative data related to the determinants of career completion by university students. Italy is an interesting case study due to the relevance of students’ initial socio-economic conditions for academic achievement which is underlined by studies based on the OECD’s PISA (Programme for International Student Assessment) scores. Obtaining evidence about how family background affects a university career is difficult since students tend to be sorted into academic and vocational tracks at the secondary level. Females exhibit higher completion rates. Experiments have been conducted in local Italian universities to study the causal impact on drop-out rates of introducing extra-tutoring.

In Chapter “Do Financial Conditions Play a Role in University Dropout? New Evidence from Administrative Data” Contini and Zotti discuss the role played by economic conditions on student university careers in Italy. They use administrative data from the University of Turin – a large public institution in the North of Italy – and information on family background collected at matriculation to analyze how

family economic conditions influence the probability of first-year dropout from university. While parents' education and parents' occupations have been shown to have a major effect on education outcomes for school-age children it seems that they do not have a sizable effect on university student drop-out. Instead, there is evidence that despite the progressive character of tuition fees and the existence of scholarships provided to low-income students, economic conditions do have a substantial impact on the likelihood of completing university studies. This suggests that current student aid policies in Italy are insufficient to close the gap that exists between high- and low-income students, and that increasing financial aid could be a tool for promoting equality of opportunity in education and eventually increase the share of young individuals with higher education degrees.

In Chapter "Drop-out Decisions in a Cohort of Italian Universities", Atzeni et al. study the determinants of students' drop-out decisions using data on a cohort of over 230,000 students enrolled in the Italian university system. The empirical analysis controls for course-of-study and university fixed effects, and shows that the probability of dropping out of university is correlated negatively with high-school grades and student age. However, it shows also that women have a lower propensity for drop-out especially among students enrolled on science, technology, engineering, and math where they are under-represented. Atzeni et al.'s data differentiate between students who leave home to enroll at university (off-site students) and students who continue to live in the family home (on-site students). They find that drop-out is significantly lower among off-site students. Self-selection into studying off-site is estimated using an instrumental variable approach to identify the causal relationship. The authors use detailed administrative data on students enrolled at the University of Sassari to investigate another self-selection channel affecting the estimation of the determinants of drop-out. They employ bivariate probit estimation to account for self-selection into the course of study, and show that estimates of the traditional determinants are modified. The unconditional comparison among degrees is misleading since some degrees attract more heterogeneously skilled and motivated students. While the estimation without selection suggests that women's dropout rates are lower after accounting for selection, the contribution of women to the drop-out rate turns either positive or negative depending on the chosen study course.

Section 3 focuses on the increasing precariousness of an academic career especially for younger researchers. The two chapters in this section exploit longitudinal administrative data. They show that the standard transitions (PhD-postdoc-assistant professors-(tenured)associate professor) are discipline specific but also are gendered since job instability has different costs as women and men age (reflected in fertility decisions, conditional on obtaining tenure).

In Chapter "Is Enter to Italian Academia Getting Harder?" Checchi and Cicero consider the traditional steps in an academic career. While a doctoral degree is often considered the first necessary step only a small fraction of doctoral graduates (less than 10%) obtain an academic position within 6 years of degree award. Despite the absence of information on labor market outcomes, the authors focus on the determinants of this transition in order to study whether entry into an academic

job is becoming more selective and/or more precarious. Merging three national administrative datasets on completed PhD degrees, postdoc collaborations, and new hiring into the academia they find a decline in appointment probability after the 2010 cohort, due to the effect of the hiring freeze imposed by fiscal austerity. They find also that combining a doctoral degree and postdoc experience increase successful application to academia. Women and foreign-born candidates are shown to be negatively discriminated, and there is evidence of career disadvantages for candidates from Southern universities.

In Chapter “Academic Careers and Fertility Decisions” De Paola et al. investigate how academic promotions affect the propensity of women academics to have a child. They use 2001–2018 administrative data on female assistant professors employed in Italian universities and estimate a model with individual fixed effects. They find that promotion to associate professor (a tenured position) increases the probability of having a child by 0.6 percentage points, which translates into a 12.5% increase at the mean. Results point in the same direction using a regression discontinuity design that exploits the eligibility requirements in terms of research productivity introduced in 2012 by the Italian National Scientific Qualification (NSQ) related to promotion to associate professor. Their study has important implications for policy by showing that reducing career uncertainty leads to increased fertility among academics.

Section 4 deals with methods designed to assess research productivity at a time when co-authorship and team production are becoming standard practice. Co-authorship and team working complicate the assessment of research quality, and the individual contribution of a research project and research output. The increased pressure to publish may induce the risk of excessive conformism in the choice of topics which can be mapped using text analysis. Gender issues may also matter since co-authorship, research networks, and research impact might not be gender neutral.

Chapter “Social Network Tools to Evaluate Individual and Group Scientific Performance” by De Stefano et al. analyzes patterns of scientific collaboration which recently has been considered an important driver of research innovation. Collaboration allows scientists to benefit from methodological and technological complementarities and synergies which can improve the quality and quantity of their research output. For example, collaboration among scientists has been shown to be increasing in all disciplines, and the rules governing international exchange programs are aimed at promoting collaboration among researchers. Collaboration among scientists can be mapped into networks and co-authorship linkages which makes social network analysis a useful theoretical and methodological approach. Several empirical studies identify a positive association between the researcher’s position in the co-authorship network and the individual researcher’s productivity, although the results differ depending on the discipline, scientific performance measure, and the data source used to construct the co-authorship network. De Stefano et al. propose the use of social SNA tools for scientific evaluation purposes. Network indices at the individual and subgroup levels are introduced to analyze the relation with both the individual research productivity and scientific output quality

measures provided by the Italian academic researchers involved in the latest national evaluation of research quality (2011–2014).

In Chapter “Topic-Driven Detection and Analysis of Scholarly Data” Ferrara et al. present a mining approach to identify academic research topics based on the idea that research topics emerge through analysis of epistemological aspects of academic publications extracted from conventional publication metadata such as title, author-assigned keywords and abstract. The authors provide a conceptual analysis of research-topic profiling according to the behaviors/trends peculiar to a given topic along a considered time interval. They define a disciplined approach and related topic mining techniques based on the use of publication metadata and natural language processing tools. This approach can be applied to various topic analysis issues such as country-oriented and/or field-oriented research analysis based on scholarly publications. To assess the effectiveness of these techniques when applied to a real situation, the authors conduct a case-study analysis based on national and international data.

Section 5 synthesizes the discussions in Sects. 3 and 4 on student achievement and teacher careers driven by research assessment. By exploiting quality measures derived from research assessment exercises conducted every 5 years by ANVUR, Chapters “The Relationship Between Teaching and Research in the Italian University System”, “Degree-Level Determinants of University Student Performance”, and “Teaching Efficiency of Italian Universities: A Conditional Frontier Analysis” focus on the conditional correlation between student mobility and academic career completion. The relationship between these dimensions can be studied in terms of the joint production of academic services to enable the evaluation of university efficiency using frontier analysis.

In Chapter “The Relationship Between Teaching and Research in the Italian University System” Carillo et al. study the relationship between the quality of research and teaching in the Italian university system at the study program level. The authors use detailed data collected by the Italian ANVUR on undergraduate and master’s degrees offered by Italian universities in the academic year 2016–2017. Their cross-sectional econometric analysis shows a positive relationship between teaching quality and research performance that emerges when taking account of yardstick competition among study programs offered by the same department. The theory suggests that despite the trade-off between teaching and research faced by individual academics, in multi-unit universities which have implemented budget sharing based on research performance and the number of students, the negative relation between teaching and research is reduced or is counterbalanced. However, in the case of universities offering only a small number of study programs the teaching-research relationship is positive and stronger. The results are even more pronounced for master-level degrees where teaching is more aligned to individual research interests.

In Chapter “Degree-Level Determinants of University Student Performance” Bratti et al. use administrative data on higher education degrees in Italy during 2013–2018 to analyze the degree-level determinants of university student performance as measured by ANVUR quality indicators. After controlling for detailed

degree subject–geographic macro area fixed effects, their analysis reveals several significant predictors of degree quality including access (i.e., selectivity), language of instruction, teaching body composition, percentage of teachers of core subjects, teachers' research performance (for second-level degrees), and spatial competition.

The last chapter on “Teaching Efficiency in Italian Universities: A Conditional Frontier Analysis” by Mastromarco and al. presents a comparative analysis of the performance of Italian university teaching by evaluating the efficiency of heterogeneous faculty courses at the national level. According to OECD data, Italian public universities are under-funded: the costs related to individual Italian students are well below the OECD average. This underlines the importance for policymakers of information on the relative efficiency of universities which can be used as an indirect evaluation of how public funding is used. Chapter “Teaching Efficiency of Italian Universities: A Conditional Frontier Analysis” uses tools developed recently in the nonparametric efficiency frontier literature. The analysis is conducted at the national level and extends traditional analyses based on mono-dimensional indicators. The efficiency scores enabled by the statistical analysis are used to interpret current trends and changes to Italian universities' teaching activities.

## 5 The Road Ahead

There are many other relevant topics that are not addressed in this volume. Public engagement and knowledge transfer are another university mission but information on these activities is scattered. However, it is being collected for the research evaluation that is currently underway. We also do not discuss the potential effects on teaching and research of the Italian National Recovery and Resilience Plan approved in July 2021 which assigns €5.4 billion to postsecondary education, slightly more than one-quarter of the total intended education budget. This plan is aimed at increasing tertiary education graduation rates, strengthening vocational education, and removing the financial obstacles to university enrolment.

The plan included a measure to raise the transition from upper secondary to higher education and to reduce university drop-out by providing more information on university careers, since the children of less-educated parents are more likely to lack confidence and knowledge about academic courses and careers. This measure is expected to increase school attendance, improve learning levels, increase university enrollment, and reduce the gender gap in university employment and participation in higher education in all fields.

At the same time, financial constraints and/or labor market opportunities also matter. For this reason, half a billion euros have been allocated to student scholarships and tuition exemptions. While empirical evidence that low graduation rates are caused by a lack of public financial support is limited, these measures will make Italy better aligned to other European countries and should promote mobility across universities in Europe. A further €1 billion has been allocated to student accommodation in Europe (to enhance student mobility) in a partnership with



the private sector.<sup>3</sup> Following the German model, the higher education system is expected to be extended to include vocationally oriented, non-academic tertiary education based on a planned investment of €1.5 billion.

To strengthen university autonomy in program design, and to increase the competition among universities in vocational training, several measures are being suggested to update study course curricula, create new cross-disciplinary programs, expand vocational training programs, and limiting the role of professional associations in the transition of graduates to the labor market.

The number of doctoral scholarships will increase from 10,000 to around 17,500 to try to increase the stock of human capital, with positive spillovers for innovation and R&D activities through partnerships with private companies and research centers, and reduce the doctoral graduation gap with European partners. Firms will be given incentives to recruit temporary junior researchers (20,000 over 3 years), and to establish research hubs and promote spin-off activity. To retain new doctoral graduates and avoid their migration to industry, additional funding will be provided for research programs led by young researchers. These incentives will be subject to gender quotas to encourage greater participation of women.

In summary, Italian policymakers plan to enhance the higher education system by promoting student mobility, providing new job opportunities for young researchers, and including new vocational programs. The former two objectives are discussed in this volume. It is hoped that these contributions will help aid to evaluating the success of the plan and the more efficient use of public resources.

**Daniele Checchi** (Ph.D. University of Siena) is Professor of Economics at the University of Milan and a former member of the Board of Directors of ANVUR, currently on leave at the Research Department of INPS (Italian Social Security Administration). His research interests are in the area of the economics of education, welfare policies, economics of inequality, research evaluation.

**Tullio Jappelli** (Ph.D. Boston College) is Professor of Economics at the University of Naples Federico II. His research interests are in the area of saving, household finance, and applied macroeconomics.

**Antonio Uricchio** is Professor of Tax Law at the University of Bari and President of ANVUR, the Italian national agency for the evaluation of universities and research institutes. His research is in the area of tax law, environmental policies and technological innovation, and public finance.

---

<sup>3</sup> The aim is to increase current available accommodation from 40,000 to over 100,000 by 2026 to reduce the gap between Italy and the EU average for share of students provided with residential facilities (18% against 3% currently for Italy).

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter'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.



# Governance Reforms in Comparative Perspective and Their Path in the Italian Case



Giliberto Capano

**Abstract** Reforming governance in higher education has been a kind of mantra that has characterised governmental policies worldwide. Under the pressure of massification, globalisation and socio-economic demands, governments have continuously intervened to redesign the characteristics of the governance arrangements of their higher education systems as well as institutional governance. This common effort has been characterised by the adoption of a common template (i.e. the ‘steering at a distance’ model), mainly based on the idea of making universities more accountable to the societal goals through the massive use of evaluation, assessment and monitoring. The final results are highly differentiated, owing to the fact that each country has implemented a common template according to its own national characteristics and legacies. In this context, the Italian case shows its own peculiarities, whereas evaluative tools have been significantly adopted in a design highly contradictory of other dimensions such as institutional governance, the rules of careers and academic recruitment and the lack of clear systemic goals to be reached.

**Keywords** Governance reforms · Hybrid governance · Italy · Recovery Resilience Plan · Policy Instruments

## 1 Introduction

Over the last three decades, governments have consistently intervened in higher education (HE). Additionally, significant changes have occurred in inherited national governance modes. In continental Europe, these governmental policies have attempted to abandon the inherited continental governance mode, which is characterised by hierarchical coordination through state-centred policies, a lack

---

G. Capano (✉)  
University of Bologna, Bologna, BO, Italy  
e-mail: [giliberto.capano@unibo.it](mailto:giliberto.capano@unibo.it)

© The Author(s) 2022  
D. Checchi et al. (eds.), *Teaching, Research and Academic Careers*,  
[https://doi.org/10.1007/978-3-031-07438-7\\_2](https://doi.org/10.1007/978-3-031-07438-7_2)

of institutional autonomy, the powerful and all-pervasive authority of academic guilds, and faculties and schools as ‘confederations of chair-holders’ (Clark, 1983), in favour of the model adopted in English-speaking countries. These reforms have been characterised as ‘autonomistic’ because universities have been granted more institutional autonomy at various levels and intensities. However, institutional autonomy does not stand alone. The other side of this phenomenon has been the changing role of governments in leading their HE systems and their university systems in particular. Governments have drastically reduced the use of the traditional direct command and control strategies in favour of leading from a distance based on national standards, procedures for monitoring and evaluation, criteria for financial rewards and changing internal institutional governance arrangements (Lazzaretti & Tavoletti, 2006; Huisman, 2009; Paradeise et al., 2009; Enders et al., 2013; Capano & Jarvis, 2020). In contrast, in the Anglo-Saxon world, governments have increased their intervention and regulation despite a historical tradition promoting institutional autonomy for universities (El-Khawas, 2002; McLendon & Hearn, 2009; Schuetze et al., 2012; Jones, 2012). It has been a long process through which some historically rooted characteristics of systemic and institutional governance have been significantly modified.

This chapter is committed to sketching out the general picture of these reforms to help readers of this book contextualise the Italian case and, consequently, the evolutions of the institutional and policy arrangements in which research, teaching and the academic profession have developed.

## **2 The Structural Problem in Governing Universities and the Old Governance Solutions**

The governance problems in higher education are twofold: one concerns the institutional dimension (i.e. how an individual university is coordinated and produces its own policies), while the other concerns the systemic dimension (i.e. how national higher education policy is designed and implemented).

Universities are *sui generis* institutions, whose constitutive nature (i.e. the fact they are federations or confederations of academic subjects and niches) has structural implications for their internal dynamics; this creates never-ending problems for institutional governance. Universities bring together groups of individuals doing very different jobs (e.g. the job of a biologist compared to that of a historian, or the job of a computer technician compared to that of a help-desk employee), many intertwined decision-making processes, and a great variety of institutional outputs that range from basic to applied research and PhD programmes to continuing education courses, etc. There is an inescapable organisational and functional complexity in universities; in order to grasp this complexity, some scholars have proposed terms such as ‘multiversity’ (Kerr, 1963) or the ‘federal or conglomerate form of organisation’ (Clark, 1995).

Because of such features, universities are considered a typical loose-coupling organisation or a form of organised anarchy. From this point of view, universities as loose-coupling institutions are characterised (Orton & Weick, 1990) by:

- Causal indeterminacy
- A fragmented external environment
- A fragmented internal environment

Causal indeterminacy means that the actions of universities are characterised by the intrinsic ambiguity and uncertainty of means-ends relations and by a contradictory variety of goals. For empirical evidence of this point, one only has to read the statutes of certain universities or the decisions taken by their collegial governing bodies in order to see how linear rationality and causality do not really apply to higher education institutions. Universities see themselves as pursuing excellence in research, the freedom of teaching, the socio-economic development of their society, equity and accountability; however, at the same time, they are subdivided into a variety of different niches and academic disciplines, each with its own mission, epistemological basis and professional rules. In such a context, causality is very often the result of chance or serendipity.

A fragmented external environment simply means that a large number of external stakeholders continuously demand several contradictory things from universities, such as local economic development, technological applications, the increased quality of the stock of human capital, its selection and education, social and political elites, social mobility, etc. This means that the expectations of the external environment may be incompatible with those of the universities themselves.

A fragmented internal environment simply refers to the constitutive variety of universities' internal components. They are composed of different academic 'tribes' that constantly seek to defend their own territory (Becher, 1989), by various groups of students' demanding very different services, and by the non-academic staff. At the same time, there is a variety of institutional levels and structures within the universities. Collegial governing bodies, faculties, departments, committees, research centres and institutes: universities are overcrowded with nested institutional arenas. This internal fragmentation is self-reproducing, self-sustaining and in accordance with a self-referential rationality.

Universities as loose-coupling organisations complicate their institutional coordination, that is, their internal governance, while at the same time explaining their ability to adapt and survive. For example, internal fragmentation enables them to register a very large range of external inputs and demands and subsequently to offer a variety of responses: this is an essential resource for institutional adaptation to external challenges. Furthermore, their loose-coupling nature provides universities with the power to buffer (i.e. to lower or to isolate) disturbances from the external world. Their buffering capacity also explains the intrinsic feature of the institutional development of universities—they are capable of change, but only by adapting to external changes. This institutional change is based on what Schon (1971) called 'dynamic conservatorism'.

It should be noted that even if they are loose-coupling organisations, universities nevertheless possess a number of internal tightening-up mechanisms (Lutz 1982). In fact, they are also bureaucratic organisations with a plethora of official internal regulations that need to be observed in order to pursue the institutional mission (for instance, time schedules for classes, rules on the recruitment of professors, rules on institutional government, etc.). This means that there are rules and practices designed to reduce the anarchic, ambiguous trend triggered by loose-coupling elements. What is now evident is the day-to-day battle between the looseness and the tightness of the institutional working and proper functioning of the university.

Thus, the governance quandary in higher education is, above all, represented by the intractable problem of how to coordinate a specific institution, the university, which is intrinsically fragmented and composed of a variety of loosely connected groups and interests, and to render it accountable and responsible—both at the institutional and the systemic level. In basic terms, the governance problem consists of getting universities to behave as ‘institutions’ and ensuring that the higher education system as a whole effectively responds to the needs of society. The three levels (infra-institutional, institutional and systemic) are strictly interconnected: each is the other face of the others.

If one examines the development of universities in the Western world over the past two centuries, one sees that the governance problem has been resolved in a variety of different ways and according to the specific national context in question. We should not forget that universities do not exist in a vacuum; they are deeply rooted within a specific economic, cultural and socio-political system. Several attempts have been made to classify governance within higher education in order to take account of this structural differentiation underlying the idiosyncratic character of higher education. The best-known attempt of such nature resulted in Clark’s triangle (1983), which consists of the interaction of three mechanisms of systemic and institutional coordination: the state, the market and the academic oligarchy. Clark proposed three ideal types of higher educational governance: the Continental, American and British types.

The Continental model’s constitutive elements are as follows: systemic, strongly hierarchical coordination through state-centred policies; no institutional autonomy; the powerful, all-pervasive authority of the academic guilds; faculties and schools constituting ‘confederations of chair-holders’. The British model, on the other hand, is characterised by substantial institutional autonomy, collegial academic predominance, and the moderate role of the state. Finally, the American model consists of the strong procedural autonomy of universities, which is counter-balanced by the substantial public monitoring of the quality of performance and results<sup>1</sup>; the important role of external stakeholders (which also means the significant role of

---

<sup>1</sup> The important influence exercised by U.S. governments at both the federal and state levels and on the institutional behaviour of universities is too often underestimated. Federal government plays a crucial role because of its earmarking of huge amounts of funds for research and for student aid programmes: federal government has used its financial weight to profoundly influence both public and private universities (especially those particularly committed to high-quality research). State

public political institutions in the case of public universities); academics' weaker role in determining universities' strategic objectives, which is counterbalanced, in accordance with the principle of 'shared governance', by their more substantial powers in relation to traditional academic matters (e.g. staff recruitment, course content, etc.).

### **3 The Challenge of Massification and Modernisation as Drivers for Radical Changes in Governing Higher Education Systems**

However, the historically rooted models of governance in Western countries, masterfully represented by Clark's ideal types, have had their limitations exposed when faced with modern-day challenges. Each inherited governance equilibrium has been obliged to change. In the past, universities were never subjected to such similar pressure to dramatically change their own hundred-year-old governance practices and equilibrium.

So the question is: what caused this tremendous and unexpected pressure to change? The answer is simple: societies and governments have started to take great interest in higher education because, within a global context of strong competition, the quality of human capital needs to be continuously improved, and new technological solutions have to be found in order to support economic development. Society and governments have started to demand increasingly more from the higher education system. Some examples are provided below:

- A rapid increase in participation rates intending to transform an elite system to a mass system and universal education, which Martin Trow theorised about more than 40 years ago (1974)
- The increased diversification of educational demands (general education, specialised education, life-long learning, distance learning courses, internationalisation of courses AND research training)
- The development of training and technology for local communities
- Education designed to spur economic development

Almost paradoxically, these new demands have arisen at a time when public funding is increasingly being cut due to the fiscal crises of the state. Public funding is of fundamental importance for all higher education systems (with the partial exception of the USA). Higher education institutions were thus strongly asked to do more than they had in the past and at a quicker rate, notwithstanding the continued reductions in public funding. Moreover, universities are suddenly being asked to be accountable. Unlike in the past, universities are now asked to report on their

---

governments play a crucial role, since they are both the 'owners' and the 'regulators' of public universities (Berdahl, 1999).

use of both public and private resources and on the results of their utilisation. Universities must be accountable for financial and physical resources; the quality of teaching innovations; student recruitment; faculty appointments; research resources, productivity, and knowledge transfer; rigour in management and quality assurance; and the well-being of students and staff.

It is this tremendous external pressure that has definitively brought down the walls of the 'ivory towers'. One of the inevitable consequences of this new trend has been the structural pressure to change the inherited and historically rooted governance arrangements.

It is no coincidence that Clark's basic assumptions have been further developed by other scholars trying to adjust the theoretical definition of governance in higher education to real changes. For example, Van Vught (1989) proposed two possible governance models: the *state control model* and the *state supervising model*. The first, which is characteristic of the continental European tradition, sees the state regulate the procedural aspects, and often the content, of student access, the recruitment and selection of academic staff, the examination system, degree requirements, the content of curricula, etc. At the same time, academics maintain considerable power over the internal life of universities. In this model, universities are weak institutions because the important power relationships are those connecting the local academic guild to the central bureaucracies. The state supervising model is characteristic of the English-speaking world, where universities are stronger and are usually governed on the basis of academics and internal management sharing governance), and the state plays a subtler role, *steering at a distance*. Other types designed to encapsulate the features of other forms of higher education governance have also been proposed (see, for example, Becher & Kogan, 1992; Braun & Merrien, 1999). In all of the aforementioned cases, the state plays an important role.

## 4 The Long March of Higher Education Reforms

New challenges have called for a radical re-thinking of governance models at the institutional and systemic levels; this, in turn, highlights the need to redesign not only the formal rules at both the institutional and systemic levels by changing the distribution of powers and responsibilities, but also the governance arrangements (i.e. the way in which decisions and policies are made, implemented and coordinated). Hence, this is not only a case of institutional reform but above all a case of policy change.

Generally speaking, the basic levers of reforms can be summarised as follows (see Amaral et al., 2002; Enders & Oliver Fulton, 2002; Gornitzka et al., 2005; Lazzaretti & Tavoletti, 2006; Cheps, 2006; Maassen & Olsen, 2007; Trakman, 2008; Huisman, 2009; Paradeise et al., 2009; Shattock, 2014; Capano et al., 2016; Capano & Pritoni, 2020a, 2020b; Capano & Jarvis, 2020): institutional autonomy, funding mechanisms, the quality assessment of research and teaching, internal institutional



governance and the changing role of the State. At the same time, it should be pointed out that governments had, and continue to have, a predominant role in the reform of governance in higher education. This is also the case for public universities in the USA, where state governments have been very active (McLendon, 2003a, 2003b; Leslie & Novak, 2003; El-Khawas, 2002).

The above basic levers have been moulded differently at the national level, although some common features have emerged:

- In European countries, governments have abandoned the state-control model in favour of steering universities from a distance by giving more autonomy to these institutions. In some countries, such as the Netherlands, Sweden, Denmark, Finland and Austria, governments have radically changed the institutional arrangements of universities by abandoning the traditional democratic mechanisms to elect the institutional leaders and the governing body for an appointment system. This has also been adopted in non-European countries such as Japan and China. The supervisory role of the state (Neave & Van Vught, 1991) is implemented by steering on the basis of new 'soft' methods of coordination no longer based on hard rules but on soft contracts, targets, benchmarks, indicators and continual assessment.
- In the English-speaking world, governments have increased their intervention and regulation, despite a tradition of institutional autonomy. In the UK, Australia and New Zealand, governments have substantially restructured the national governance framework by creating national agencies for the assessment of research and teaching, and through a strong commitment to realigning the behaviour of universities to socio-economic requirements. At the same time, public universities in the USA have been strongly encouraged to adopt a more competitive stance in order to obtain more funding from private sources, in a substantial process of marketisation (Capano & Jarvis, 2020).

Within this context of the substantial re-design of the borders and the general framework of higher education's systemic coordination, certain other features are present in all of the most important countries, with the partial exception of the USA (because of the intrinsic difficulty in defining the incredible variety of American higher education institutions as a system):

- Institutional autonomy does not mean 'independence' or 'academic freedom'; rather, it means the capability and right of a higher education institution to determine its own courses of action without undue interference from the state, but within a context that is strongly influenced by the same state. In this sense, the common interpretation of institutional autonomy is that of a policy instrument designed to increase the effectiveness of higher education policies; so what clearly emerges is that in those countries belonging to the Continental mode, where institutional autonomy was either weak or non-existent, governments have started to grant greater institutional autonomy; on the other hand, in those systems where university institutions have traditionally been very autonomous (e.g. in the English-speaking world), governments have started to

interfere in institutional behaviour through the introduction of new regulations, the assignment of targets, pressure for more inter-institutional competition, and so on.

- Funding traditionally earmarked for the functioning of universities has been abandoned in favour first of lump-sum grants.
- The entity of public funds assigned to universities is based on output-oriented criteria and performance-based contracting systems.
- There exists strong pressure to increase private funding (by increasing tuition fees and by selling services and research to private actors).
- National agencies or committees for the evaluation and assessment of the quality and performance of teaching and research in higher education institutions have been established in all Western countries (with the exception of the USA and Canada).

At the institutional level, under the pressure of governmental policies, a common trend has emerged even in those countries where pre-existing institutional-governmental structures have not changed or are changing very slowly, as in Italy (Capano, 2008), Spain (Mora & Vidal, 2005), France (Mignot Gerard, 2003) and Germany (Kehm and Lansendorf 2006): environmental pressure from society, governments, economic requirements, etc. shift the balance of power and authority within universities. The centralisation of institutional authority has grown steadily over the years. This implies the following:

- The strengthening of the role of individual leaders (presidents, rectors, vice-chancellors, deans)
- The reinforcement of the role of central administration and management
- The strengthening of the power of governing boards both in the English-speaking world and in the reformed European systems
- A decline in the influence of academics and academic guilds on institutional decision-making (the said guild often conflictS with and resistS centralisation trends). This creates a structural risk of stalemate in the internal decision-making process
- The introduction of new management tools, such as strategic plans, budgeting and financial management, internal audits and quality assessment systems
- Play an increasingly important role in governance

## **5 The Hybridity of New Systemic Governance in Higher Education: Same Instruments but Different Policy Mixes**

What clearly emerges in the comparative picture sketched above is that the forms of governance within higher education policy are changing radically: the question is, how are they changing? If one examines the plethora of comparative studies of governance shifts in higher education that have been produced over the last

30 years, it is clear that at the systemic level, the governance models of the past have been clearly abandoned in favour of a new template, the steering at a distance model that, however, has been adopted in different ways according to the context and the national traditions. This variety has justified different and sometimes radically divergent assessments of these reforms. For example, there are studies that underscore how, in recent years, there has been a strong re-regulation of the field in many countries (Enders et al., 2013; Donina et al., 2015). Other scholars consider governance reforms in higher education a product of the neoliberal age and thus emphasise the predominance of privatisation, deregulation, managerialisation and the limitation of academic freedom (Marginson, 2009; Olssen & Peters, 2005; Harvey, 2005). These positions are slightly extreme in assessing reality and very often consider only some dimension of the adopted governmental policies. It is not the case that recent research that compares many European countries has shown very differentiated results in terms of existing systemic governance arrangements, and that every country has adopted its national interpretation of the steering at a distance model by mixing evaluative, information and regulatory tools (Capano & Pritoni, 2020b). This variety can be ordered by focusing on the instrumental composition of the governmental policies adopted over time in the last decades. By following Capano and Pritoni (2019), this kind of instrumental perspective leads to the extraction of three different hybrid types through which the steering at a distance model has been implemented from a comparative perspective: the performance-oriented mode, the re-regulated mode, and the goals-oriented mode. Table 1 presents these three types of hybrid steering at a distance mode.

The performance-oriented mode focuses on performance, which means that a significant part of public funding is based on the assessment of teaching and research. Someone might expect this mode to be the most diffused hybrid due to all the rhetoric about evaluation that characterises the public discourse on evaluation worldwide), but this expectation does not correspond to the empirical evidence. In fact, it appears that among the European countries, only England and parts of Italy fit this hybrid (Capano & Pritoni, 2019). The peculiarity of this hybrid circumscribes it to these few cases; it does not appear that other systems in the Americas (perhaps except Brazil) and in Asia have really emphasised performance as the pillar criterion for governing their HEs (clearly, with the exception of New Zealand, which has been the pioneer in shifting towards a performance-oriented hybrid since the 1980s) (Capano and Jarvis, 2020).

The re-regulated mode is characterised by a strong proceduralisation imposed by governments, a relevant presence of target and performance funding and the tendency to not increase tuition fees. In this hybrid, evaluative practices are procedural and push more for compliance than for performance. This hybrid is adopted by governments that cannot invest too much in higher education and that try to steer their HEs by mixing common procedural rules and different types of evaluation and quality assurance. Additionally, this hybrid appears to be the one with more potential diffusion worldwide (especially in countries with a legacy of bureaucratic systemic governance in higher education). Regarding Western countries, it looks at

**Table 1** Types of hybrid systemic governance modes in higher education

Types of “steering at a distance”	Main and leading instruments <i>Different mixes of Regulation, Expenditure, Taxation, and Information + public funding + tuition fees</i>
<i>Performance-oriented mode</i>	<ul style="list-style-type: none"> <li>• Significant percentage of public funding; based on the results of research assessment</li> <li>• Use of information tools</li> <li>• Many regulations of administrative procedures</li> <li>• Significant percentage of public funding based on evaluation of teaching performance</li> <li>• Student support based on loans</li> <li>• Relatively high tuition fees</li> </ul>
<i>Re-regulated mode</i>	<ul style="list-style-type: none"> <li>• Many procedural constraints on the main activities (recruitment, promotion, postdoc, teaching content and organisation of degrees, student admissions)</li> <li>• Proceduralisation of quality assurance</li> <li>• Target funding and performance funding</li> <li>• Average/low public funding</li> <li>• Low tuition fees</li> </ul>
<i>Goal-oriented mode</i>	<ul style="list-style-type: none"> <li>• Clear systemic goals stated by governments</li> <li>• Many opportunities in admissions, curricula, and institutional autonomy</li> <li>• High public funding</li> <li>• Information instruments (monitoring, reporting)</li> <li>• Strategic use of target/performance funding</li> <li>• Student support based mostly on grants</li> <li>• High performance and target funding</li> <li>• No/low tuition fees</li> </ul>

the prevailing mode in Austria, Ireland, France, Greece, Portugal, Italy (partially) and the Netherlands (Capano & Pritoni, 2019).

The goal-oriented hybrid is foremost characterised by the presence of clear goals stated by governments that then design their systemic steering by mixing high public funding, a strategic use of evaluation and enormous student support. This hybrid is likely to be another European peculiarity since it is present in the Nordic European countries, which are the motherland of the broad welfare state. However, what makes the difference here is the strong capacity of the government in designing clear systemic goals that the institutions are asked to contribute towards achieving.

These three types of hybrid governance can be a useful point of departure for further research and for analysing systemic governance from a comparative perspective. Overall, for example, many Asian governments (e.g. China, Japan, Malaysia) seem to have been steering their HEs through a re-regulatory approach, while others such as Singapore and Hong Kong have been doing so through a goal-oriented approach. It would also be interesting to apply this framework to Latin America and to the states and provinces of the USA and Canada, respectively. For example, Quebec has clearly adopted a re-regulative mode, while in most of

the other provinces, the goal-oriented hybrid appears to prevail, although with the substantial difference that many of them have increased their tuition fees.

Clearly, the three hybrids could be biased because they are ‘continental’-specific and, thus, they cannot be considered exhaustive, especially because in European HEs, the private sector is marginal, whereas in other continents and national systems, the private sector can be large in size.

In this general context, there are some interesting national peculiarities that deserve attention. For example, there is a very relevant point of the performance funding linked to the quality of research that many observers consider as the pillar of every steering at a distance governmental policy and as the main innovation introduced in the last decades. On this crucial issue, it has to be noted that many countries have also introduced strong systems of performance evaluation for university research based on a period of institutional research assessment. However, among these countries, only a few link recurrent assessment to performance funding: Australia, Belgium, Hong Kong, Italy, Norway, the Slovak Republic, Spain, New Zealand and the UK. Among them, two countries allocate a significant portion of public funding to universities on the basis of national research assessment: Italy (30% in 2021) and the UK (approximately 50% of the direct public grant).

Thus, the role of evaluation, and the evaluation of research in particular, has become a pillar of the new existing governance arrangements; however, its impact is very different according to the specific national choice with respect to the financial relevance of the related public funding.

In this context, it is relevant to observe how in every country’s governments are also trying to implement national ways to make systemic performance stronger. For example, various countries (e.g. France, Germany, the Netherlands and Italy) have adopted contracts between the ministry and the individual universities to push towards institutional profiling. To increase the competitiveness of the national system, Germany has adopted the Excellence Initiative. France has created a national champion by merging a few higher education institutions in Paris and creating the University of Paris-Saclay. Italy has assigned extra funding to the best university departments.

## **6 The Evolution of Systemic Governance in Italy: A Long Process of Reforms with Contradictory Results**

According to what sketched above, Italy emerges as a contradictory case because it looks that the waves of reforms have created an apparently contradictory systemic governance arrangement: significantly performance-oriented but also deeply re-regulated. To understand this contradiction, which is the product of a specific national sequence of reforms, it is useful to summarise the diachronic evolution of the designed changes in the governance arrangements.

The Italian university system was characterised by centralised bureaucratic control and a self-governing academic guild (Clark, 1977). Thus, it was subject to a virtually pure type of bureaucratic governance mainly because the government had never indicated any clear goals for universities to pursue. From the 1960s through the 1980s, Italy's university system developed in an anarchical manner under the pressure of demand without being governed at all by the political centre. As a result, at the end of the 1980s, the situation was truly chaotic (Capano, 1998).

Suddenly, after a brief parliamentary debate, a new Ministry of University and Technological Research (MUTR) was created in 1989 under Italian Law no. 168. This law can be thought of as a watershed moment in Italian higher education policy and the beginning of a process of radical innovation, at least at the legislative level. In fact, Law 168 provided for a general framework of didactic, organisational and scientific autonomy for every university and thus can be considered the point of departure from the previous governance mode. The development of policy design in Italian higher education is characterised by constant legislation; this is understandable given that the original governance mode was highly centralised and bureaucratic. Table 2 presents the main policy design decisions made during the period 1989–2018 through which the Italian governance arrangements have been changed to deal with those global challenges that have been sketched above (Capano, 2011; Rebora & Turri, 2009; Capano et al., 2016; Capano, 2018).

As seen from the list of decisions, Italian policy design dynamics in the field of higher education have been characterised by constant reforms of the governance mode.

The new governmental goal is to shift to a steering at a distance model, which has been justified more in ideological terms than from a practical point of view. In other words, the idea of giving universities greater autonomy does not derive from a perception of any specific systemic need but rather from the general idea that the system could perform better if universities were more independent of bureaucratic, centralised control (Capano, 1998). Consequently, there was no clear idea for how to redesign the system according to the new governance mode; this led to the constant changes in national regulations that were designed to give greater powers to Italy's universities during the 1990s.

However, universities' perceived performance, especially in the teaching field, remained unsatisfactory; thus, a complete redesign of the features of institutional governance was approved in 2010 based on the idea that by strengthening institutional governance, universities would perform better and could thus be genuinely steered at a distance. At the same time, this attempt to correct how the steering at a distance model had worked in the previous 20 years was accompanied by substantial financial retrenchment and clear over-regulation of financial and recruitment matters, together with substantial bureaucratisation of the accreditation processes (Capano, 2014; Rebora & Turri, 2013; Turri, 2014; Reale & Primeri, 2014). Therefore, what emerges from the policy design dynamics of the Italian attempt to shift towards a steering at a distance model of higher education governance is that:

**Table 2** Main policy design decisions in Italian higher education governance

Year	Decisions
1994–95	<ul style="list-style-type: none"> <li>• Several budgetary laws were introduced: The lump-sum budget, the establishment of a national body to assess universities' performance (the National Committee for the evaluation of universities), the creation of internal assessment units within universities, a provision stating that the student fees at each university should not exceed 20% of the public funding received by the universities, and the introduction of evaluations of institutional performance for the allocation of public funding. It should be noted that this provision was terminated in 2003 when the percentage to be allocated reached 7%).</li> </ul>
1998–99	<ul style="list-style-type: none"> <li>• A new law reformed the recruitment of academic staff by decentralising the existing system of national, centralised competition for posts.</li> <li>• A ministerial decree introduced a system of undergraduate/postgraduate curricula according to the Bologna declaration; this provision was partially reformulated in 2004 and 2007 to correct the misbehaviour of universities in implementing the new system.</li> </ul>
2004	<ul style="list-style-type: none"> <li>• A new ministerial decree introduced minimum requirements to establish a degree course.</li> </ul>
2007	<ul style="list-style-type: none"> <li>• A new ministerial decree provided for further restrictive regulations for establishing a degree course.</li> <li>• The National Agency for evaluation and accreditation of higher education was established.</li> </ul>
2008	<ul style="list-style-type: none"> <li>• The budgetary law was amended to cut public funding to universities, particularly through a cap on turnover. Public funding was reduced by 20% between 2009 and 2014.</li> </ul>
2010	<ul style="list-style-type: none"> <li>• A new law reformed the institutional governance of universities with the aim of encouraging 'corporate' behaviour and increasing their institutional accountability. The main changes provided for by this law were as follows: <ul style="list-style-type: none"> <li>– Strengthening the role of the boards with respect to the senates and the role of the rector, who is still elected by the university community;</li> <li>– Abolishing faculties, with departments provided with all of the most important powers regarding academic and teaching affairs;</li> <li>– Creating a national system of university accreditation, evaluation and self-evaluation;</li> <li>– Establishing strong financial provisions regulating the recruitment and promotion of university lecturers and professors;</li> <li>– Establishing the structure of an academic career composed of a tenure track assistant professor position and associate and full professor positions as previously, there were three tenure positions;</li> <li>– Establishing a national system to obtain the scientific qualifications necessary to apply to the institutional calls for associate and full professor.</li> </ul> </li> <li>• A new salary system for academics based on periodic (i.e. every 3 years and, from 2020 onwards, every 2 years) disbursement to be implemented autonomously by universities.</li> </ul>
2011	<ul style="list-style-type: none"> <li>• First round of the National Assessment of Research (2004–2010).</li> </ul>
2013	<ul style="list-style-type: none"> <li>• A new law established that public funding (up to 30% in the next few years) would be assigned through a competitive mechanism based all on the quality of research and recruitment. Furthermore, over the next few years, the remaining 70% of public funding will be allocated on the basis of standard costs per student.</li> </ul>
2016	<ul style="list-style-type: none"> <li>• Second round of the National Assessment of Research (2011–2014).</li> </ul>
2017	<ul style="list-style-type: none"> <li>• Selection of the best 180 departments to which 271 million of Euro have been distributed for the period 2017–2022.</li> </ul>
2019	<ul style="list-style-type: none"> <li>• Third round of National Assessment of research (2015–2019).</li> </ul>

- It has been slow in developing the new governance mode and has done so in an incoherent manner. Too often, regulation has been contradictory (i.e. strong procedural regulation—unclear goals, ambiguous voluntary compliance—or strong regulation without any real warnings or sanctions for misconduct). There has never been any package design besides bricolage or layering. For example, the law reforming institutional governance is written in an ambiguous manner (i.e., universities have been given considerable room to design their own forms of internal governance), and thus universities have had the chance to design institutional governance in a way that does not guarantee any greater institutional accountability (Moscati, 2014). Very often, new instruments layered onto the existing governance mode create tensions that require further governmental intervention.
- It has never truly specified its systemic targets in either teaching or research, with public documents speaking merely of ‘improving’. This lack of clear national policy goals leaves the systemic results in the hands of the universities.
- It has displayed, in some cases, a considerable instrumental capacity to formulate genuine policy design; however, it has performed very poorly from the technical point of view. None of the most important design decisions have been made on the basis of real, sufficient evidence-based knowledge. The main approach adopted in the formulation of the new policy intervention is the typical one adopted by the Italian governments: a select few people, appointed by the minister in question, work with certain senior ministerial bureaucrats on the preparation of a first draft, which is then shared with an expert and discussed with the parties’ representatives (and with the Italian Conference of Rectors). In the case of laws, the parliamentary stage of formulation has been characterised by the government’s need to address the customary splits within the parliamentary coalition. This process has never changed; thus, from a design perspective, little has been learnt over the last 30 years.
- Reform laws and rules have been written in an ambiguous way or have left significant room for universities to interpret them. In Italy, it seems that the traditional lessons of the top-down implementation school are still being heeded today; effective policy design must structure implementation so as to enhance compliance by the implementers (Pressman & Wildavsky, 1973; Sabatier & Mazmanian, 1981).
- There has not been a clear political choice with respect to the characteristics of the system and, above all, in terms of the differentiation of Italian higher education. This lack of political vision and guidelines allows every university of the country to decide what to be independent from its real resources and related socio-economic context.

Furthermore, it has to be emphasised that the adopted instruments have been incapable of developing complementarities and thus very often have clashed with each other. Subsequently, the chosen reforms have become ineffective, thus obliging Italian governments to intervene again and again.



Thus, what emerges is that the actual governance arrangements in Italian higher education are characterised by merging different policy tools in a very incoherent way. All in all, there has not been a clear political choice with respect to what the system should do and how it should do it. Substantially, there has also not been a clear political choice about the way of working in the higher education system as well as respect for its social and economic mission.

A first clear example of this ambiguity is found in the emphasis of the financial incentives and of the evaluation of research while maintaining the attitude of bureaucratic regulation. As mentioned above, Italy is one of the countries in which the financial impact of the public funding of the periodic national research assessment is higher; additionally, it is one of the few countries in which there has been an assignment of extra money to university departments on a meritocratic basis. This adoption of performance-based funding is characterised as having been introduced like it was a neutral instrument capable of inducing systemic better performance. There has not been a clear political strategy through which systemic goals have been established to be reached. The main idea was that evaluation *per se* should have contributed to improving the system. Thus, while a policy tool such as evaluation should be a means to reach policy goals according to political preferences (i.e. as a means with which to steer a policy), the adoption of this tool of evaluation in the Italian case has represented a way through which the policy tools themselves have been attributed the role of ruler. The consequence has been, for example, that there has been a structural push to the already existing delineation among universities in a context in which historically universities based in Centre-Northern Italy were in better organisational and financial conditions than those based in Southern Italy (Viesti, 2016; Fadda et al., 2021).

A second relevant example concerns the lack of serious attention to the fact that to make the steering at a distance model function, student mobility should be increased to create the conditions for a real academic market. Regarding student mobility, it is well known that the Italian higher education system has never invested enough money in grants for students. Due to the fact that this allocation is a region-specific task, significant differences exist among Northern and remaining Italian regions; in particular, there have been less financial opportunities for students in Southern regions. As to the academic market, it cannot be understated how, in the last 20 years, academic mobility has been minimal because the current rules of the game do not favour it at all (Seeber & Mampaey, 2021). This indicates that reforms approved in 2010 did work to change the long-lasting localism of academic recruitment. In fact, the new system introduced by that reform has established a national research qualification procedure, *Abilitazione Scientifica Nazionale* (ASN), to impose minimum standards for potential candidates applying for local competitions; thus limiting the traditional discretion of committees and universities. However, this new system did not change the prevalence of localistic interests or the asymmetric chances of being promoted. At the systemic level, 83% of the competitions for associate or full professor posts have been won by scholars belonging to the institutions that launched the calls. In sum, the new system works mostly to promote internal candidates (Abramo & D'Angelo, 2020). The way the

current career and recruitment system represents a structural constraint to the full potential of universities to act strategically in terms of searching for the human resources they would need to pursue innovation in their missions.

It appears that the new systemic governance is problematic in terms of outcomes. Furthermore, it has to be observed that while various attempts at re-regulating the system have been adopted over time, especially in terms of procedural regulation, the strengthening of institutional governance (i.e. the other pillar of the steering at a distance model) has not been significantly reached. This can be seen, for example, in the way in which universities have implemented the power to decide whether or not to attribute to their professors the periodic increase of salary; in fact, in all the universities, the adopted rules for this are not demanding.

This way of designing and implementing the reforms of governance arrangements of higher education has produced contradictory dynamics and results. Generally speaking, the actual situation is characterised by a significant conflict on the role of evaluation, by a recurrent attempt of the centre of the system to regulate the behaviour of institutions in terms of procedures, while the institution can enjoy relatively high autonomy in complying with these attempts at re-centralisation. In sum, it is evident that the adopted variant of the steering at the distance model has not been capable of massively overcoming the past legacies characterised by a significant bureaucratic role of the centre of the system and by a low capacity of universities to behave as corporate organisations. Thus, the impact of these reforms on the main dimensions of universities' performance in Italy (e.g. teaching, research, third mission) is still very problematic in terms of effectiveness.

## **7 The Gordian Knots of Systemic Governance in Italian Higher Education and the National Plan of Recovery and Resilience**

The governance of the Italian university system has undergone a significant redesign of its arrangements both at the national and at the institutional level; however, the final results do not look very satisfactory. Universities have more autonomy now while, at the same time the centre of the system is not very demanding in terms of accountability of local choices and results. Evaluation is pervasive but ineffective in terms of pushing universities towards strategic choices; some rules, especially those regarding academic recruitment and career, clearly represent constraints in terms of institutional strategic capacity. It should be noted that these rules are welcome inside universities because they increase the expectations of internal promotion.

The system's current governance arrangements and ways of working will be challenged by two new events: the proposed increase of the public funding in the years 2022, 2023, 2024 (more than the 20%) and the investment of more than 5 billion Euros due to the National Plan of Recovery and Resilience that, as noted in the introduction of this book, is firmly committed to resolving the problem of

access to higher education, increasing the systemic amount of applied research and partnering with universities and actors in the economic system to increase the offer of vocational degrees. This plan can be considered the first real and ambitious attempt to shift the Italian university system from a traditional way of working towards structural integration to better serve the national needs of socio-cultural and economic development (Capano & Regini, 2021). However, the success of this plan, as well as the efficient and effective investment of the new public funding, is linked not only to external variables (e.g. the governance capacity that the Italian government will show in managing the implementation of the NPRR and the pressure of the EU level) but also by the characteristics of the governance arrangement of the higher education system. Thus, it is necessary to rethink how this governance system works and eventually consider the opportunity to take those choices that have been postponed or excluded by the decisional agenda regarding, for example, the issue of the institutional differentiation of universities (Capano et al., 2016). Furthermore, there is the problem of determining whether most universities are truly capable of becoming strategic actors (as theoretically imposed by the logic of the NPRR and by the global competition in higher education). Apparently, modern institutional governance does not differ much from its past; thus it is prone to distributive and democratic-corporatist logics of actions. This problem cannot be dealt with only by assuming that strong action at the centre of the system will result in due peripheral, ripple-effect reactions. To increase the chances of the best implementation of the NPRR and to ensure the efficient use of the new financial sources, significant intervention regarding a clear political decision with respect to institutional profiling, new rules and incentives to design a real academic market and a significant restyling of the arrangements of institutional governance are necessary.

These changes are necessary not only to properly evaluate and assess universities but also to unlock the potential of evaluative tools (that is masterfully shown in the chapters of this book) to assist decision-makers towards improving the overall performance of the university system and all its fundamental missions.

## References

- Abramo, G., & D'Angelo, C. A. (2020). Were the Italian policy reforms to contrast favoritism and foster effectiveness in faculty recruitment successful? *Science and Public Policy*, 47(5), 604–615.
- Amaral, A., Jones, G., & Karseth, B. (Eds.). (2002). *Governing higher education: National perspectives on institutional governance*. Kluwer Academic Publishers.
- Becher, T. (1989). *Academic tribes and territories*. Open University Press.
- Becher, T., & Kogan, M. (1992). *Process and structure in higher education* (2nd ed.). Routledge.
- Berdahl, R. (1999). Universities and governments in the 21st century: The US experience. In D. Braun & F.-X. Merrien (Eds.), *Towards a new model of governance for universities? A comparative view* (pp. 59–77). Jessica Kingsley.
- Braun, D., & Merrien, F. X. (1999). *Towards a new model of governance for universities? A comparative view*. Jessica Kingsley.
- Capano, G. (1998). *La politica universitaria*. Il Mulino.

- Capano, G. (2008). Looking for serendipity: The problematical reform of government within Italy's universities. *Higher Education*, 55(4), 481–504.
- Capano, G. (2011). Government continues to do its job. A comparative study of governance shifts in the higher education sector. *Public Administration*, 89(4), 1622–1642.
- Capano, G. (2014). The re-regulation of the Italian university system through quality assurance. A mechanistic perspective. *Policy and Society*, 33(3), 199–213.
- Capano, G. (2018). Policy design spaces in reforming governance in higher education: The dynamics in Italy and the Netherlands. *Higher Education*, 75(4), 675–694.
- Capano G and Pritoni A. (2019). Varieties of Hybrid Systemic Governance in European Higher Education. *Higher Education Quarterly* 73(1): 10–28.
- Capano, G., & Jarvis, D. (Eds.). (2020). *Convergence and diversity in the governance of higher education. Comparative perspectives*. Cambridge University Press.
- Capano, G., & Pritoni, A. (2020a). What really happens in higher education governance? Trajectories of adopted policy instruments in higher education over time in 16 European countries. *Higher Education*, 80(3), 989–1010.
- Capano, G., & Pritoni, A. (2020b). Exploring the determinants of higher education performance in Western Europe: A qualitative comparative analysis. *Regulation & Governance*, 14(4), 764–786.
- Capano, G., & Regini, M. (2021). Le politiche universitarie al tempo del Covid-19: le risposte parallele. *Politiche Sociali/Social Policies*, 8(2), 281–303.
- Capano, G., Regini, M., & Turri, M. (2016). *Changing governance in universities. The Italian case in comparative perspective*. Palgrave-MacMillan.
- Cheps. (2006). *The extent and impact of higher education governance reform across Europe*. Final Report to the Directorate General for Education and Culture of European Commission. <http://ec.europa.eu/education/>.
- Clark, B. (1977). *Academic power in Italy*. Chicago University Press.
- Clark, B. (1983). *The higher education system. Academic organization in cross national perspective*. University of California Press.
- Clark, B. (1995). Complexity and differentiation: The deepening problem of university integration. In D. Dill & B. Sporn (Eds.), *Emerging patterns of social demand and university reform: Through a glass darkly* (pp. 159–169). Pergamon Press.
- Donina, D., Meoli, M., & Paleari, S. (2015). Higher education reform in Italy: Tightening regulation instead of steering at a distance. *Higher Education Policy*, 28(2), 215–234.
- El-Khawas, E. (2002). Governance in US universities: Aligning internal dynamics with today's needs. In G. Amaral, G. A. Jones, & B. Karseth (Eds.), *Governing higher education: National perspectives on institutional governance* (pp. 261–278). Kluwer Academic Publishers.
- Enders, J., de Boer, H. F., & Weyer, E. (2013). Regulatory autonomy and performance: The reform of higher education re-visited. *Higher Education*, 65(1), 5–23.
- Enders, J., & Oliver Fulton, O. (Eds.). (2002). *Higher education in a Globalising world*. Kluwer.
- Fadda, N., Marinò, L., & Pischedda, G., et al. (2021). The effect of performance-oriented funding in higher education: Evidence from the staff recruitment budget in Italian higher education. *Higher Education*, published online 27 May <https://doi.org/10.1007/s10734-021-00725-4>
- Gornitzka, A., Kogan, M., & Amaral, A. (2005). *Reform and change in higher education*. Springer.
- Harvey, D. (2005). *A brief history of neoliberalism*. Oxford University Press.
- Huisman, J. (2009). *International perspectives on the governance of higher education*. Routledge.
- Jones, G. A. (Ed.). (2012). *Higher education in Canada: Different systems, different perspectives*. Routledge.
- Kerr, C. (1963). *The uses of university*. Harvard University Press.
- Kehm BM, Lanzendorf U (2006) Germany – 16 Länder approaches to reform. In: Kehm BM, Lanzendorf U, Reforming University Governance. Changing Conditions for Research in Four European Countries. Lemmens, Bonn, pp 135–186

- Lutz W.F. (1982). Tightening up Loose Coupling in Organizations of Higher Education. *Administrative Science Quarterly*, 27(4), pp. 653–669
- Lazzaretti, L., & Tavoletti, E. (2006). Governance shifts in higher education: A cross national comparison. *European Educational Research Journal*, 5(1), 18–36.
- Leslie, D. W., & Novak, R. J. (2003). Substance versus politics: Through the dark mirror of governance reform. *Educational Policy*, 17(1), 98–120.
- Maassen, P., & Olsen, J. P. (Eds.). (2007). *University dynamics and European integration*. Springer.
- Marginson, S. (2009). Hayekian neo-liberalism and academic freedom. *Contemporary Readings in Law and Justice*, 1(1), 86–114.
- McLendon, M. (2003a). Setting the governmental agenda for state decentralization of higher education. *Journal of Higher Education*, 74(5), 479–515.
- McLendon, M. (2003b). State governance reform of higher education: Patterns, trends, and theories of the public policy process. *Higher Education Handbook of Theory and Research*, 27, 57–143.
- McLendon MK, Hearn JC. (2009). Viewing Recent U.S. Governance Reform Whole: “Decentralization” in a Distinctive Context. In: Huisman J. (ed) *International Perspectives on the Governance of Higher Education*, pp. 161–181. Routledge, New York
- Mignot Gerard, S. (2003). Who are the actors in the government of French universities? The paradoxical victory of deliberative leadership. *Higher Education*, 45(1), 71–89.
- Mora, J.-G., & Vidal, X. (2005). Two decades of change in Spanish Universities: Learning the hard way. In A. Gornitzka, M. Kogan, & A. Amaral (Eds.), *Reform and change in higher education* (pp. 135–152). Springer.
- Moscato, R. (2014). Autonomy for what? The university mission in a centralised higher education system – The case of Italy. In M. Shattock (Ed.), *International trends in university governance* (pp. 89–104). Routledge.
- Neave, G., & Van Vught, F. (1991). *Prometheus bound*. Pergamon Press.
- Olssen, M., & Peters, A. (2005). Neoliberalism, higher education and the knowledge economy: From the free market to knowledge capitalism. *Journal of Education Policy*, 20(3), 313–345.
- Orton, D., & Weick, K. E. (1990). Loosely coupled system: A reconceptualization. *Academy of Management Review*, 15(2), 203–223.
- Paradeise, C., et al. (2009). *University governance. Western European comparative perspectives*. Springer.
- Pressman, J. L., & Wildavsky, A. (1973). *Implementation*. University of California Press.
- Reale, E., & Primeri, E. (2014). Reforming universities in Italy. Towards a new paradigm? In C. Musselin & P. N. Teixeira (Eds.), *Reforming higher education: Public policy design and implementation* (pp. 39–64). Springer.
- Rebora, G., & Turri, M. (2009). Governance in higher education: An analysis of the Italian experience. In J. Huisman (Ed.), *International perspectives on the governance of higher education. Alternative frameworks for coordination* (pp. 13–31). Routledge.
- Rebora, G., & Turri, M. (2013). The UK and Italian research assessment exercises face to face. *Research Policy*, 42(9), 1657–1666.
- Sabatier, P. A., & Mazmanian, D. A. (1981). *Effective policy implementation*. Lexington Books.
- Schon, D. A. (1971). *Beyond the stable state*. Random House.
- Schuetze, H. G., Bruneau, W., & Grosjean, G. (Eds.). (2012). *University governance and reform – Policy, fads, and experience in international perspective*. Palgrave Macmillan.
- Seeber, M., & Mampaey, J. (2021). How do university systems’ features affect academic inbreeding? Career rules and language requirements in France, Germany, Italy and Spain. *Higher Education Quarterly*, published online on 18 January.
- Shattock, M. L. (Ed.). (2014). *International trends in university governance*. Routledge.
- Trakman, L. (2008). Modelling university governance. *Higher Education Quarterly*, 62(1–12), 63–83.
- Trow, M. (1974). *Problems in the transition from elite to mass higher education*. Oecd.

Turri, M. (2014). The difficult transition of the Italian university system: Growth, underfunding and reforms. *Journal of Further and Higher Education*, 40(1), 83–106.

Van Vught, F. (Ed.). (1989). *Governmental strategies and innovation in higher education*. Jessica Kingsley.

Viesti, G. (Ed.). (2016). *L'università in declino*. Roma.

**Gilberto Capano** is Professor of Public Policy at the University of Bologna, Italy. His research focuses on governance dynamics and performance in higher education, policy design and policy change, policy instruments' impact, the social role of political science, and leadership as an embedded function of policymaking.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.



**Part II**  
**Teaching and Students' Careers in Italy**



# Do Financial Conditions Play a Role in University Dropout? New Evidence from Administrative Data



Dalit Contini and Roberto Zotti

**Abstract** A large strand of research in the economics and sociology of education has highlighted the existence of deeply rooted inequalities in educational choices along socioeconomic lines, even when net of prior performance. These disparities may take different forms at different stages of schooling and across institutional systems. Yet, due to the lack of data, it is often difficult to disentangle the role played by the various dimensions of socioeconomic background on students' educational careers. While parental education and occupation may shape aspirations (and thus the wish to undertake ambitious educational programmes), lack of income could represent a material obstacle to the continuation of study. In this chapter, we focus on the effect of financial conditions on the probability of dropping out from university. Italy is an interesting study case, because the education system is mainly public and university tuition fees are relatively low and income progressive. Because direct costs for disadvantaged students are low, we would expect income not to be highly relevant in this context. By exploiting a unique data set from the University of Torino (in northern Italy) linking administrative data from students' university careers and information on parental characteristics collected at matriculation, we analyse how socioeconomic background influences the first-year dropout probability. While extremely relevant in earlier educational outcomes, parental education and occupation no longer exert a sizable effect at this point in students' lives. Instead, we find that economic conditions greatly influence the chances of completing university. This result suggests that low tuition fees may be insufficient to foster the participation of low-income high school graduates and that additional forms of support might be needed to ensure equity and, at the same time, raise the share of young people with higher education degrees, which is still too low in Italy.

---

D. Contini (✉) · R. Zotti

Department o of Economics and Statistics “Cognetti De Martiis”, University of Torino, Torino, TO, Italy

e-mail: [dalit.contini@unito.it](mailto:dalit.contini@unito.it); [roberto.zotti@unito.it](mailto:roberto.zotti@unito.it)

© The Author(s) 2022

D. Checchi et al. (eds.), *Teaching, Research and Academic Careers*,  
[https://doi.org/10.1007/978-3-031-07438-7\\_3](https://doi.org/10.1007/978-3-031-07438-7_3)

**Keywords** Higher education · University dropout · Household financial conditions

## 1 Introduction

There is a huge literature from the economics and sociology of education analysing the role played by family background and economic resources on individuals' schooling and college choices. Overall, this body of work provides overwhelming evidence that educational choices are strongly influenced by family background. It is widely recognised that, on average, children from higher socioeconomic status backgrounds perform better at school: this pattern is attributed to the capability of more advantaged parents to purchase better quality education, offer cultural stimuli, and support their children in case of difficulties. Yet, students from advantaged backgrounds make more ambitious school choices and exhibit better outcomes net of prior scholastic results. Further differences in educational choices across family backgrounds may emerge because, acknowledging their own ability, rational individuals take decisions according to costs and expected benefits, maximising a utility function.

Breen and Golthorpe (1997) conceptualise utility in terms of expectations concerning the social class destinations of their offspring and, emphasising the role of aspirations, assume that individuals aim at minimising the risk of social demotion (i.e. ending up in a lower class than that of their parents). Parental education is also valued as a major driver of aspirations, and most empirical analyses of the effects of family background on educational outcomes either focus on the role of parental education or control for it. Other channels might exacerbate differences across family backgrounds in retention. Tinto (1975, 1993) highlights the role played by academic and social integration. Student academic performance and interaction with faculty, as well as involvement in informal peer-group interactions, may lead to either positive or negative experiences that affect feelings of inclusion. Students who feel more disconnected are more likely to withdraw: because first-generation university students often lack good knowledge of and familiarity with the higher education system, they tend to have a higher chance of experiencing poor integration and eventually drop out.

A large body of the economic literature is centred on the role played by family income, and the utility function is defined in terms of children's future earnings. As discussed in Becker (1975), low-income families may face limited borrowing opportunities. Credit constraints may discourage college attendance among youth from low-income families, even when the financial returns are high. However, Cameron and Heckman (2001) and Carneiro and Heckman (2002) find relatively small gaps by family income after controlling for children's ability. They conclude that the long-run factors associated with family income—family environment, early investments in children's education—are what play a prominent role in explaining differential college enrolment rates by family income compared to short-term

borrowing constraints. Similarly, Stinebrickner and Stinebrickner (2008) study college dropout decisions and report little evidence of credit constraints on most students. Instead, other scholars find that financial constraints are important drivers of university enrolment and completion (Ellwood & Kane, 2000; Belley & Lochner, 2007). Comparing cohorts from the mid-seventies studied in Heckman and colleagues' work with cohorts of students from the mid-nineties, Belley and Lochner (2007) find that family income has become substantially more important over time. They conclude that it is likely that borrowing constraints have become more stringent, although they acknowledge that other factors such as social networks, imperfect information and college admissions policies might have played a major role as well. Bound et al. (2012) find that growing difficulties in financing a college education, especially among students from low-income families, have contributed to increasing student employment to cover a greater share of college costs, and in turn to increasing time to degree. Examining college dropout, Stinebrickner and Stinebrickner (2012) argue that students learn about their academic ability from grade performance while in college and provide evidence that a substantial share of withdrawals can be attributed to the gained awareness of poor performance. Indeed, families invest more in their children's education the higher the expectations are of their ability (Checchi, 2000). While affluent parents might still find it worthwhile to keep financing their offspring's education even when they perform poorly, low-income ones are more likely to give up.

The issue of credit constraints is addressed mainly in research on the USA and UK, where the tertiary education system is strongly differentiated, and tuition fees are generally much higher. In European countries, where higher education institutions are mainly public and direct costs are much lower, the explanations put forward by scholars of the potential influence of family income on university attendance (conditional on prior ability and schooling careers) are more generically related to the inability to face costs, including the cost of living, and to foregone earnings (Glocker, 2011; Barone et al., 2014). Where financial difficulties and no efficient student aid system exist, disadvantaged students often need to cover their costs by working, increasing time to degree and/or leading to dropout (Glocker, 2011; Triventi, 2014). In favourable labour market conditions, pull factors may also operate, as in particular, low-income students might be induced to accept good job offers and leave university. Indirect evidence of an impact of family income on higher education attendance and completion is also provided by the numerous studies showing the beneficial effect of student aid in different countries (e.g. Dynarski, 2003; Glocker, 2011; Mealli & Rampichini, 2012; Singell, 2014; Bettinger et al., 2019; Denning et al., 2019; Modena et al., 2020).

Against this background, in this chapter, we analyse whether family economic conditions affect the probability of dropout from Italian university courses upon enrolment. Italy is an interesting study case because the education system is mainly public and university tuition fees are relatively low and income progressive. While parental education and occupation may shape aspirations—and thus the wish to undertake ambitious educational programmes—lack of income could represent a material obstacle to the continuation of study. However, because the direct costs for

disadvantaged students are low, we would expect income not to be highly relevant in this context. As we will show, this is not the case: economic conditions appear strongly associated to student dropout, even after controlling for other dimensions of socioeconomic background, prior school achievement and school type. To our knowledge, there is little existing evidence in Italy on the role played by financial conditions on student academic careers in university. One reason is the lack of appropriate data. Although administrative data provide a measure of family income, it is difficult to identify its independent effect because of the potential confounding of other family background characteristics.

Our research focuses on student educational careers upon enrolment in higher education. Exploiting a unique data set from the University of Torino that links administrative data from students' university careers, information on family income and wealth and information on mothers' and fathers' characteristics collected at matriculation, we disentangle the effects of income, parental education and parental occupation on the probability of dropping out in the first academic year. Information on the financial situation of the family is provided by the ISEE indicator (*Indicatore della Situazione Economica Equivalente*), which is an official document released by the tax authorities delivering a measure of the household economic condition, based on official records of family members' labour income, property and real estate assets, and normalised by the number of components. This document is used to determine tuition fees due for each student.

Parental education and occupation are not available in university registries. To overcome this limitation, the University of Torino has been collecting data on parental education and occupation since the 2014/2015 academic year through an online questionnaire that students fill in at matriculation. Although this section is not mandatory, the large majority (approx. 90%, evenly distributed across subgroups) provide this information. However, nearly 30% of the students do not disclose the ISEE documentation. We show that these data are not randomly missing and that a non-negligible share can be attributed to early dropout decisions. Because in this case complete case analyses or naïve solutions will deliver biased estimates of income effects, we tackle this problem by implementing an appropriate ad hoc imputation strategy.

The rest of the chapter is organised as follows. In Sect. 2, we summarise the existing evidence for Italy. In Sect. 3, we describe the data, and in Sect. 4, we illustrate the problem of missing information investing income data and how we tackle it. In Sect. 5, we describe the empirical strategy, and in Sect. 6, we present our findings. Conclusions follow in Sect. 7.

## 2 The Italian Context

Despite the absence of formal barriers to track choice and access to university, the Italian educational system is flawed by strong socioeconomic inequalities (Cobalti & Schizzerotto, 1993; Checchi & Flabbi, 2007). In comparative research,

Italy stands as a country with particularly large inequalities across parental class and education in upper secondary school choice and access to tertiary education (Jackson, 2013). Family background critically influences students' high school choices (Gambetta, 1987; Schizzerotto & Barone, 2006). Even if inequalities in access to upper secondary education have consistently declined and the share of students enrolling to the academic track has increased over time, class inequalities in track choices have not changed much (Panichella & Triventi, 2014). Horizontal segregation in high school has strong consequences on inequalities in university enrolment, as the transition rate to tertiary education varies largely across tracks (around 80% for students with a lyceum diploma, and below 30% for students with a vocational/technical diploma). Overall, there is evidence of increasing participation in higher education and slightly decreasing inequalities up to the 2000s (Argentin & Triventi, 2011; Guetto & Vergolini, 2016), but in the most recent decade, probably due to the economic crisis, transition rates have been declining and differences across high school tracks have increased, which has determined a change in the composition of the enrolled population (ANVUR, 2016).

Research on student academic careers has been limited by the lack of appropriate longitudinal data at the national level. For this reason, the existing literature on university dropout is largely based on retrospective survey data on high school graduates, periodically run by the National Statistical Institute (Cingano & Cipollone, 2007; Di Pietro & Cutillo, 2008; Cappellari & Lucifora, 2009; Ghignoni, 2017; Contini et al., 2018). This literature reports substantial differentials related to family background and shows that disadvantaged groups in terms of enrolment are also disadvantaged in terms of persistence. These groups include students who attended technical institutes and vocational schools (largely composed of students of lower socioeconomic background), although parental education and social class also influence university attendance and retention, conditional on prior schooling experience. Disadvantaged students are also less likely to enrol in a second tier, once they have obtained a bachelor's degree (Bratti & Cappellari, 2012).

Only a few studies have been based on micro-level administrative data (Belloc et al., 2010; Clerici et al., 2014; Carrieri et al., 2015; Zotti, 2016; Contini & Salza, 2020; Scagni, 2021). Because the archives on schooling and university careers are not linked together, it is not possible to study enrolment choice and consider selection effects. Moreover, a major limitation is that, while it is possible to obtain data on family income, there is no information on parental characteristics. Parental education and occupation influence individuals' aspirations and shape their expectations about future life chances. Economic conditions influence the possibility of bearing the direct and indirect costs of schooling. To disentangle these effects, data on all of these dimensions are needed.

While parental education and class strongly influence high school choices, in Italy there is no evidence of income effects at this stage (Checchi, 2000). This is hardly surprising, because schooling is free up to high school completion, and the expansion of the educational system has now made high school attendance almost universal, as nearly 85% of the young attain a high school qualification. The evidence on the role of economic resources in higher education is mixed. Analysing

a national sample in the survey on Household Incomes and Wealth, Checchi (2000) reports that family income does not seem to play a significant role in preventing the enrolment of cohabiting children in Italian public universities. Instead, Aina (2013) finds sizable effects on enrolment probability but small effects on dropout. Using administrative data from single institutions, Zotti (2016) and Scagni (2021) report income effects on dropout probability. Although analysing the data of single institutions has limited external validity, focusing on more homogeneous environments has the advantage of better controlling for contextual confounding effects. Analysing the University of Salerno, Zotti (2016) reports significant differences between low- and medium-income families in dropout probability. Scagni (2021) analyses data from the University of Torino and finds a sizable effect of income on dropout choices. Belloc et al. (2010), however, report the opposite finding—that low-income students drop out less—for the University Roma La Sapienza. Yet, this result is derived from including university performance (a mediator of dropout) as a control, and thus it is not comparable with the other studies. From a different perspective, Barone et al. (2018) use measures of material deprivation to study university enrolment and find that economic deprivation, as such, matters, even controlling for other variables meant to capture the rational choice mechanisms, in line with the Breen and Goldthorpe’s theoretical model, although it does not play a major role.

Indirect evidence of the role of financial conditions on student academic careers is provided by the compelling evidence that income support provided to low-income students is effective in preventing dropout and fostering in-time graduation (Mealli & Rampichini, 2012; Vergolini & Zanini, 2015; Martini et al., 2021; Modena et al., 2020). Scholarships may favour college enrolment and persistence by providing income that allows students to allocate more time to school activities instead of work.<sup>1</sup>

### 3 Data

We exploit administrative data provided by the Ministry of Education on the entire career of the cohorts of students first enrolled at the University of Torino in a bachelor’s programme in the three academic years from 2015/2016 to 2017/2018. The archive contains full information on the students’ progression (including exam transcripts and credits earned, degree changes, timing of degree attainment or withdrawal); demographic characteristics (gender, age, place of birth and place of residence); and information on previous schooling (type of high school and final examination marks). These data have been integrated with information on

---

<sup>1</sup> As shown by Triventi (2014), students from upper-middle classes have a lower probability of working while studying, and working students have much poorer performance outcomes than full-time students.

family income and tuition payments, with information on scholarship reciprocity<sup>2</sup> and with a unique piece of information on parental education and occupation collected independently by the University of Torino at matriculation since 2014.<sup>3</sup> This makes it possible to improve our understanding of socioeconomic inequalities in higher education, assess the independent contribution of each of these family characteristics and disentangle the effect of economic conditions.

We analyse the determinant of first-year dropout, with a particular focus on the role played by family income. Withdrawal is defined implicitly, based on whether we observe re-enrolment in year 2. Because we have access only to microdata from the University of Torino, we cannot distinguish between changes of institution and withdrawal from higher education altogether.<sup>4</sup> Previous analyses based on more comprehensive data have, however, shown that, among bachelor students, only a small share of the observed dropouts belong to the former group, so we believe we can safely interpret the results in terms of system-level dropout.

In Italian public universities, tuition fees are progressive, depending on household economic conditions. Students make a first payment of a fixed amount at the beginning of each academic year. In late fall, they are asked to provide the ISEE document reporting the family equivalized indicator, based on family members' labour income, properties and real estate assets.<sup>5</sup> Students whose ISEE exceeds a given threshold (currently set around 85,000 euros) or not providing the document are requested to pay the maximum fee (approximately 2500 euros per year). Nearly 30% of students do not provide the ISEE declaration. In the next section, we deal with this issue: as we will show, this piece of information is clearly not missing randomly. This implies that we cannot ignore the issue and conduct a complete case analysis: instead, missing data will be imputed, based on the available information on the following academic years, on parental education and occupation and tuition payments.

## 4 Missing Data on Family Income

If we could assume that, conditional on observed variables, data on income were “missing at random” (MAR), we could conduct a complete case analysis including all of the relevant explanatory variables in the models. There are, however, good reasons to believe this is not the case. First, because high-income students have

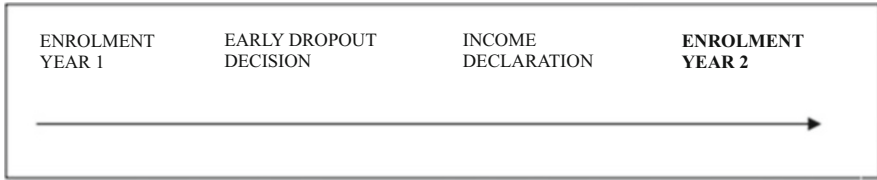
---

<sup>2</sup> Data on scholarships was made available by EDISU-Piemonte (Ente Regionale per il Diritto allo Studio Universitario).

<sup>3</sup> This data collection was spurred by the project EqualEducToEmploy, financed by the Compagnia di San Paolo in 2012–2016.

<sup>4</sup> Students changing their degree programme are not considered dropouts.

<sup>5</sup> Students may figure as an independent household only if they have lived on their own for at least 2 years and if they have earned at least 7000 euros/year. This rule was introduced in the early 2000s to discourage the previous common practice of changing residence to figure as a separate, low-income household and pay low tuition fees.



**Fig. 1** Decision-making timeline

no tuition reductions, they have no incentive to provide an income declaration. Let us label these students *rich*. Indeed, if we could assume that all individuals with missing ISEE exceed the highest threshold, it would not be a big problem, because we would have relevant information on income that we could exploit. Unfortunately, there is evidence against this assumption. When we analyse the characteristics of the students with missing ISEE we find that: (a) many of the students with missing ISEE come from disadvantaged family backgrounds in terms of parental education and occupation (see Table 9 in Appendix A); and (b) many students not disclosing income in year 1 do so in subsequent years, often reporting a low ISEE value (see Table 10 in Appendix A). If economic conditions are fairly stable over a short time span, we may assume that in year 1 they had missed the deadlines, so we call these students *sloppy* and exploit the information provided in later years.

Second, students who decide to leave their studies within the first couple of months of the academic year also have no incentives to declare ISEE, because ISEE determines the second tuition payment, due in late fall. We call these students *early dropouts*. The choice timeline is depicted in Fig. 1.

While the *rich* and *sloppy* can be easily handled by imputing high income or subsequent ISEE values, *early dropouts* involve an endogeneity issue that must be considered. Endogeneity results from the fact that, although we are dealing with missing values for an independent variable, whether this variable is observed or not may depend on the dependent variable itself.<sup>6</sup> Hence, we cannot simply ignore the issue and exclude these cases from the analysis, because we would end up with potentially highly biased estimates of the effect of income on the dropout probability. As we will see later, this practice would lead to substantial underestimation of the effect of interest.

We now describe how to identify the students in these subgroups and our imputation strategy. We classify the students in the cohorts of interest in terms of whether they have or have not provided the income declaration in academic years 1 and 2, whether they have or have not enrolled in year 2 and, when relevant, whether they have paid the second tuition instalment: this piece of information is useful to identify early dropout students. Details are provided in Table 1.

<sup>6</sup> Endogeneity refers to situations in which an **explanatory variable** is **correlated** with the **error term**. In other words, an endogenous variable is a variable whose value is determined by the model, while an exogenous variable is one whose value is determined outside the model.



**Table 1** Classification of students (matriculated population in BA degrees, 2015–2017)

YEAR 1	YEAR 2		Enrolled in year 2 (not dropout)		TOTAL
	Not enrolled in year 2 (dropout)		ISEE provided	ISEE not provided	
ISEE provided	N=3120 9.32%		N=19424 58.01%	N=1105 3.30%	N=23649 70.63%
ISEE not provided	Second tuition payment		N=1698 5.07% <i>(SLOPPY)</i>	N=5873 17.54% <i>(RICH)</i>	N=9836 29.37%
	Yes	N=770 2.30% <i>(OTHER DROPOUTS)</i>			
	No	N=1495 4.46% <i>(EARLY DROPOUTS)</i>			
TOTAL	N=5385 16.08%		N=21122 63.08%	N=6978 20.84%	N=33485 100%

Note: Authors' elaboration.

Most of the students (more than 70% of the entire student population matriculated in bachelor's degree courses) provide ISEE in year 1. Consider the students not declaring ISEE in year 1 (29.37% of the total population); as discussed above, we may identify three relevant clusters: the *rich*, the *sloppy* and the *early dropouts*, as well as an additional residual group. In the following lines, we describe how we identify them and the imputation strategy. Let us start with those who do not drop out by year 2.

1. *SLOPPY*. As argued above, we assume that those who did not declare income in year 1 but declare income in year 2 had previously missed the deadlines: the *sloppy* represent 5.07% of the total population. Assuming short-term stability of economic conditions, we impute ISEE in year 1 using the value reported in year 2.
2. *RICH*. Some students fail to provide the information even in year 2 (and in subsequent years). These students (17.5% of the total population) are labelled *rich*, under the assumption that if a student does not disclose ISEE more than once, it is because there would be no substantial tuition reduction justifying the burden required to produce the documentation. For these individuals, we impute ISEE with a conventional value exceeding the maximum threshold. To keep it simple, we impute the value 100,000 and run robustness checks with alternative values (see Sect. 6).

After these imputations, the share of students with no information on economic condition drops from 29.37% to 6.76%. Even if the size of the missing ISEE population is small at this point, we must still account for the most problematic subgroup of students: those who do not enrol in year 2.

3. **EARLY DROPOUTS.** To identify this group, we exploit an additional piece of information: whether students have made the second tuition payment, due in late fall. We assume that those who did not (4.46% of the total population) have taken the dropout decision before the ISEE deadline. Our imputation strategy for the early dropout students relies on the available information on parental education and occupation and on the observed relation between these family background characteristics and ISEE. Let us define  $I$  as the household economic condition indicator and  $z$  as the vector of dummy variables describing mother's and father's education and occupation. Assuming a linear relation  $I_i = a + bz_i + u_i$ , we estimate model parameters, predict ISEE for given combinations of parental background characteristics and use the estimated  $E(I|z)$  to impute missing ISEE. Yet, to address the endogeneity issue, we must acknowledge that the relation between  $I$  and  $z$  is generally different in the dropout population from that in the student population at large, because economic conditions and other dimensions of family background may themselves affect dropout (see proof in Appendix B). Against this background, we estimate the relation between  $I$  and  $z$  among those dropouts disclosing income and impute the predicted expected value  $\hat{E}(I|z, \text{drop out})$ , under the additional assumption that the same relation holds for early and late dropouts.
4. **OTHER DROPOUTS.** There is an additional small residual group of dropouts (2.3% of the total population), who did not declare ISEE in year 1, but, having paid the second instalment, should not be considered as early dropouts. In principle, we could exploit the observed relation between parental characteristics and income and impute expected income as for the early dropouts; however, this would imply neglecting their decision not to disclose their income. Instead, we may acknowledge that this group is likely to be composed of *sloppy* and *rich* students. However, because they drop out, we cannot observe their behaviour in year 2, so we have no means of identifying them. Hence, we will assume they are all rich. Although this is unlikely to be true for all of the students in this group, by imputing a high value of income to all of them, we tend to narrow the economic differences between dropouts and non-dropouts, delivering a conservative estimate of the true income effect.

## 5 Empirical Strategy and Variables Description

The original sample included 33,485 individuals who first matriculated in bachelor's degree programmes between 2015 and 2017. We excluded from the analyses the students not reporting parental occupation or parental education for both parents (approximately 10% of the original sample, apparently randomly selected) and those who attained a high school degree abroad, because most of them did not report family background information (final sample size  $N = 29,719$ ).

In Table 2, we show descriptive evidence on the ISEE and the parental education distributions of dropouts and non-dropouts. On average, the former display

substantially less favourable economic conditions and a smaller share have parents with higher education degrees. In the last columns, we report the share of dropouts within the population at large and among those providing and not providing the income declaration. As we can see, dropouts are overrepresented among those not disclosing income, confirming the suspicion that provision of the income declaration may be endogenous to the early dropout decision.

To analyse the role of family economic conditions on dropout probability, we estimate logit models where the dependent variable is a binary indicator taking the value 0 if the students enrolled in year 1 re-enrol in year 2 and the value 1 if they do not re-enrol, focusing on students who first matriculated between 2015 and 2017 in 3-year degree programmes. We consider the following baseline specification:

$$D_i^* = \beta_0 + \beta_1 I_i + \beta_2 x_i + \beta_3 z_i + \beta_4 f_i + \beta_5 c_i + u_i \quad (1)$$

$$D_i = \begin{cases} 1 & \text{if } D_i^* > 0 \\ 0 & \text{if } D_i^* < 0 \end{cases} \quad (2)$$

where  $D^*$  is the latent utility of dropout,  $D$  is the observed binary counterpart and the error term  $u$  is distributed as a logistic random variable. The explanatory variable of main interest is  $I = \ln(\text{income})$ , while the control variables are  $x$ =parental education and occupation,  $z$ =socio-demographic characteristics and prior schooling,  $f$ = field of study and  $c$ =matriculation cohort.

Given that we can control for a large array of explanatory variables capturing all of the main determinants described in the existing literature (including other dimensions of socioeconomic background), we are able to estimate the independent effect of family economic conditions on the probability of withdrawal. What often prevents researchers from being able to interpret the income effect as causal is the unavailability of information on parental education and occupation. In the absence of such controls, due to the association between these variables and family income, we would not be able to disentangle income effects from other effects related to family background. Moreover, there are possible selection effects that might affect our results, because by observing only university students we cannot model the enrolment decision. We address these limitations in Sect. 6.3. The explanatory variables are defined as follows:

- *Income* is defined as the natural logarithm of the ISEE indicator, determining family economic conditions from household income, parental wealth and family size. When missing, we use the imputation strategy described in Sect. 4.
- *Parental education* is recorded separately for mothers and fathers, according to the following classification: up to lower secondary school, upper secondary school and higher education. However, in the estimation, we include the highest level between mother and father, further distinguishing between households where one parent or both parents have a university degree.

**Table 2** Family economic condition (ISEE) and parental education by dropout status and share of dropouts among students providing and not providing the income declaration

	ISEE			Parental education			Decl. ISEE %	Not decl. ISEE %
	25th percentile	50th percentile	75th percentile	Both $\geq$ high school	At least 1 higher education	Both higher education		
Dropouts	9886	18,325	29,025	75.8%	22.2%	6.8%	12.7	16.7
Non-dropouts	15, 819	28,990	65,116	80.7%	29.4%	10.2%	87.3	83.3
							All %	
							15.5	
							84.5	

Note: Authors' elaboration

- *Mother and father occupation* is categorised as: blue collar, low-skilled white collar, high-skilled white collar and self-employed.<sup>7</sup> For the mother, we also add the category housework.
- *Female* is dummy variable identifying female students to account for the widespread evidence of gender differences in educational outcomes.
- *Age at matriculation* is included because there is extensive evidence that individuals not enrolling right after the end of high school (possibly after a period of occupation or while working) or, more generally, at an older age (perhaps because they previously experienced grade repetition) are more likely to leave university before degree completion. The variable is included in a categorical version (<=19 years old, 20 years old, 21–25 years old and more than 25 years old) to capture possible non-linear effects.
- *High school track* is included because prior schooling has been shown to strongly affect higher educational choice and outcomes. It is classified into traditional lyceums (classic and scientific), other lyceums (linguistic, human science, artistic), technical schools and vocational schools. Students who attended high school abroad ( $n = 881$ ) were excluded from the analyses.
- *High school final grade*, ranging between 60 (pass) to 100 (excellent) is a proxy of academic preparedness and has been shown to be an important predictor of students' outcomes.
- *Area of origin* may influence the dropout probability for several reasons. First, because there is evidence from national and international standardised assessments that the level of competencies reached in school widely differs across the country (highest in the North and lowest in the South, see Bratti et al., 2007). Second, because students leaving their family of origin and bearing higher costs of living, on the one side are more exposed to changes in family economic conditions, but on the other side, they might be more motivated than stayers. The area of origin has been based on information related to high school location. We adopt the classification: Turin, Piedmont, North-West, North-East, Centre and South.
- *Field of study*. University careers—withdrawal/completion, credit attainment speed, grades—vary across majors and disciplines. We classify the field of study into broad categories: Scientific, Political and Social Sciences, Economics, Humanities, Health and Psychology.<sup>8</sup>
- *Scholarship* is a binary variable taking value 1 if the student receives financial aid in the form of a (small) scholarship and 0 otherwise. In some specifications,

---

<sup>7</sup> 'Blue collar' includes workers; 'Low-skilled white collar' includes clerks and service workers; 'High-skilled white collar' includes senior officials, professionals, teachers and managers; and 'Self-employed' includes business owners, self-employed and freelance.

<sup>8</sup> Scientific includes Mathematics, Physics and Natural sciences; Political and Social Sciences includes Law and Political Sciences; Economics includes Business, Management and Economics & Statistics; Humanities includes Philosophy, History, Languages, Arts and Educational sciences; Health includes all healthcare professions (Nursing, Speech Therapy, Physiotherapy, Dental Hygiene, etc).

we include the variable in the model to account for the evidence that financial aid has a beneficial effect on student progression.

- *Working student* is a binary variable taking the value 1 if the student declares being a working student and 0 otherwise.<sup>9</sup> In some specifications, we include this variable because this condition often entails worse academic outcomes and higher chances of withdrawal.

Descriptive statistics on the full set of variables are presented in Table 3.

## 6 Results

In Table 4, we summarise the results of logit model estimation relative to the effect of income on the dropout probability. All models control for parental education and occupation, gender, age at enrolment, high-school type and final grade and area of origin, as well as including field of study and cohort fixed effects. For comparative purposes, we start with two naïve strategies: a complete case analysis (column 1) and a model including all observations, with a variable taking the observed *ISEE* value if available and 0 if missing and a dummy indicator for missing *ISEE* (column 2).<sup>10</sup> We then move to models using the imputed *ISEE*, according to the procedure described in the previous section: a model with the baseline explanatory variables (column 3) and models adding as control variables an indicator of the student being a scholarship recipient and whether the individual is a working student (columns 4–6).<sup>11</sup>

The effect of income is negative and highly significant in all models, implying that students from more affluent families experience lower chances of withdrawal.<sup>12</sup> The effect appears weaker in the complete case model than in the models where we address the missing data issue with appropriate imputation. The effect is even weaker when we estimate the naïve model in column 2: interestingly, the estimates reveal that the dropout probability for individuals not disclosing *ISEE* is substantially larger even than the probability experienced by those reporting

---

<sup>9</sup> Working students may be eligible for part-time status, which means that they are given twice the time to complete their degree programmes and are entitled to pay reduced tuition fees. Unfortunately, although the administrative data report whether a student declares being a working student, we do not know whether they apply for part-time status.

<sup>10</sup> In this way, the income coefficient describes the effect of *ISEE* among those who declared it, and the missing *ISEE* dummy coefficient captures the difference between those who do not provide *ISEE* and individuals with *ISEE* = 0.

<sup>11</sup> In the first year, the scholarship is granted according to family income, although only approximately half of the eligible students apply for it. From the second academic year upon enrolment, merit restrictions also apply.

<sup>12</sup> By making a single imputation for each missing *ISEE* value, the standard error of the estimates will be underestimated to some extent. However, due to the large sample size, we are confident that the estimates will still be highly statistically significant.

**Table 3** Descriptive statistics

Variable	Obs	Mean	Std. Dev.
Dropout	29,719	0.155	–
Age at enrolment			
<=19 years old	29,719	0.693	–
20 years old	29,719	0.182	–
21–25 years old	29,719	0.102	–
(>25 years old	29,719	0.021	–
Gender			
Females	29,719	0.600	–
Males	29,719	0.399	–
Secondary school			
Lyceum	29,719	0.528	–
Other Lyceum	29,719	0.149	–
Technical	29,719	0.210	–
Vocational	29,719	0.082	–
High school missing	29,719	0.029	–
High school grade	29,719	77.51	11.06
Education of parents			
Lower secondary	29,719	0.203	–
Upper secondary	29,719	0.521	–
One higher education	29,719	0.177	–
Both higher education	29,719	0.097	–
Occupation father			
Blue collar	29,719	0.260	–
Low skilled white collar	29,719	0.238	–
High skilled white collar	29,719	0.139	–
Self-employed	29,719	0.326	–
Occupation missing	29,719	0.035	–
Occupation mother			
Blue collar	29,719	0.184	–
Low skilled white collar	29,719	0.350	–
High skilled white collar	29,719	0.154	–
Self-employed	29,719	0.171	–
Housework	29,719	0.117	–
Occupation missing	29,719	0.020	–
Income (log)	29,719	10.21	0.98
Scholarship	29,719	0.056	–
Student worker	29,719	0.161	–
Field of study			
Scientific	29,719	0.249	–
Political science	29,719	0.211	–
Economics	29,719	0.195	–
Humanities	29,719	0.209	–

(continued)

**Table 3** (continued)

Variable	Obs	Mean	Std. Dev.
Health	29,719	0.105	–
Psychology	29,719	0.029	–
Residence			–
Torino	29,719	0.606	–
Piedmont	29,719	0.240	–
North-West	29,719	0.058	–
North-East	29,719	0.007	–
Centre	29,719	0.008	–
South	29,719	0.078	–
Cohorts			
Cohort 2015	29,719	0.303	–
Cohort 2016	29,719	0.340	–
Cohort 2017	29,719	0.355	–

Note: Authors' elaboration

very poor economic conditions, confirming the suspicion that missing income is at least partially endogenous. In column (3), we find our preferred estimates, which we explain in further detail below. The average marginal effect (AME) is  $-0.234$ ; thus, between the 5th and the 95th income percentile (8.45 and 11.51), the dropout probability of two otherwise identical individuals in terms of demographic characteristics, prior schooling, field of study and parental background, differs by 7.16 percentage points.<sup>13</sup> The effect size is large, if we consider that the overall dropout share in the first academic year is 15–16%. In columns (4)–(6) we include the additional controls: the income effect increases when we include the scholarship variable and decreases slightly when we include the variable student worker. Interestingly, the effects of both controls are large and highly significant. *Ceteris paribus*, scholarship recipients have a dropout probability which is approximately 8 percentage points lower than that of non-recipients: this result confirms the findings of rigorous impact evaluation studies reporting a positive impact of scholarships on student academic careers. Student workers also have a much higher dropout probability (13 percentage points) than non-workers.

We believe the overall effect of income is best captured by the model that does *not* include being a scholarship recipient and being a working student as explanatory variables (Table 4, Column 3), because these variables are endogenous to income and play the role of mediators. Both receiving the scholarship and being a working student are influenced by income: by including them in the model as controls, we would capture the direct effect of income on dropout probability, while failing to acknowledge the—positive or negative—indirect effects. Let us be more specific.

<sup>13</sup> Robustness checks with alternative values of imputed income (80,000, 100,000 and 120,000) are shown in Table 11 in Appendix A. Only marginal changes are observed.



**Table 4** The effect of economic conditions on first year dropout probability (AME)

	(1)	(2)	(3)	(4)	(5)	(6)
(log) Income	-0.0153***	-0.0124**	-0.0234***	-0.0302***	-0.0215***	-0.0276***
	(0.0041)	(0.0042)	(0.0031)	(0.0028)	(0.0026)	(0.0022)
Income missing		0.0783***				
		(0.0095)				
Scholarship				-0.0845***		-0.0779***
				(0.0070)		(0.0075)
Working students					0.1384***	0.1343***
					(0.0073)	(0.0071)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Field of study	Yes	Yes	Yes	Yes	Yes	Yes
Parental occupation	Yes	Yes	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,114	29,714	29,714	29,714	29,714	29,714

Note: Robust standard errors in parentheses clustered at field of study level. \*\*\*  $p$ -value<0.001, \*\*  $p$ -value<0.01, \*  $p$ -value<0.05. Cohort fixed effects include 2015, 2016 and 2017. Field of study includes Scientific, Political Science, Economics, Humanities, Health and Psychology. Parental occupation includes Blue collar, Low-skilled white collar, High-skilled white collar, Self-employed for the father, and Blue collar, Low-skilled white collar, High-skilled white collar, Self-employed and Housework for the mother. Parental education includes Lower secondary, Upper secondary, Higher education or both Higher education. Individual characteristics include age ( $\leq 19$ , 20, 21–25 and  $> 25$  years old), Female, High school type (Lyceum, Other lyceum, Technical and Vocational), High-school grade, Residence (Turin, North-west, North-East, Centre and South)

(1) Scholarships are typically granted to less affluent students, with the explicit aim of supporting their studies. Including the variable in the model would result in inflating the estimate of the income effect, because in this way the income effect would capture the difference in the dropout probability between more affluent and less affluent non-recipient students (or recipients, although this comparison seems less salient). In other words, in doing so we would end up interpreting the income effect as if income support policies did not exist. (2) Working students are generally less affluent than non-workers (Triventi, 2014); moreover, as we have seen, they have a much higher likelihood of leaving university before completion. By interpreting the income effect when controlling for this variable (and thus comparing students with different incomes, but either both working or both non-working), we would then end up underestimating the income effect by ascribing part of the negative effect of the lack of income to the condition of being a student worker, although being a student worker is itself influenced by the lack of income.

## 6.1 Heterogeneity of the Income Effect

Does income influence dropout probability for all students, or is the observed average effect driven by the behaviour of specific subgroups? To answer this question, we conduct separate analyses by gender, high school type, parental education, area of origin and field of study. Overall, income seems to exert a sizable influence on all subgroups, with only minor differences between them and only a few exceptions. We also estimate the income effect by the levels of the two mediator variables, indicating whether the student is a scholarship recipient or a working student. The results are shown in Tables 5, 6, 7 and 8.

Gender differences are small (Table 5). Income seems to have a slightly lower impact on the dropout probability of girls than boys, but the difference is not statistically significant. Income has a stronger effect on students holding technical and vocational high school degrees. Having previously self-selected into less academically oriented high school types, these students are likely to be more exposed to difficulties and may be able to count on lower family support than students from lyceums (Table 6, Columns 1a–4a). There are no sizable differences across parental education levels (Table 6, Columns 1b–4b). Income does not seem to exert an influence on students coming from central south Italy: we interpret this result in terms of self-selection as well. Although these students display a higher propensity to leave their studies compared to students from the North (results not presented here)—perhaps because, as shown by standardised assessments, they reach lower competence levels (Bratti et al., 2007)—they are likely to be especially positively selected in terms of aspirations and motivation and might thus be less exposed to the detrimental effects of low economic resources (Table 6, Columns 1c–4c).

Income plays a role in all fields of study except for health degrees (Table 7). This is not surprising, because of the selective admission to these programmes regulated by *numerus clausus*. Being strongly self-selected at entrance, these students are

**Table 5** Heterogeneous effects by gender

	(1)	(2)
	Females	Males
(log) Income	−0.0215***	−0.0261***
	(0.0024)	(0.0055)
Observations	17,842	11,872
Cohort FE	Yes	Yes
Field of study	Yes	Yes
Parental occupation	Yes	Yes
Parental education	Yes	Yes
Individual characteristics	Yes	Yes

Note: Robust standard errors in parentheses clustered at field of study level. \*\*\*  $p$ -value<0.001, \*\*  $p$ -value<0.01, \*  $p$ -value<0.05. Controls as in Table 4, Column 3

**Table 6** Heterogeneous effects by high school type, parental education and area of origin

	(1a)	(2a)	(3a)	(4a)
	Lyceum	Other Lyceum	Technical	Vocational
(log) Income	-0.0213*** (0.0031)	-0.0142*** (0.0029)	-0.0296** (0.0096)	-0.0403*** (0.0061)
Observations	15,712	4431	6254	2454
	(1b)	(2b)	(3b)	(4b)
	Lower secondary	Upper secondary	One parent HE	Two parents HE
(log) Income	-0.0245*** (0.0049)	-0.0243*** (0.0043)	-0.0217*** (0.0051)	-0.0204** (0.0076)
Observations	6044	15,510	5266	2886
	(1c)	(2c)	(3c)	(4c)
	Turin	Piedmont	North	Centre-south
(log) Income	-0.0212*** (0.0033)	-0.0338*** (0.0026)	-0.0252* (0.0128)	-0.0006 (0.0086)
Observations	18,016	7147	1968	2583
Cohort FE	Yes	Yes	Yes	Yes
Field of study	Yes	Yes	Yes	Yes
Parental occupation	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes
Individual characteristics	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses clustered at field of study level. \*\*\*  $p$ -value<0.001, \*\*  $p$ -value<0.01, \*  $p$ -value<0.05. Controls as in Table 4, Column 3

**Table 7** Heterogeneous effects by field of study

	(1)	(2)	(3)	(4)	(5)	(6)
	Scientific	Political science	Economics	Humanities	Health	Psychology
(log) Income	-0.0159*** (0.0048)	-0.0329*** (0.0055)	-0.0248*** (0.0045)	-0.0250*** (0.0051)	-0.0069 (0.0064)	-0.0267*** (0.0098)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Field of study	Yes	Yes	Yes	Yes	Yes	Yes
Parental occupation	Yes	Yes	Yes	Yes	Yes	Yes
Parental education	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7408	6273	5800	6093	3004	815

Note: Standard errors in parentheses. \*\*\*  $p$ -value<0.001, \*\*  $p$ -value<0.01, \*  $p$ -value<0.05. Controls as in Table 4, Column 3

**Table 8** Heterogeneous effects by scholarship and working student

	(1a)	(2a)
	Scholarship	No scholarship
(log) Income	-0.0172***	-0.0325***
	(0.0046)	(0.0033)
Observations	1538	28,039
	(1b)	(2b)
	Working student	Not working student
Income	-0.0051	-0.0253***
	(0.0047)	(0.0031)
Observations	4785	24,929
Cohort FE	Yes	Yes
Field of study	Yes	Yes
Parental occupation	Yes	Yes
Parental education	Yes	Yes
Individual characteristics	Yes	Yes

Note: Robust standard errors in parentheses clustered at field of study level. \*\*\*  $p$ -value<0.001, \*\*  $p$ -value<0.01, \*  $p$ -value<0.05. Controls as in Table 4, Column 3

highly motivated and generally display very low dropout rates. Similarly, although still sizable and statistically significant, we observe a smaller income effect among students enrolled in the scientific fields, where in many degree programmes there are selective admission tests.

We find no income effects for working students, who are usually engaged in full-time jobs and display much higher dropout probabilities than full-time students. We interpret the absence of income effects for this subgroup as being related to the fact that, earning their own income, they are less dependent on family economic conditions. Income effects are weaker for scholarship recipients (AME = 0.017) than for non-recipients (AME = 0.032). This result provides additional evidence of the beneficial effect of student aid policies, as the scholarship contributes to making recipients less exposed to the negative impact of lack of family economic resources (Table 8).

## 6.2 *The Effect of Parental Education and Occupation*

Although the role of economic conditions emerges clearly, the effect of parental education and occupation is less clear. In Table 12 in Appendix A, we show the estimated effects for all family background dimensions. The effects of parental education go in the expected direction, but they are small and barely significant,

and even weaker results are observed for parental occupation.<sup>14</sup> Hence, we may conclude that at this point of the educational career—after a strong previous social selection that may be represented as an obstacle course for low-SES and a flat road for high-SES individuals—parental education and occupation do not seem to exert any substantial residual effect on the decision to complete the bachelor’s degree.

### **6.3 Potential Limitations**

#### **6.3.1 Peer Effects**

It might be argued that because we have not controlled for peer characteristics, we cannot rule out that our estimates of the effect of financial conditions also capture peer effects. Let us examine this point more closely. Students of higher socioeconomic background have, on average, better peers in terms of academic and soft skills, and better peers foster persistence in education. Hence, the link between socioeconomic background and persistence in education is likely to be causal, but (at least partly) indirect. Yet, if first-year university students’ relevant peers are high school friends and classmates, it is reasonable to consider parental education and high school track—taken jointly—as good proxies for peer quality. If we believe this is the case (this is our standpoint), the issue no longer exists. If instead we believe that income as such may influence the capability of making friends and which friends young individuals make, the income coefficient might indeed also incorporate peer effects. What would the policy implications be in this latter case? If the relevant peers have been established during high school, providing financial aid upon university enrolment might not help reduce dropout, because the aid comes too late. Instead, income support could contribute to reducing dropout if the relevant peers are made after university enrolment, because this additional source of income could foster social integration in university and the acquisition of better peers. In this scenario, the income coefficient captures the total causal effect of income (direct + indirect). Thus, the policy implications may depend not only on whether the relation between economic conditions and retention is truly causal but also on the mechanisms underlying this causal link.

#### **6.3.2 Self-Selection Issues**

By exploiting administrative data on university students, we cannot account for selection effects related to previous educational decisions—the choice of the high school track, high school completion and university entrance. Hence, our estimates

---

<sup>14</sup> Even when parental occupation and ISEE are not controlled for. This result is not shown here and is available upon request.

of the effect of economic conditions on university dropout are not estimates of a causal effect in the usual sense: being conditional only on observed features, they do not capture the differences across the income distribution among individuals otherwise identical in terms of both observed and unobserved characteristics. The comparison is not fully ‘like with like’, because—due to the strong social selection operating along the entire schooling career—upon university enrolment, low-socioeconomic status individuals are likely to be positively selected and thus more endowed in terms of unobserved traits such as motivation and resilience than students from advantaged backgrounds (Cingano & Cipollone, 2007). For this reason, we expect our estimates to be conservative estimates of the total causal effect of income (by total effect we mean the effect inclusive of the potential effects of mediators). This conclusion holds under the assumption that motivation is independent of financial conditions after controlling for parental education and occupation (see Appendix B for proof).

## 7 Conclusions and Discussion

As maintained by Manski (1989) and more recently by Bertola (2021), college dropout need not be considered a social problem, because ‘students contemplating college entrance do not know whether completion will be feasible or desirable. Hence, enrolment is a decision to initiate an experiment, one of whose possible outcomes is dropout’ (Manski, 1989, p. 1). While we do agree with this point, we believe that dropout becomes a social problem if it is mainly experienced by students from disadvantaged backgrounds. If this is the case, we need to gain a better understanding of how to weaken barriers to higher education attainment among young individuals who have taken the decision to enrol in college and thus reduce intergenerational transmission of education and income.

Exploiting the unique administrative data from the University of Torino, which augments administrative university data with information on mothers’ and fathers’ educational level and occupation since academic year 2014/2015, we have been able to analyse whether and how family economic condition, parental education and occupation influence university students’ dropout probability and disentangle their effects. We highlight the existence of a severe missing data problem, elicited by the lack of incentives to provide ISEE documentation if the student’s income exceeds a certain threshold, and most importantly, in case of an early dropout decision. This source of missing data cannot be ignored. We deal with the endogenous missing data issue with an ad hoc imputation strategy and find that at this stage of the schooling career—after a strong previous social selection operating up to university enrolment—parental education and occupation no longer exert a sizable effect on educational choices. Instead, there is evidence that, despite the progressive character

of tuition fees and the existence of scholarships provided to low-income students, financial conditions have a substantial impact on university dropout.

Our results suggest that low tuition fees and current student aid policies, although beneficial, are not sufficient to eliminate the negative effect of a lack of economic resources on student academic careers. Further investigation is needed to gain a better understanding of why this is the case. While still preliminary, our analyses reveal that scholarship recipients are much less exposed to family income effects than non-recipients, even if a sizable effect also exists among them. Moreover, despite all eligible applicants receiving a scholarship in recent years, the take-up rate is low, as only about half of the students meeting the income requirements apply for a scholarship (Laudisa, 2017). Whether this is due to a lack of information or to other reasons remains to be determined, which is necessary if we wish to promote equity and at the same time raise the share of young people with tertiary education, which is still dramatically low in Italy.

## Appendix A: Additional Tables

**Table 9** Individuals with ISEE missing, father education and occupation (%)

Occupation of father	Education of the father			
	Lower secondary	Upper secondary	Higher education	Total
Blue-collar	9.31	3.47	0.19	12.98
Low-skilled white-collar	3.28	11.69	3.03	18.00
High-skilled white-collar	0.56	7.27	13.55	21.38
Self-employed	15.14	20.52	11.98	47.64
Total	28.29	42.95	28.75	100.00
Occupation of mother	Education of the mother			
	Lower secondary	Upper secondary	Higher education	Total
Blue-collar	7.08	3.65	0.26	10.99
Low-skilled white-collar	4.55	22.39	5.85	32.79
High-skilled white-collar	0.19	6.38	13.67	20.25
Self-employed	6.28	12.03	8.04	26.34
Housework	4.62	3.92	1.09	9.63
Total	22.71	48.37	28.91	100.00

Note: Authors' elaboration

**Table 10** ISEE distribution in year 2

Percentile	ISEE
5	4776
10	8320
25	18091
50	32599
75	53427
90	81883
95	98306

Individuals with missing ISEE in year 1 revealing ISEE in year 2  
 Note: Authors' elaboration

**Table 11** The effect of economic conditions on first-year dropout probability (AME)—using different values of imputed income

	(1)	(2)	(3)
	80,000	100,000	120,000
(log) Income	-0.0246*** (0.0031)	-0.0234*** (0.0031)	-0.0224*** (0.0031)
Cohort FE	Yes	Yes	Yes
Field of study	Yes	Yes	Yes
Parental occupation	Yes	Yes	Yes
Parental education	Yes	Yes	Yes
Individual characteristics	Yes	Yes	Yes
Observations	29,714	29,714	29,714

Note: Robust standard errors in parentheses clustered at field of study level. \*\*\*  $p$ -value<0.001, \*\*  $p$ -value<0.01, \*  $p$ -value<0.05. Controls as in Table 4, Column 3. Benchmark estimates in Column (2).



**Table 12** The effect of economic conditions on first-year dropout probability (AME) - All controls

	(1)	(2)	(3)	(4)
(log) Income	-0.0208***	-0.0199***	-0.0239***	-0.0234***
	(0.0033)	(0.0031)	(0.0032)	(0.0031)
Education of the parents (ref: Upper secondary)				
Lower sec		0.0069***		0.0124***
		(0.0012)		(0.0009)
1 parent HE		-0.0032		-0.0064*
		(0.0034)		(0.0026)
2 parents HE		-0.0053		-0.0113
		(0.0086)		(0.0113)
Occupation of the father (ref: Low-skilled white-collar)				
Blue-collar			-0.0041	-0.0065
			(0.0034)	(0.0036)
High-skilled white-collar			0.0000	0.0033
			(0.0091)	(0.0093)
Self-employed			0.0174**	0.0171**
			(0.0060)	(0.0058)
Occupation missing			0.0011	0.0002
			(0.0087)	(0.0086)
Occupation of the mother (ref: Low-skilled white-collar)				
Blue-collar			-0.0044	-0.0081
			(0.0043)	(0.0045)
High-skilled white-collar			0.0026	0.0064
			(0.0075)	(0.0095)
Self-employed			0.0169***	0.0162***
			(0.0043)	(0.0048)
Housework			-0.0142*	-0.0171**
			(0.0056)	(0.0054)
Occupation missing			-0.0359**	-0.0367**
			(0.0122)	(0.0120)
Age (ref. <=19 years old)				
20 years old	0.0278***	0.0278***	0.0279***	0.0281***
	(0.0034)	(0.0034)	(0.0035)	(0.0035)
21-25 years old	0.0602***	0.0598***	0.0614***	0.0610***
	(0.0117)	(0.0116)	(0.0117)	(0.0116)

(continued)

**Table 12** (continued)

	(1)	(2)	(3)	(4)
>25 years old	0.1781*** (0.0389)	0.1747*** (0.0384)	0.1834*** (0.0393)	0.1784*** (0.0394)
Female	-0.0240*** (0.0042)	-0.0246*** (0.0043)	-0.0234*** (0.0041)	-0.0240*** (0.0042)
High school type (ref: lyceum)				
Other-lyceum	0.0677*** (0.0080)	0.0670*** (0.0078)	0.0683*** (0.0082)	0.0673*** (0.0079)
Technical	0.0781*** (0.0096)	0.0765*** (0.0095)	0.0804*** (0.0099)	0.0785*** (0.0096)
Vocational	0.1306*** (0.0062)	0.1288*** (0.0058)	0.1346*** (0.0071)	0.1325*** (0.0065)
High school missing	-0.0363 (0.0329)	-0.0370 (0.0327)	-0.0348 (0.0331)	-0.0357 (0.0329)
High-school grade	-0.0036*** (0.0003)	-0.0036*** (0.0003)	-0.0036*** (0.0003)	-0.0036*** (0.0003)
Area of origin (ref: Turin)				
Piedmont	0.0015 (0.0090)	0.0012 (0.0089)	0.0005 (0.0092)	0.0003 (0.0091)
North-West	0.0310*** (0.0062)	0.0311*** (0.0063)	0.0285*** (0.0061)	0.0287*** (0.0062)
North-East	0.0554*** (0.0157)	0.0577*** (0.0158)	0.0520*** (0.0150)	0.0558*** (0.0156)
Centre	0.1011*** (0.0089)	0.1030*** (0.0090)	0.0981*** (0.0074)	0.1006*** (0.0080)
South	0.0651*** (0.0116)	0.0652*** (0.0116)	0.0653*** (0.0126)	0.0653*** (0.0126)
Field of study (ref: Scientific)				
Political Science	-0.0012 (0.0030)	-0.0013 (0.0031)	-0.0012 (0.0030)	-0.0016 (0.0031)
Economics	-0.0663*** (0.0014)	-0.0664*** (0.0013)	-0.0671*** (0.0014)	-0.0672*** (0.0014)
Humanities	-0.0107*** (0.0019)	-0.0111*** (0.0019)	-0.0105*** (0.0019)	-0.0111*** (0.0020)
Health	-0.0955*** (0.0015)	-0.0958*** (0.0016)	-0.0951*** (0.0015)	-0.0956*** (0.0016)
Psychology	-0.0858*** (0.0024)	-0.0859*** (0.0024)	-0.0872*** (0.0024)	-0.0875*** (0.0024)
Cohort (ref: 2015)				
2016	-0.0101* (0.0050)	-0.0100* (0.0051)	-0.0106* (0.0050)	-0.0104* (0.0050)
2017	0.0115** (0.0042)	0.0116** (0.0043)	0.0115** (0.0040)	0.0118** (0.0041)
Observations	29,714	29,714	29,714	29,714

Note: Robust standard errors in parentheses clustered at field of study level. \*\*\*  $p$ -value <0.001, \*\*  $p$ -value <0.01, \*  $p$ -value <0.05

## Appendix B: Proofs

- (a) The relation between income and parental education and occupation in the dropout group differs from the entire student population.

Calling  $D$  the binary variable describing dropout after year 1,  $I$  income and  $x$  the vector of dummy variables capturing mother and father education and occupation, we now prove that:

$$E(I|x) = a + bx \neq E(I|x, D = 1) \quad (3)$$

Model (1–2) for the dropout decision assumes that the dropout probability depends on income, parental education and occupation, prior schooling characteristics, other individual variables like age at enrolment and area of residence, field of study and matriculation cohort. Simplifying the notation, we indicate with  $C$  the vector of all explanatory variables other than parental education and occupation.

If  $D^*$  is the latent propensity of dropping out after year 1, and

$$D_i^* = \beta_0 + \beta_1 I_i + \beta_2 x_i + \beta_3 C_i + u_i \quad (4)$$

$$P(D = 1|I, x, z) = P(D^* > 0|I, x, C) = P(u > -(\beta_0 + \beta_1 I + \beta_2 x + \beta_3 C)) \quad (5)$$

If  $I = a + bx + v$ , where  $v$  is the error term following the usual assumptions:

$$\begin{aligned} E(I|x, D = 1) &= a + bx + E(v|D = 1) \\ &= a + bx + E(v|u > -(\beta_0 + \beta_1 I + \beta_2 x + \beta_3 C)) \\ &= a + bx + E(v|u > -(\beta_0 + \beta_1(a + bx + v) + \beta_2 x + \beta_3 C)) \\ &= a + bx + E(v|\beta_1 v > -(\beta_0 + \beta_1 a + (\beta_1 b + \beta_2)x + \beta_3 C + u)) \end{aligned} \quad (6)$$

Even if  $\rho(v, u) = 0$ , the relation between  $I$  and  $x$  in the population of dropouts differs from that holding in the population of university students at large. The relation is weaker among dropouts because in this group  $v$  is negatively correlated with  $x$ . If income negatively affects the dropout decision (i.e.,  $\beta_1 < 0$ ), other things being equal, individuals from advantaged parental education and occupation need a relatively low income to make the dropout choice (if income positively affected the dropout choice the opposite would hold; however, there are no theoretical reasons for this to occur).

**(b) The effect of sample selection on the estimation of the income coefficient.**

We consider the following specification for the university enrolment choice:

$$E_i^* = bSES_i + gC_i + \varepsilon_i \quad (7)$$

$$E_i = \begin{cases} 1 & \text{if } E_i^* > 0 \\ 0 & \text{if } E_i^* < 0 \end{cases} \quad (8)$$

where  $SES$  is socio-economic status, for simplicity defined as binary (high  $SES = 1$ , low  $SES = 0$ ) and  $C$  the full array of control variables.

The dropout choice is modelled as:

$$D_i^* = \beta SES_i + \gamma C_i - u_{Mi} - u_{Li} \quad (9)$$

$$D_i = \begin{cases} 1 & \text{if } D_i^* > 0 \\ 0 & \text{if } D_i^* < 0 \end{cases} \quad (10)$$

where  $u_M$  is an unobserved factor representing the individual motivation component that is not captured by the other controls such as the high school track and the final grade, and  $u_L$  is the usual idiosyncratic unobserved component representing pure luck.

The causal  $SES$  effect is defined as the difference in the propensity to drop out between high and low  $SES$ , net of all individual observed characteristics and (unobserved) motivation:

$$E(D^* | SES = 1, C, u_M) - E(D^* | SES = 0, C, u_M) = \beta \quad (11)$$

Instead, the estimable effect is:

$$\begin{aligned} & E(D^* | SES = 1, C, E = 1) - E(D^* | SES = 0, C, E = 1) = \beta^* \\ & = \beta - [E(u_M | SES = 1, \varepsilon > -b - gC_i) - E(u_M | SES = 0, \varepsilon > -gC_i)] \end{aligned} \quad (12)$$

Since  $\rho(u_M, \varepsilon) > 0$  (because more motivated individuals are more likely to attend university) the expression in square parenthesis is negative, as it takes a smaller  $\varepsilon$  for high  $SES$  individuals to enroll, and smaller  $\varepsilon$  entails a smaller  $u_M$ . As  $\beta^* > \beta$  and  $\beta < 0$ ,  $\beta^*$  will be closer to 0 (if negative) than the true causal effect  $\beta$ . In other words, without controlling for sample selection we will obtain an underestimate of the true (negative) effect of  $SES$  on the dropout probability. This argument has been made in a slightly simpler form by Cingano and Cipollone (2007).

Yet, we must acknowledge that these conclusions hold conditional on the additional hypothesis that  $\rho(u_M, SES) \leq 0$ . However, one might argue that higher *SES* individuals display higher aspirations and are more motivated to attain the university degree than lower *SES* individuals—to avoid social demotion, higher *SES* individuals are more prone to make ambitious educational plans (Breen & Goldthorpe, 1997). If this is true,  $\beta^*$  need not be a conservative estimate of the *SES* effect, as  $E(u_M | SES = 1) > E(u_M | SES = 0)$ . Here, even if on average among the enrolled  $\varepsilon$  is larger for low *SES* than for high *SES* (because the condition  $\varepsilon > -b - gC_i$  is less stringent than  $\varepsilon > -gC_i$ ), the expression in square parenthesis in (12) need not be negative.

On the other hand, what we are interested in here is the effect of economic conditions net of parental education and occupation. The caveats just made above should apply to the family background dimensions directly shaping educational aspirations, most likely related to the social position (parental education and social class, usually operationalized in terms of occupation) rather than to economic resources.

Against this background, our conclusion is that the effect of economic resources estimated on a sample of university students, controlling for parental education and occupation, but not accounting for sample selection, can be safely interpreted as a conservative estimate of the total causal effect of financial income on the dropout decision.

## References

- Aina, C. (2013). Parental background and university dropout in Italy. *Higher Education*, 65(4), 437–456.
- ANVUR. (2016). *Rapporto biennale sullo stato del sistema universitario e della ricerca*. Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca.
- Argentin, G., & Triventi, M. (2011). Social inequality in higher education and labour market in a period of institutional reforms: Italy, 1992–2007. *Higher Education*, 61, 309–323.
- Barone, C., Abbiati, G., & Azzolini, D. (2014). Quanto conviene studiare? Le credenze degli studenti su costi, redditività economica e rischi di fallimento dell'investimento in istruzione universitaria. *Quaderni di Sociologia*, 64, 11–40.
- Barone, C., Triventi, M., & Assirelli, G. (2018). Explaining social inequalities in access to university: A test of rational choice mechanisms in Italy. *European Sociological Review*, 34(5), 554–569.
- Becker, G. (1975). *Human capital*. Chicago University Press.
- Belley, P., & Lochner, L. (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human Capital*, 1(1), 37–89.
- Belloc, F., Maruotti, A., & Petrella, L. (2010). University drop-out: An Italian experience. *Higher Education*, 60(2), 127–138.
- Bertola, G. (2021). *University dropout. Problems and solutions*. CESifo Working Paper No. 8880.
- Bettinger, E., Gurantz, O., Kawano, L., Sacerdote, B., & Stevens, M. (2019). The long-run impacts of financial aid: Evidence from California's Cal Grant. *American Economic Journal: Economic Policy*, 11(1), 64–94.
- Bound, J., Lovenheim, F., & Turner, S. (2012). Increasing time to baccalaureate degree in the United States. *Education Finance and Policy*, 7(4), 375–424.

- Bratti, M., & Cappellari, L. (2012). Reform of higher education and social gradients in university graduates outcomes. In J. Ermisch, M. Jannti, & T. Smeeding (Eds.), *From parents to children: The intergenerational transmission of advantage*. Russel Sage.
- Bratti, M., Checchi, D., & Filippini, A. (2007). Geographical differences in Italian students' mathematical competencies: Evidence from Pisa 2003. *Giornale degli Economisti e Annali di Economia*, 66(3), 299–333.
- Breen, R., & Golthorpe, J. (1997). Explaining educational differentials: Towards a formal rational action theory. *Rationality and Society*, 9, 275–305.
- Cameron, S. V., & Heckman, J. J. (2001). The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy*, 109(3), 455–499.
- Cappellari, L., & Lucifora, C. (2009). The Bologna Process and college enrollment decisions. *Labour Economics*, 16, 638–647.
- Carneiro, P., & Heckman, J. (2002). The evidence on credit constraints in postsecondary schooling. *Economic Journal*, 112, 705–734.
- Carrieri, V., D'Amato, M., & Zotti, R. (2015). On the causal effects of selective admission policies on students' performances: Evidence from a quasi-experiment in a large Italian university. *Oxford Economic Papers*, 67(4), 1034–1056.
- Checchi, D. (2000). University education in Italy. *International Journal of Manpower*, 21(3/4), 177–205.
- Checchi, D., & Flabbi, L. (2007). *Intergenerational mobility and schooling decisions in Germany and Italy: The impact of secondary school tracks*. IZA Discussion Papers 2876, Institute for the Study of Labor (IZA).
- Cingano, F., & Cipollone, F. (2007). *University drop-out: The case of Italy*. Temi di discussione del Servizio Studi, Banca D'Italia – Eurosystem, 626.
- Clerici, R., Giraldo, A., & Meggiolaro, S. (2014). The determinants of academic outcomes in a competing risks approach: Evidence from Italy. *Studies in Higher Education*, 40(9), 1535–1549.
- Cobalti, A., & Schizzerotto, A. (1993). Inequality of educational opportunity in Italy, in Shavit Y. and H.P. Blossfeld Persistent inequality. In *Changing educational attainment in thirteen countries* (pp. 155–176). Westview Press.
- Contini, D., Cugnata, F., & Scagni, A. (2018). Social selection in higher education. Enrolment, dropout and timely degree attainment in Italy. *Higher Education*, 75, 785–808.
- Contini, D., & Salza, G. (2020). Too few university graduates. Inclusiveness and effectiveness of the Italian higher education system. *Socio-Economic Planning Sciences*, 71, 100803.
- Denning, J. T., Marx, B. M., & Turner, L. J. (2019). ProPelled: The effects of grants on graduation, earnings, and welfare. *American Economic Journal: Applied Economics*, 11(3), 193–224.
- Di Pietro, G., & Cutillo, A. (2008). Degree flexibility and university drop-out: The Italian experience. *Economics of Education Review*, 27(5), 546–555.
- Dynarski, S. M. (2003). Does aid matter? measuring the effect of student aid on college attendance and completion. *The American Economic Review*, 93(1), 279–288.
- Ellwood, D. T., & Kane, T. J. (2000). Who is getting a college education? Family background and the growing gaps in enrollment. In S. Danziger & J. Waldfogel (Eds.), *Securing the future: Investing in children from birth to college*. Sage Found.
- Gambetta, D. (1987). *Were they pushed or did they jump? Individual decision mechanisms in education*. Cambridge University Press.
- Ghignoni, E. (2017). Family background and university dropouts during the crisis: The case of Italy. *Higher Education*, 73, 127–151.
- Glocker, S. (2011). The effect of student aid on the duration of study. *Economics of Education Review*, 30, 177–190.
- Guetto, R., & Vergolini, L. (2016). *Educational expansion without equalization: Social origins and children's choice of the upper secondary track in Italy (1958-1989)*. No. 2016-01. Research Institute for the Evaluation of Public Policies (IRVAPP), Bruno Kessler Foundation.

- Jackson, M. (2013). *Determined to succeed? Performance, choice and education*. Stanford University Press.
- Laudisa, F. (2017). *Gli studenti iscritti al primo anno che non presentano domanda di borsa di studio pur avendone potenzialmente diritto*. Osservatorio regionale per l'Università e per il Diritto allo Studio Universitario, IRES Piemonte.
- Manski, C. F. (1989). Schooling as experimentation: A reappraisal of the postsecondary dropout phenomenon. *Economics of Education Review*, 8(4), 305–312.
- Martini, A., Azzolini, D., Romano, B., & Vergolini, L. (2021). Increasing college going by incentivizing savings: Evidence from a randomized controlled trial in Italy. *Journal of Policy Analysis and Management*, 40(3), 814–840. <https://doi.org/10.1002/pam.22260>
- Mealli, M., & Rampichini, R. (2012). Evaluating the effects of university grants by using regression discontinuity designs. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 175(3), 775–798.
- Modena, F., Rettore, E., & Tanzi, G. M. (2020). The effect of grants on university dropout rates: Evidence from the Italian case. *Journal of Human Capital*, 14(3), 343–370.
- Panichella, N., & Triventi, M. (2014). Social inequalities in the choice of secondary school: Long-term trends during educational expansion and reforms in Italy. *European Societies*, 16(5), 666–693.
- Scagni, A. (2021). Performance universitaria, rischio di abbandono e background familiare: i nuovi microdati di UNITO. In M. Attanasio, M. Porcu, & G. Ragozini (Eds.), *Verso Nord. Le nuove e vecchie rotte delle migrazioni universitarie*. Franco Angeli.
- Schizzerotto, A., & Barone, C. (2006). *Sociologia dell'istruzione*. Il Mulino.
- Singell, L. D. (2014). Come and stay a while: Does financial aid effect retention conditioned on enrollment at a large public university? *Economics of Education Review*, 23, 459–471.
- Stinebrickner, R., & Stinebrickner, T. (2008). The effect of credit constraints on the college dropout decision: A direct approach using a new panel study. *American Economic Review*, 98(5), 2163–2184.
- Stinebrickner, T., & Stinebrickner, R. (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4), 707–748.
- Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition*, 2nd ed., University of Chicago Press, Chicago, IL.
- Tinto, V. (1975). Drop out from higher education: A theoretical synthesis of recent research. *Review of educational research*, 45(1), 89–125.
- Triventi, M. (2014). Does working during higher education affect students' academic progression? *Economics of Education Review*, 41, 1–13.
- Vergolini, L., & Zanini, N. (2015). Away, but not too far from home. The effects of financial aid on university enrolment decisions. *Economics of Education Review*, 49, 91–109.
- Zotti, R. (2016). Should I stay or should I go? Dropping out from university: An empirical analysis of students' performances. In G. Coppola & N. O'Higgins (Eds.), *Youth and the crisis - unemployment, education and health in Europe*. Routledge.

**Dalit Contini** is Professor of Social Statistics at the University of Torino. Her current research interests cover the area of educational inequalities, higher education choices and academic careers, gender inequalities, school guidance, school systems, impact evaluation of social policies.

**Roberto Zotti** (Ph.D. University of Salerno) is Assistant Professor of Public Economics at the University of Torino. His research interests are in the field of economics of education with a focus on higher education and local economic development, and in the field of public finance with a focus on regional economics, elections, taxation and tax competition.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.





# Drop-Out Decisions in a Cohort of Italian Universities



Gianfranco Atzeni, Luca G. Deidda, Marco Delogu, and Dimitri Paolini

**Abstract** In this chapter, we study the determinants of student drop-out decisions using data on a cohort of over 230,000 students enrolled in the Italian university system. The empirical analysis reveals that the probability of dropping out of university negatively correlates with high school grades and student age, controlling for the course of study and university fixed effects. The benchmark estimation suggests a negative correlation between high school final grade and drop-out probability. We also find that enrolling late at the university increases the likelihood of dropping out. In line with the literature, our results suggest that women have a lower propensity to drop out. Our dataset allows differentiating between students who leave their homes to enroll at university (off-site students) and on-site students. We find that off-site students drop out significantly less than those who study in their hometowns. We provide significant evidence that off-site students are a self-selected

---

We wish to thank Bianca Biagi, Claudio Deiana, Claudio Detotto, Alessandra Faggian, Masood Gheasi, and Jacques Poot for their precious suggestions. We would like to acknowledge the participants to the workshops “Anvur: III concorso pubblico idee per la ricerca” (2019, Rome) and “International and Internal Migration: Challenges and Opportunities in Europe” (2020, L’Aquila). We thank for precious research assistance Paolo Deledda (UNISS). The authors gratefully acknowledge financial support by Regione Autonoma della Sardegna (Legge n. 7), Anvur, III Concorso Pubblico Idee di Ricerca (ANVUR) and Università degli studi di Sassari (bando una tantum ricerca 2019).

---

G. Atzeni · L. G. Deidda  
DiSEA and CRENoS, Università di Sassari, Sassari, Italy  
e-mail: [atzeni@uniss.it](mailto:atzeni@uniss.it); [deidda@uniss.it](mailto:deidda@uniss.it)

M. Delogu  
DiSEA and CRENoS, Università di Sassari, Sassari, Italy  
DEM, University of Luxembourg, Esch-sur-Alzette, Luxembourg

D. Paolini (✉)  
DiSEA and CRENoS, Università di Sassari, Sassari, Italy  
CORE, Université catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium  
e-mail: [dpaolini@uniss.it](mailto:dpaolini@uniss.it)

sample of the total population. Accordingly, we use an instrumental variable (IV) approach to identify the causal relationship. The IV estimation shows that studying off-site negatively affects drop-out decisions and more so for students growing up in the south of Italy who typically study off-site in the Center-North of Italy. Taking advantage of a more detailed dataset concerning students enrolled at the Università di Sassari, we show that the choice of the degree is also important to predict the magnitude of drop-out. Specifically, we resort to a bivariate probit specification to account for self-selection into the course of study, finding that the estimates of the determinants of drop-out and the predicted probabilities are heavily affected. Accounting for self-selection, we show that an unconditional comparison among degrees is misleading, as some degrees attract more heterogeneous students than others, as far as skills and motivation are concerned. For instance, regarding the effect of gender, we show that while the estimation without selection suggests that women drop out less, once we account for selection, the contribution of women to drop-out becomes either positive or negative, depending on which course of study they choose. In line with these results, policymakers should tailor drop-out reducing policy interventions to the specificities of each course of study.

**Keywords** Drop out · Location choice · Instrumental variable · Higher education

**JEL Codes** A22, C26, I20, I21

## 1 Introduction

There is robust evidence that more educated individuals earn higher salaries and enjoy higher employment rates, see OECD (2019). Empirical studies indicate a sizable effect, with an average increase in annual earnings of around 10% per additional year of education (see Card 2001). Nevertheless, in “[..] all developed countries the percentage of students dropping out of university or graduating beyond legal terms is very large [..],” see Aina et al. (2018), page 2. In general, delayed completion of studies reduces the average and the overall skill levels of the working population. Reducing drop-out rates could therefore have a positive impact on the skill composition of the workforce. In turn, this may trigger a positive feedback effect on the economy in terms of both efficiency and inequality. First of all, a more educated workforce would facilitate technological change and technology adoption, see Acemoglu (2002). Second, it could push down the wage skill premium, thereby reducing inequality, see Katz and Murphy (1992). Along with the USA, Italy is one of the OECD countries where the drop-out phenomenon reaches dramatic levels, with more than one student in two dropping out of university before completion, see Aina et al. (2018).

The focus of the chapter is the impact of studying off-site on drop-out behavior. We define off-site students as those who leave their homes to pursue higher

education. Although Italian universities are evenly distributed across the national territory, a nonnegligible fraction of students enroll in universities located in a region or province different from residence.<sup>1</sup>

We exploit the Anagrafe Nazionale Studenti (ANS), a dataset produced by the Ministero dell'Università e della Ricerca (MUR), to study the determinants of the drop-out rate of undergraduates enrolled in Italian universities. The ANS collects information about all students who enrolled in the Italian university system. We rely on three years of data regarding undergraduate (i.e., bachelor) students who enrolled in the academic year 2013–2014. In particular, we study the correlation between drop-out rates and characteristics of students, courses, and universities. Regrettably, the ANS dataset does not provide specific information on the off-site status of the student. However, it provides precise information on the place of residence of the student. Linking this information with the university's geographical location, we construct several indicators that work as proxy variables of the off-site status of the students. In our dataset, 22% of the individuals enrolled in universities located in a region different from that of their residential place. Similarly, 53.5% of the students study in a province different from their residence place. Italian inter-regional student mobility is probably eased by the homogeneous distribution of university fees across all public universities, see (Beine et al. 2020). Indeed, financial barriers to education access are quite low in Italy as poor students have access to a generous system of government grants (Cecchi 2000).

Using the region of origin to define the off-site status, we estimate a reduction of the probability of dropping out associated with the off-site status of 1.62%. The results are also robust to other measures of the off-site status,<sup>2</sup> different estimation strategies, and when we cluster individuals by macro-area.

Our empirical analysis reveals that the probability of dropping out of university is negatively correlated with the high school grades and the age of the students. Our benchmark estimation suggests that one additional point in the high school final grade reduces the probability of dropping out by 4%.<sup>3</sup> Furthermore, enrolling one year later at the university increases the probability of dropping out by 9.8%. Flunking out of high school is the main reason that explains late university enrollment in Italy.<sup>4</sup> Consistently with the literature, our results also show that women have a lower probability of dropping out than men. Interestingly, our results

---

<sup>1</sup> 51 out of the 108 Italian provinces host a university. Furthermore, each Italian region hosts at least one university. For all municipalities, the geodesic distance from the nearest university is less than 108 km (our computation).

<sup>2</sup> Other measures for the off-site status include (i) defining off-site students either as students studying in a university outside their home district and (ii) defining off-site students as the ones studying in a university more than 150 km or 200 km far from their place of origin.

<sup>3</sup> Other studies that found the inverse relationship between high school grades and drop-out rates include Belloc et al. (2010).

<sup>4</sup> Differently from the USA, where grade repetition is usually limited to a particular subject, in Italy it is common practice to let students entirely repeat the high school year when the student fails one or more subjects. The percentage of Italian students reporting having failed at least one

suggest that men have a larger probability of drop-out, slightly less than 3 percent. In line with the literature, we find that individuals who attended a Liceum have a substantially lower probability of dropping out than their peers who attended vocational high school. Indeed, these estimates do not change in all variants that we consider and remain stable under our instrumental variable analysis.

Leaving home to pursue a university education may affect the educational outcomes in several ways. On the one hand, studying far from home requires additional efforts in organizing daily life, building new relationships, and so on. On the other hand, studying off-site requires more financial support, often provided by parents, that may motivate off-site students. Checchi (2000) and Contini and Zotti (2021) report that economic conditions greatly influence the likelihood of completing university studies.

It is widely known that there exist sizable differences between the North and the South of Italy, both in terms of wages and in terms of job opportunities. We interpret these findings in the light of Roy's model of self-selection, see Borjas (1987), with Roy's model predicting self-selection in the flow of migrants. We document that students from the South of Italy are more likely to enroll outside their home region or district than their peers from the country's North. Moreover, southern students tend to move to a university located in the Center-North of Italy. In line with Roy's model predictions, we show that the flow of students follows mostly the South-Center\North direction and that very few northern students move to the South to pursue higher education. Besides, we document that off-site students' skills are higher than the overall population in terms of high school grades. Also, students who attended a Liceum are overrepresented among off-site students. As postulated by the Roy model, evidence of self-selection is reinforced when we run separated estimates by macro-area of origin. For instance, for the northern students, we do not obtain a significant negative coefficient for the off-site proxies, and this can be partially explained by a lower strength of the selection channel for these students compared to what happens in the southern ones.

Our results are in line with Johnes and McNabb (2004), one of the few existing contributions that explicitly address the impact of the off-site status on drop-out rates. In particular, they find that the probability of dropping out is lower for students attending a university far from the one in their parental hometown. Similarly, Modena et al. (2018) report a negative correlation between drop-out rates and studying off-site.<sup>5</sup>

The above discussion leads us to the conclusion that addressing causality with OLS estimates is problematic for two reasons. First, our OLS significant negative

---

year during high school was equal to 16% in 2016, above the OECD average, see <https://www.openpolis.it/quantitative-sono-i-ripetenti-nelle-scuole-italiane/>.

<sup>5</sup> Looking solely at students enrolled at the Università di Sassari, Bussu et al. (2019) find that students who are not from Sassari have a statistically significant lower propensity to drop out. They define students not from Sassari as students whose parental home is located more than 30 km away from Sassari. Zotti (2015) reports a similar relationship focusing on students enrolled at the Università di Salerno.

coefficient for the off-site status proxies in our drop-out regression is potentially an artifact of sample selection bias. Second, off-site students go through a significant change in their daily life that, *ceteris paribus*, may affect their studies. We attempt to tackle this issue by resorting to an instrumental variable (IV) procedure which, taking advantage of a variable correlated with the decision of studying off-site but independent from the outcome (drop-out behavior), should allow isolating the effect of studying off-site on drop-out behavior, removing from the estimate the confounding effects mentioned above.

Technically, we instrument the off-site status proxy with the minimum distance from the closest university (our instrument), controlling for characteristics of the districts by fixed effects. Our IV estimates still uncover a negative relationship, with an impact larger in magnitude than the one suggested by the standard OLS procedure. We also implement the IV procedure by splitting our dataset according to the macro-origin of the students. Interestingly, for the subsample of southern students, the off-site status coefficient substantially increases in terms of magnitude while remaining statistically significant and negative. We suggest interpreting this result as evidence that going off-site eventually positively affects students' motivation coming from more distressed districts. Indeed, aside from identification issues, the causal effect of studying off-site is potentially ambiguous. Studying off-site is more costly in terms of the organization of daily life and from an economic viewpoint. Extra financial support is therefore necessary, which is often provided by off-site students' parents. The extra costs have two opposing effects. On the one hand, the fact that off-site faces a higher cost of studying compared to their peers who study in their hometown undermines the sustainability of the off-site choice, which induces higher drop-out rates. On the other hand, the extra costs might provide extra motivation to the off-site students, which would result in a lower drop-out rate. Accordingly, a negative and significant effect is compatible with the idea that the second effect dominates. Nevertheless, we are fully aware that uncovering robust causal relationships regarding the determinants of drop-out requires particular care due to the pervasiveness of self-selection and unobservables.

Self-selection bias relates to the fact that students choose where to study and which course to enroll in based on unobservable factors that can also affect drop-out. To investigate this issue, we take advantage of a more detailed dataset concerning 16 cohorts of students enrolled at the Università di Sassari. Specifically, we are interested in investigating whether the magnitude of drop-out is also affected by the choice of course of study. It is well known that students' choice of which course to enroll in is influenced by factors such as the likelihood of finding a job after graduation or the popularity of certain studies among teenagers. This may cause a systematic mismatch between the student's abilities and those required to complete a degree successfully. If this deviation were systematic, it would generate a higher level of drop-out in the courses affected by this phenomenon, not depending on the organization's quality or teaching. Our results show that the estimated probability of drop-out in the five most popular departments, i.e., with an above-average enrolment rate, is always lower than that estimated without taking the selection mechanism into account. These results suggest that an unconditional comparison among degrees is

misleading, as some degrees attract more heterogeneous students in terms of skills and motivation. The selection approach also shows that a univariate probit model's estimated parameter without selection may be biased. There is abundant evidence that women drop out less than men. However, this finding may result from women being overrepresented in degrees where drop-out is below average. Once we account for selection, we find that the contribution of women to drop-out is either positive or negative, depending on the choice of the course of study.

The chapter is organized as follows. In Sect. 2, we describe our data and provide some stylized facts on drop-out rates. In Sect. 3, we outline our econometric approach. In Sect. 4, we present the OLS empirical estimates along with several robustness checks. In Sect. 5, we describe and implement the IV estimation procedure to tackle the causality issue. In a separate box, we present the synthesis of the analysis on the relationship between drop-out and choice of study course. Section 6 concludes.

## 2 Data and Variables

In the following, we describe our dataset and the definition of the variables employed in our empirical analysis. Then, we provide some descriptive evidence coming from our data.

### 2.1 Dataset

Our data from the ANS contain information about all population students enrolled in all Italian universities for the cohort of bachelor degree students enrolled in 2013–14 for the first time. We follow the students along with their academic career until the 21st of March 2018. Abstracting from PhD programs, which we do not deal with in this study, Italian universities offer three types of degrees: “Laurea triennale,” which is equivalent to a Bachelor degree, “Laurea specialistica,” which is equivalent to a 2-year Master degree, and “Laurea a ciclo unico,” which combines bachelor and master degrees.

We choose to exclude students enrolled in “Laurea specialistica” or “Laurea a ciclo unico,” because we lack information about the final grade they got in their previous careers as bachelor students. Moreover, we exclude international students, as they seem to be selected from a different population compared to national students and constitute a self-selected group so that drop-out mechanisms would probably be different from those that characterize domestic students. We also exclude students enrolled in online universities.<sup>6</sup> Finally, the above choices lead to a dataset that contains information on 230,336 students.

---

<sup>6</sup> Note that in 2013–2014, online universities accounted for only the 4.53% of the total population of students enrolled in bachelor courses. And, there is no clear meaning for the off-site status when a student enrolls for an online course.

The next step is to provide a precise definition of university drop-out. First of all, notice that due to the peculiar characteristics of the Italian university system, differently from Johnes and McNabb (2004), we cannot differentiate between voluntarily and involuntarily drop-out. We proceed as follows. First, we classify students in four main categories: (A) students who successfully completed their degree by the 21st of March 2018, (B) students who were still enrolled by the 21st of March 2018, having not completed their degree yet, (C) students who changed course/university the year after the first year of enrollment, and (D) students who left the Italian university system.

We build a dummy variable  $D_{i,j,c,t}$  which takes value 1 if a student  $i$  enrolled in course  $c$  at university  $j$  drops out at time  $t$  and 0 otherwise. Concerning the measurement of the student's off-site status, unfortunately, our dataset does not contain direct information on whether the student is actually off-site or not. Hence, to capture the off-site status, we combine information on both the place of residence of the student and university location. We use this information to construct the following three alternative discrete proxies of the off-site status:

1. *OD* (out of district): This variable takes value 1 if the student enrolls in a university located outside her home district. Notice that for a sizable percentage of students, this variable always takes value 1 given that 52 out of the 110 Italian districts do not host any university.
2. *OR* (out of region): This variable takes value 1 when the student  $i$  enrolls in a university located outside the home region. Each Italian region hosts at least one university. Therefore the value of this variable is not prearranged as it is the case for the *OD* variable for a sizable fraction of districts.
3. *OFF<sub>km</sub>*: This variable takes value 1 when the student  $i$  enrolls in a university located further away than a threshold distance from the student's home. We take advantage of the ANS information on students' home residence for all students enrolled in any given university  $j$ . Then, after having obtained geographic coordinates for university  $j$ , we compute travel distance, between the university  $j$  and the home of student  $i$ .<sup>7</sup> This measure rules out the cases of students whose house is close to the regional border who enroll outside the region without changing residence. To deal with this shortcoming, we consider two thresholds, 150 km and 200 km, that give rise to two indicators, *OFF<sub>150</sub>* and *OFF<sub>200</sub>*.

In addition, to capture the off-site status of the student, we also construct two continuous variables. We consider both the travel and geodesic distances between the university  $j$  and the home student  $i$ .

---

<sup>7</sup> We take advantage of the STATA routine developed in Weber and Péclat (2017).

## 2.2 Descriptive Statistics

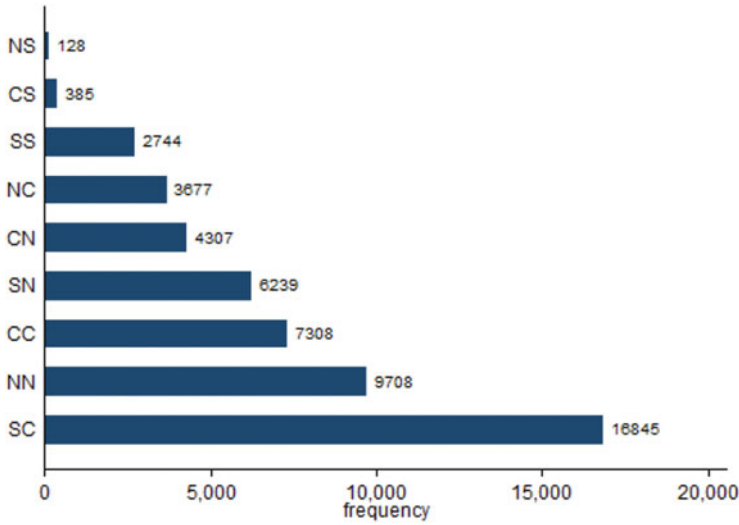
Due to missing values in some variables, we end up with a dataset containing information on 226,094 individuals, representing the 98% of the population of students we initially included. We find that 38.40% of the students completed their degree by the 21st of March 2018, the 17.8% of the student changed course/university, the 38.3% completed higher education, and finally the 12.9% left the university system. We define this last set of students as the droppers. Students enroll in 708 different courses, which belong to 46 different classes, clustered in the four general subject areas: (1) Health, (2) Science, (3) Social Science, and (4) Humanities. Science is the area with more students, representing 38.4% of the sample. Interestingly, slightly more than the majority of students are enrolled either in Humanities or in Social Science. Regarding gender, 54.2% of students are female, while men mainly enroll in Science and only 2.5% enroll in Health. We find that the percentage of women who leave the university, 14.8%, is lower than that of men, 11.2%. Our data show a significant difference in the percentage of drop-outs across the areas of study. While drop-outs are equal only to 5.3% in the Health\Medical area, they reach the sizable figure of 15.1% in Humanities. To account for these patterns, we include fixed effects for the area of study in our empirical estimations. Men leave graduate studies more compared to women in any of the four areas of study. For instance, although women are underrepresented in the area of Science, the percentage of men who drop out is substantially larger than that of women. Accordingly, in our estimation, we include a dummy variable that captures the students' gender. Another finding is that drop-out rates are much larger for students from vocational high schools; this holds for all areas. Students coming from a Liceum show a drop-out rate that is 10% lower. Conversely, students from vocational schools show a much larger drop-out rate, which reaches 21% for the Science area.

One may expect individuals with a low high school grade are overrepresented among the droppers, leading us to include a continuous variable capturing the students' high school grades among the drop-out determinants.

Besides, we find that the drop-out rates exhibit significant variation across the home regions of the individuals. To account for this heterogeneity, fixed effects for the district and region of origin of the student are included in our econometric specification.

The percentage of students studying off-site is unevenly distributed across Italian districts. Measuring off-site students through the variable  $OFF_{150}$ , we find that off-site students reach the sizable figure of 33% among the students who come from the South. Instead, for those both coming from the Center and the North of Italy, the percentages are much lower and amount to 16% and 17%, respectively. Figure 1 confirms that most of the off-site students move from South Italy to study in the North. Very few individuals (only 128) move from the North to the South. We count 23,084 students from the South and enroll in universities located either in the Center





NS: North to South; CS: Centre to South; SS: South to South; NC: North to Centre; CN: Centre to North; SN: South to North; CC: Centre to Centre; NN: North to North; SC: South to Centre

**Fig. 1** Migration Corridors: number of students enrolled out of region by Macro-Regions—North, Center, and South

or in the North of Italy. Also, we document that internal mobility of students<sup>8</sup> is sizable in the Center\North of Italy and modest in the South.<sup>9</sup>

The variables that we use for our estimation are:

- $G_i$ , which is a dummy variable that takes value 1 if the gender of student  $i$  is male and 0 otherwise.
- $HT_i$  is a dummy variable that takes value 1 if the high school attended by the student is a *Liceum* and 0 otherwise.
- $HG_i$  is the high school grade rescaled, see Table A.1. It is a discrete variable that measures the high school grade, which takes values in the interval [0 41]. A student enrolled in an Italian high school needs to achieve a minimum final grade of 60/100 in order to graduate.<sup>10</sup>
- $AGE_i = -1 (Year\ of\ birth - 1995)$ , which is a variable aimed at capturing late enrollment at the university. Late enrollment can be the result either of grade repetition in the high school or general late enrollment. Most of the Italian students end high school at the age of nineteen. However, some may

<sup>8</sup> We define intra-mobility as the relocation among Italian macro-regions.

<sup>9</sup> In Fig. 1 to capture the off-site status, we employ the indicator  $OR$ .

<sup>10</sup> Students may get a mention. Under this case, the grade is coded as 101.

start university earlier, given the possibility to anticipate entrance at the primary school.

According to Rosenzweig et al. (2006), two main reasons explain why students move elsewhere to complete higher education.<sup>11</sup> First, individuals move elsewhere due to the lack of higher education institutions in their home region. However, this does not apply to Italy, given that universities are evenly distributed within the country's territory. At the same time, we may expect that the percentage of off-site students is larger in better universities, as there is substantial evidence that university quality is a key pull factor of student mobility (Beine et al. 2020). Moreover, Italian universities with the best rankings are located in the Center-North of Italy. The second model explains student migration with individuals intending to move to areas where skilled labor is better paid. This model fits better the Italian experience where many individuals leave the South to join universities located in most of the Center-North area in Italy, which provides better working opportunities after graduation. We also check whether drop-out rates are different, conditioning for the area of the primary area of study, concerning the off-site status (defined here by the dummy variable *OR*). Except the area of Health,<sup>12</sup> where the percentage of droppers is only slightly lower among off-site students (5.4 for off-site and 5.1 for on-site), for the other areas, the average drop-out rate of off-site students is always considerably lower. In the area of Science, the average drop-out rate is equal to 9.6% among on-site students while equal to 12.6 among on-site students. In Social Science, the percentage of droppers is equal to 9.9 among off-site students, while among on-site students it is equal to 15.6. Finally, in Humanities, the percentage of droppers is equal to 11.5 among off-site students and is substantially larger among on-site students (16%).

Descriptive statistics seem to suggest that off-site students are a self-selected sub-population. Additional support to this hypothesis is obtained by computing the difference in means and computing the t-test. Similar results obtain if we define off-site students either using the indicator *OR* or using the indicator *OFF*<sub>150</sub>. For instance, *HG* takes a mean value equal to 20.60 among students for which the variable *OFF*<sub>150</sub> takes value 1. On the contrary, among on-site students, the value is substantially lower equal to 18.08. The difference in means highlights that among off-site students, the fraction of students who attended a *Liceum* high school is larger than for other types of high school, and the same pattern holds when we consider the age of the students with off-site students being on average younger. We also find that the percentage of female students is slightly larger among off-site students,

---

<sup>11</sup> Rosenzweig et al. (2006) deal with international students' mobility flows, but similarities with internal student mobility are easily recognizable.

<sup>12</sup> The majority of these students are enrolled in nursing degrees. In such courses, enrollment is usually allowed after passing a test organized at the local university level. Differently, nowadays, admission to medical school is conditioned to passing a test with a national ranking. Notice that our analysis considers only bachelor's degree students, disregarding students enrolled in medical studies.

which holds for all the indicators that we employ. This preliminary analysis suggests interpreting with extreme caution analysis to uncover a causal link between off-site status and drop-out behavior.

### 3 Empirical Analysis

The existing literature provides evidence that the characteristics of universities, the field of study, and the social and economic conditions of the students' home districts are correlated with drop-out rates.<sup>13</sup> Within this literature, we aim to document the relationship between distance, namely studying off-site, and drop-out rates in the case of Italian students. In order to do so, in this section, we discuss the results of our benchmark estimations complemented with several robustness checks. Then, we address the causality issues due to self-selection and omitted variables using an instrumental variable approach.

To uncover this relation, we set up the following empirical specification:

$$D_{i,u,o,f,c} = \alpha + A_u + A_f + A_o + \beta_1 G_i + \beta_2 AGE_i + \beta_3 HT_i + \beta_4 HG_i + \beta_5 OffSite_{i,t} + \varepsilon_i, \quad (1)$$

where  $\varepsilon_i$  is the error term, and we recall that  $D_{i,u,o,f,c}$  is the dummy variable that captures the drop-out decision of student,  $i$ , coming from the place of origin,  $o$ , enrolled in university,  $u$ , the field of study,  $f$ , and course  $c$ . The variables on the RHS of Eq. 1 include gender,  $G_i$ , age,  $AGE_i$ , type of high school,  $HT_i$ , and high school grade,  $HG_i$ , which were already defined.

- $A_u$ , which is a set of fixed effects that we include to control for differences in university characteristics.
- $A_f$ , which is a set of fixed effects we include to control for the different fields of study.
- $A_o$ , which is a set of fixed effects controlling for all factors specific to the home districts of students. With fixed effects, we also aim to capture differences in high school education quality among Italian districts.
- $OffSite_i$ , which is the measure of the off-site status of students. We code this variable, the focus of our analysis, in different ways:
  1.  $OD$ , which takes value 1 if the student enrolls in a university located outside the home district, and zero otherwise.
  2.  $OR$ , which takes value 1 when the student enrolls in a university located outside the home region, and zero otherwise.

---

<sup>13</sup> See Aina et al. (2018).

3.  $OFF_{km}$ , which takes value 1 when the student enrolls in a university located more than  $km$  away from her home, and zero otherwise. We consider two thresholds: 150km and 200km.
4.  $TD$ , the travel distance between the university and the student's place of residence measured in hundreds of km. We also employed  $GD$ , which is the geodesic distance between the university and the student's place of residence. One unit is equal to 100 km, results are quite similar, and we do not report the ones obtained using the latter.

A detailed table, see appendix at the end of the chapter, provides a brief description detailing definition, data source, and remarks for all the variables employed in our analysis. According to the above description, specification 1 controls for university, district of origin, and field of studies characteristics through fixed effects, as well as for other individual characteristics, for which the ANS dataset provides information including, gender, the final high school grade, the age of the individual, and the type of the high school attended.<sup>14</sup> We note that a limitation of the ANS dataset is the lack of information on both family income and parental background.<sup>15</sup> Also, we lack unambiguous information on the amount of tuition fees charged to each student.<sup>16</sup>

We obtain our baseline estimates of Eq. 1 through an OLS estimation procedure. Several reasons lead us to stick with the LPM (Linear Probability Model) as a baseline. Among others, Angrist and Pischke (2009) advocate the use of the LPM. Nonlinear estimation methods may provide an efficiency gain, but at the cost to commit to a precise distributional assumption of the error term and, notably, Probit and Logit average marginal effect estimates, quite often, do not differ much from LPM estimates and the interpretation of the regression coefficients is much more straightforward with the LPM.<sup>17</sup> Also, we evaluate the robustness of our findings to selection employing the method developed in Oster (2019). Finally, to tackle

---

<sup>14</sup> ANS differentiates university courses in 46 distinct fields of studies.

<sup>15</sup> Checchi (2000) highlights the role of both family income and parental background among the determinants of university drop-out rates.

<sup>16</sup> In Italy, tuition fees depend on several factors. Among others, we recall household income, the field of study, and the year of enrollment. In Italy, private universities are allowed to charge much higher levels of tuition fees, see Beine et al. (2020). Our fixed effects capture the heterogeneity in fees due by different universities' policies. However, we do not have specific information to the amount of tuition fees charged to each student present in the data, and to avoid losing observations, our estimations do not include such information. Modena et al. (2018) employing a similar dataset show that earning an education-grant significantly reduces early drop-out rates.

<sup>17</sup> To deal with the well-known issue of heteroskedasticity of the LPM, we employ robust standard errors.

the endogeneity of our variable capturing the off-site status, we complement our estimation results by means of an IV procedure.

## 4 Results

In what follows, we present and discuss the empirical estimates of the benchmark model described by Eq. 1. We consider all different measures of studying off-site.

Columns 1–3 of Table 1 report the estimation results when we use the dummy variables  $OD$  and  $OR$  to measure the off-site status of the students.<sup>18</sup> The drop-out rates are negatively correlated with the high school grade, with the age of the student, with being a woman, and with a diploma from a *Liceum*. Interpreting our coefficient estimates as marginal effects, we find that, *ceteris paribus*, one additional point in the high school grade reduces the probability of dropping out by 0.4%. Being graduated in a *Liceum* is correlated with a reduction of drop-out by 10%. Concerning the correlation between drop-out and being an off-site student, we find a significant negative sign. When we employ  $OD$ , we find that the off-site status is associated with a 1.25% reduction of the probability of dropping out. When we proxy the off-site status with the dummy  $OR$ , the estimated correlation becomes stronger neither the sign nor the magnitude of any of the other coefficients changes across the two specifications. A comparison of columns (1)–(3) of Table 1 shows that our estimates are robust to different measurements of the off-site status.

As pointed out in the introduction, for many students, the home district does not host any university, so that the only option is to leave the district to pursue a university education. Specifically, this implies that for students coming from 52 out of the 110 Italian districts, the dummy,  $OD$ , always takes one as value. In that respect,  $OR$ , which is based on regions, provides a more conservative definition of the off-site status. Still, both  $OR$  and  $OD$  might not be meaningful measures of the off-site status for various reasons. For instance, using either  $OR$  or  $OD$ , we might end up classifying them as off-site students who enroll in universities that, while located in a different district or region, might be geographically very close to their home location close enough to allow for daily commuting. Therefore, we also consider alternative measures of the off-site status based on travel and geodesic distance between the student's home and the student's university. Specifically, in columns 4 and 5 of Table 1, the dummy variables  $OD$  and  $OR$  are replaced with the continuous variable  $TD$ , respectively, where  $TD$  is the travel distance.<sup>19</sup> Column 4 of Table 1 suggests that a 100 km increase in the average travel distance is associated with a 0.3% reduction in the probability of drop-out. In Column (5) of Table 1, we also report the results for a regression model that include the

---

<sup>18</sup> Johnes and McNabb (2004) and Bussu et al. (2019) employ similar indicators.

<sup>19</sup> Similar results, available upon request, are obtained when we employ geodesic distance in place of travel distance.

**Table 1** Determinants of drop-out rates. Benchmark (1)

	(1)	(2)	(3)	(4)	(5)
$HG_i$	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)
$Age_i$	0.0098*** (0.000)	0.0098*** (0.000)	0.0098*** (0.000)	0.0097*** (0.000)	0.0097*** (0.000)
$HT_i$	-0.1034*** (0.001)	-0.1030*** (0.001)	-0.1038*** (0.002)	-0.1038*** (0.002)	-0.1038*** (0.002)
$G_i, M = 1$	0.0268*** (0.002)	0.0269*** (0.002)	0.0267*** (0.002)	0.0268*** (0.002)	0.0267*** (0.002)
$OD_i$	-0.0125*** (0.002)				
$OR_i$		-0.0162*** (0.002)	-0.0161*** (0.002)		
$TD_{u,o}$				-0.0033*** (0.0000)	-0.0078*** (0.0000)
$TD_{u,o}^2$					-0.000005*** (0.00000)
University fixed effects	Yes	Yes	Yes	Yes	Yes
Field fixed effects	Yes	Yes	Yes	Yes	Yes
Region fixed effects	No	Yes	No	No	No
District fixed effects	Yes	No	Yes	Yes	Yes
$R^2$	0.0917	0.0917	0.0927	0.0925	0.0926
$N$	226,094	226,094	226,094	226,094	226,094

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . OLS estimates. Robust standard errors in parentheses

square of distance. Including this variable, we test the hypothesis of a nonlinear relationship, and we find that the marginal effect of distance diminishes with the distance. Finally, we also report the results we obtain measuring the off-site status with the dummy variable  $OFF_{km}$ . We consider two specifications of this indicator:  $OFF_{150}$  and  $OFF_{200}$ . Notice that  $OFF_{150}$  and  $OFF_{200}$  take value equal to 1 if the student is enrolled in a university more than 150 and 200 km distant from her home, respectively. Table 2 reports the empirical estimates obtained using these two measures of studying off-site.

Two results stand out from Table 2. First, the magnitude of the coefficients capturing the off-site status is strikingly close to the one delivered by the empirical estimate of  $OR$ , see Table 1. Also, we notice that the magnitude of the coefficient  $OFF_{200}$  is smaller than the one of  $OFF_{150}$ .

To summarize, all our measures of studying off-site confirm a strong negative and significant correlation between the drop-out decision and off-site status. The estimates of the other variables of interest are in line with the findings in the literature. Women show a lower propensity to drop out. Also, there is evidence that older individuals tend to leave university more frequently and that the high school

**Table 2** Determinants of drop-out rates. Benchmark (2)

	(1)	(2)
<i>HG</i>	-0.0040*** (0.000)	-0.0040*** (0.000)
<i>AGE</i>	0.0098*** (0.000)	0.0098*** (0.000)
<i>HT</i>	-0.1037*** (0.002)	-0.1038*** (0.002)
<i>G, M = 1</i>	0.0268*** (0.002)	0.0268*** (0.002)
<i>OFF</i> <sub>150</sub>	-0.0195*** (0.003)	
<i>OFF</i> <sub>200</sub>		-0.0165*** (0.003)
University fixed effects	Yes	Yes
Field fixed effects	Yes	Yes
District fixed effects	Yes	Yes
<i>R</i> <sup>2</sup>	0.0926	0.0925
<i>N</i>	226,094	226,094

\**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001. OLS estimates. Robust standard errors in parentheses

grade negatively correlates with drop-out rates, with students that earned a better high school grade eventually dropping out less.<sup>20</sup> Finally, students who attend a *Liceum* tend to drop less than students coming from the vocational schools.

Our findings concerning the off-site status can be questioned on several grounds. First, we evaluate whether the correlations reported in Tables 1 and 2 remain stable independently of the home macro-area of the off-site students. Accordingly, we run regressions clustering students depending on their home macro-area. We consider three macro-areas: North, Center, and South of Italy. To capture the off-site status, we use two indicators: *OFF*<sub>150</sub> and *OR*. Table 3 shows that the magnitude of our proxy varies substantially once we consider regressions by macro-area.

The use of *OR* or *OFF* yields almost identical results. Interestingly, the off-site status of the students is not significantly associated with drop-out when we run the regressions considering only students from the North of Italy. Also, it is important to notice that the magnitude of the *HG* coefficient is larger, in absolute value, for the sub-population of students from the South. Remarkably, the coefficient of *HG* is almost identical when we run regressions separately for Center and Northern students.

Several reasons may explain the lack of significance of both *OR* and *OFF* coefficients for the sample of North students. One for all, the vast majority of off-

<sup>20</sup> Notice that Belloc et al. (2010) found a positive correlation between high school grade and drop-out rates.

**Table 3** Determinants of drop-out rates: estimates by macro-area

	(South)	(Center)	(North)	(South)	(Center)	(North)
<i>HG</i>	-0.0046*** (0.000)	-0.0037*** (0.000)	-0.0036*** (0.000)	-0.0046*** (0.000)	-0.0037*** (0.000)	-0.0036*** (0.000)
<i>Age</i>	0.0112*** (0.001)	0.0090*** (0.000)	0.0096*** (0.000)	0.0112*** (0.001)	0.0090*** (0.000)	0.0096*** (0.000)
<i>HT</i>	-0.1071*** (0.003)	-0.1018*** (0.003)	-0.1003*** (0.002)	-0.1072*** (0.003)	-0.1019*** (0.003)	-0.1003*** (0.002)
<i>G, M = 1</i>	0.0285*** (0.003)	0.0270*** (0.003)	0.0248*** (0.002)	0.0286*** (0.003)	0.0270*** (0.003)	0.0248*** (0.002)
<i>OFF</i> <sub>150</sub>	-0.0243*** (0.006)	-0.0183*** (0.005)	-0.0068 (0.004)			
<i>OR</i>				-0.0311*** (0.009)	-0.0185*** (0.005)	-0.0067 (0.004)
<i>R</i> <sup>2</sup>	0.1074	0.0888	0.0835	0.1073	0.0889	0.0835
<i>N</i>	77,238	67,850	80,929	77,238	67,850	80,929
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Field fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . OLS estimates. Robust standard errors in parentheses

site students from the North opt to enroll in a university still located in the North and therefore at a short distance from the student's home. Maybe distance from home is so short that it does not affect students' life in a particular way, and therefore it does not affect their performance.

Also, the literature often estimates drop-out determinants through nonlinear models.<sup>21</sup> As a robustness,<sup>22</sup> we compute the marginal effects by estimating a Logit specification of Eq. 1. We find that the estimated marginal effects do not change significantly when we employ a Logit specification in place of our benchmark LPM. In line with previous results, we obtain negative and significant coefficients for our measures of off-site status. In the introduction, we highlighted how the possibility of self-selection and omitted variables induce particular caution in interpreting our results; thus, this analysis does not allow interpreting the partial correlation between off-site status and drop-out as evidence of a causal relationship.

To evaluate the role of selection on unobservables, we employ the procedure outlined in Oster (2019). Two reasons may explain the negative correlation between off-site status and drop-out rates: (1) *selection*, the best and the brightest leave

<sup>21</sup> For examples we refer the reader to Belloc et al. (2010) and Zotti (2015).

<sup>22</sup> Results available upon requests.



their hometown to get higher education and (2) *omitted variable bias*, our off-site indicators are absorbing the role of omitted variables, such as family income. The method outlined in Oster (2019) assumes that the relationship between observables and the treatment is informative of the relationship between treatment and unobservables. Therefore, we assume that  $HG$ ,  $HT$ , and  $AGE$  are similarly related with the treatment, the off-site status, as the observable. More clearly, in our estimation, the unobservable includes parents' education, family income, and unobserved ability.<sup>23</sup> The implementation of the Oster (2019) method confirms previous results, suggesting that the off-site status affects drop-out behavior. If selection on unobservables has the same strength as the selection of observables, our estimate of the off-site status coefficient is only slightly reduced. Selection on unobservables should be at least five times stronger than selection on observables to make the relationship between the off-site status and the drop-out behavior negligible.

## 5 Causality: Instrumental Variable Approach

The descriptive evidence previously discussed suggests that off-site students' sub-population is a self-selected group with systematic characteristics different from the overall population. Due to the possibility of self-selection and omitted variables, interpreting the evidence from the regression models already presented is problematic.

In other words, the evidence of a strong negative correlation between drop-out rates and off-site status does not legitimate us to conclude anything about the causality direction of that relationship due to both unobservables and self-selection.

Notably, the decision to study far from home implies sunk costs, both monetary and non, which, see Checchi (2000), affect students' effort. Off-site students leave back home both family and friends, need to get used to the new city social norms and, last but not least, a substantial monetary investment is required (think about rent of the room/apartment, transportation cost). These costs are likely to be positively correlated with distance. As Checchi (2000) shows, students' effort is sensitive to monetary costs in general, and those studying off-site may exert more effort in their studies because in the event of dropping out, the sunk cost is higher compared to the ones faced by on-site students. Also, Garibaldi et al. (2012) show that an increase in tuition fees reduces late graduation providing evidence that students' effort depends on investments incurred.<sup>24</sup> Besides, among off-site students, there

---

<sup>23</sup> The Oster (2019) method requires to set a value of the  $R^2$  that the model would have attained whether all predictors were available. Following the literature, we set this value 1.3 and 2.2 times higher of the  $R^2$  that we got from our estimations.

<sup>24</sup> This paper does not find similar evidence for drop-out rates. However, it only considers students enrolled in one of the most expensive Italian private universities.

may be heterogeneity concerning the sunk cost. We may have type 1 students, with higher ability and motivation, who choose to study off-site to enroll in a better university, and type 2 students, from high-income households, who may choose to study off-site merely because they can afford it, although equipped with average (or below average) motivation and ability. For type 1 students, the decision to study off-site is driven by motivation. For type 2 students, it is driven by family wealth. It is evident that both motivation and wealth are negatively correlated to drop-out. As motivation and family income fell in the unobservable component in specification 1, we are not able to say whether the negative correlation between drop-out rates and offsite status is fostered by the link between higher costs and motivation or between higher costs and family wealth or by both.

The above discussion suggests the need of an appropriate estimation strategy to address the bias that self-selection along with omitted variables generates.<sup>25</sup> Following Card (1993), we exploit information on the distance from the closest university to construct an instrument for the off-site status. For each student, we determine the distance from her place of residence to the closest university. Taking advantage of this information, we identify two possible instruments:<sup>26</sup>

1. The distance from the closest university, which we label *minD*
2. A dummy variable that we set equal to one if the closest university is distant more than 20 km from the student's place of residence.<sup>27</sup> We label this instrument *dD*.

We acknowledge that there are some arguments that question the validity of our instrument, similar to the one mentioned in Card (1993) and Card (2001).<sup>28</sup> First, we collect some evidence on the validity of the exclusion restriction. Subsequently, we present and discuss our IV estimates.

---

<sup>25</sup> Focusing on the self-selection, one may suggest estimating the model with an Heckman type correction model. We prefer to stick to an IV procedure. By doing so, the validity of our estimates does not rely on any assumption concerning the distribution of the error term, see Angrist and Pischke (2009).

<sup>26</sup> Further research may build new instruments developing measures of spatial competition for each degree program, see Bratti et al. (2021).

<sup>27</sup> When using this instrument, one may be prone to suggest to run a Probit in place of an OLS in the first stage. Angrist and Pischke (2009) and Wooldridge (2010) shows that this procedure would be incorrect, namely we would run a kind of *forbidden regression*. Differently, another feasible alternative would be a bivariate probit. However, our rich structure of fixed effects generates collinearity issues. Therefore, we consider solely estimation obtained only through a two-stage least squares procedure.

<sup>28</sup> Typically one may argue the validity of the exclusion restriction saying that when deciding where to settle households internalizes the offsprings' decision of whether to enroll at the university. However, in Italy, the mobility of households is minimal, with individuals showing a very low propensity to move once settled.

## 5.1 Exclusion Restriction and Reduced Form

Our model is just identified, thus preventing us from performing the Sargan–Hansen to check whether the correlation among the error term and the instrument are statistically not different from zero. Despite the impossibility of performing the overid test, we can check how *minD* correlates with the other drop-out determinants to evaluate the exclusion restriction assumption. A good instrument should not be correlated with strong determinants of the dependent variable. Our instrument *minD* is almost uncorrelated with the determinants of the drop-out rate previously discussed (*HT*, *HG*, and *Age*). Also, we check for the reduced-form estimates. We compute such regressions for both our instruments, *minD* and *dD*. Our reduced-form estimates are both negative and slightly statistically significant.

## 5.2 IV: Results

Table 4 reports our empirical estimates, where we instrumented the measures of off-site status.

Table 4 reports our empirical estimates, where we instrumented the measures of off-site status. First-stage estimates confirm that our instruments are strong.<sup>29</sup> Column (1) instruments the dummy variable *OR* with the instrument *dD*. Notice that the sign of the *OR* coefficient is still negative and significant and increases in magnitude compared to the OLS estimation with no instrumental variables.

Significantly, the standard errors increase, a typical consequence of the IV procedure. Column (2) reports similar findings. Here, we instrument *OR* with the actual minimum distance, *minD*. Column (3) and column (4) report results when we employ the variable *TD* as a proxy of the off-site status. Notice that the magnitude and the sign of all the other control variables stay almost unchanged as we vary either the instrument or the variable measuring the off-site status. In conclusion, we notice that the coefficients on distance lose statistical significance for all cases, which may be due to the lower precision implied by IV estimation. To check whether it is sensible to run the IV procedure, we report the Wu–Hausman test. The null hypothesis is that both estimators, OLS and IV, are consistent. We do not obtain strong evidence for the non-consistency of the OLS estimates. However, even if we fail to reject the null hypothesis, the test does not allow us to claim that the OLS estimates are consistent. Hence, such values of the WU–Hausman test do not invalidate our IV estimates. Indeed, this situation is typical when the standard error of the IV estimator is large as it is for Table 4 estimates. In Columns (5) and (6), we use as a proxy of the off-site status the variable *OFF*<sub>150</sub>, while columns (7)

---

<sup>29</sup> The value of the *F* statistics is always larger than 104, as suggested in Lee et al. (2020).

**Table 4** IV estimates (1)—drop-out decision: instrumented indicators:  $OR$ ,  $TD$  and  $OFF_{km}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$HG_i$	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)	-0.0040*** (0.000)
$AGE_i$	0.0099*** (0.000)	0.0099*** (0.000)	0.0098*** (0.000)	0.0098*** (0.000)	0.0099*** (0.000)	0.0099*** (0.000)	0.0099*** (0.000)	0.0099*** (0.000)
$HT_i$	-0.1025*** (0.002)	-0.1027*** (0.002)	-0.1037*** (0.002)	-0.1037*** (0.002)	-0.1026*** (0.002)	-0.1030*** (0.002)	-0.1024*** (0.002)	-0.1029*** (0.002)
$G_i, M = 1$	0.0265*** (0.002)	0.0266*** (0.002)	0.0268*** (0.002)	0.0268*** (0.002)	0.0267*** (0.002)	0.0268*** (0.002)	0.0269*** (0.002)	0.0269*** (0.002)
$OR_i$	-0.1068** (0.053)	-0.0906* (0.048)						
$TD_i$			-0.0093** (0.005)	-0.0069* (0.004)				
$OFF_{150}$					-0.0941** (0.046)	-0.0632* (0.034)		
$OFF_{200}$							-0.1294** (0.064)	-0.0903* (0.048)
Univ fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV: dummy (> 20 km)	Yes	No	Yes	No	Yes	No	Yes	No
IV: distance	No	Yes	No	Yes	Yes	No	Yes	No
F first stage	502	547	4969	6236	717	1024	494	686
Hausman test	2.99	2.38	1.86	1.09	2.61	1.69	3.16	2.36
$N$	226,094	226,094	226,094	226,094	226,094	226,094	226,094	226,094

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

IV estimates

Robust standard errors in parentheses

and (8) employ  $OFF_{200}$ . Notice that the results reported in Table 4 slightly change depending on the indicator used.

The exclusion restriction of our IV might be questioned on several grounds. Our estimations control for several drop-out determinants, the high school grade, the type of the high school attended, age, and gender. However, we already acknowledge that we lack information on some determinants such as family income. Furthermore, it may be that households that give more weight to education have a larger propensity to live closer to a university.

We resort to the method proposed by Conley et al. (2012) to account for possible deviations from the exclusion restriction. This method allows considering the parameter capturing the exclusion restriction (the IV's coefficient in the structural equation) as a random parameter drawn from a given distribution. Also, the method allows considering asymmetric deviations from the exclusion restriction. We employ the method labeled *Union of Confidence Intervals* taking advantage of the STATA routine developed in Clarke and Matta (2018). We conducted this robustness with the instrument  $dD$  employing as an indicator of the off-site status the dummy variable  $OFF_{150}$ . As long as the interval is sufficiently tiny, our estimates remain statistically significant, and the coefficient's magnitude is only slightly affected.<sup>30</sup> Once we consider wider intervals for the parameter capturing the exclusion restriction, our estimates lose statistical significance. Furthermore, this procedure allows for assessing the instrument's validity when the degree of the over-identification is not positive.

Our previous findings suggest that the impact of the off-site status on drop-out rates is much stronger among students coming from the South. In Table 5, we report IV estimation clustering individuals along the home macro-areas. We employ the variable  $OFF_{150}$  as a proxy of the off-site status, which we instrument using  $minD$ . Column (1) considers only students from the South. It reports a highly significant and negative estimate for the off-site measure,  $OFF_{150}$ . This suggests that, once we account for the selection effect, for a southern student, going off-site has a considerable impact on the decision of not leaving the university. Interestingly, this does not happen to be the case for students originating from the Center and the North of Italy, for whom we do not find any significant impact of the off-site status on the decision to drop out. Our results are in line with the model and empirical findings of Checchi (2000). Students moving from the South to the North face larger sunk cost. Large sunk cost appears to have eventually a positive effect in the decision to not drop out. Similar evidence is not obtained once we consider separately students originating either from the Center or from the North. Most of them attend universities located in the same area, and therefore they face lower sunk cost and, as our estimate suggest, the positive effect on the drop-out decision eventually does not materialize.

So far, our interpretation of our IV estimates builds on the basic homogeneous treatment effect framework. However, in the more general case of heterogeneous

---

<sup>30</sup> Results available upon request.

**Table 5** IV estimates, by macro-area (3)

	(South)	(Center)	(North)
<i>OFF</i> <sub>150</sub>	-0.2507*** (0.069)	0.1794 (0.134)	-0.0754 (0.091)
<i>HG</i>	-0.0046*** (0.000)	-0.0037*** (0.000)	-0.0036*** (0.000)
<i>Age</i>	0.0114*** (0.001)	0.0087*** (0.000)	0.0096*** (0.000)
<i>HT</i>	-0.1070*** (0.003)	-0.1046*** (0.003)	-0.0992*** (0.003)
<i>G, M = 1</i>	0.0282*** (0.003)	0.0270*** (0.003)	0.0248*** (0.002)
University Fixed effects	Yes	Yes	Yes
Field Fixed effects	Yes	Yes	Yes
District Fixed effects	Yes	Yes	Yes
N	77238	67850	80929
First Stage	577	179	109
Hausman Test	11.14	2.23	0.57

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

IV estimates

Robust standard errors in parentheses

treatment effects, the IV estimates only capture the *LATE*, local average treatment effect, the impact of studying off-site on the sub-population of compliers. In our case, compliers are individuals who enrolled in an off-site university because there was no university close to their place of residence. Notice that our IV estimates are substantially larger than the OLS ones. However, one may argue that after taking into account endogeneity issues, we should uncover, at least, an estimate with a lower magnitude. In our benchmark estimations, the off-site status was absorbing the impact of variables negatively correlated with our outcome variable (i.e., parents' income, individuals' ability). The same counterintuitive effect materializes in Card (1993); the impact of education on wages gets larger once endogeneity issues are tackled. However, it is legitimate to expect a larger effect of the off-site status on the outcome with a heterogeneous treatment effect. *Compliers* should come, on average, from families with a lower average income than the rest of off-site students. *Ceteris paribus*, families with low income, incur a relative higher education cost, leading off-site students to think twice before dropping university and putting more effort into their studies. Notably, this interpretation accounts for the substantial difference obtained once we separate estimations clustering individuals by macro-area of origin.

### **BOX: Course Heterogeneity, Selection Bias, and Drop-Out**

As widely discussed, dropping out of university has a course- and student-specific causes. The former include those typical for each degree, which may require a relatively different level of effort. Student-specific causes include, for example, the student's own abilities, the financial viability of those who finance the studies, and the impact on the effort that studying off-site may generate. Moreover, drop-out can be largely influenced by the mismatch between the student's abilities and his/her most suitable degree. For these mismatches to have a significant effect, they have to be systematic. A possible explanation for why students may systematically make such misjudgments about the adequacy of their abilities with the skills and knowledge required by a degree is, for instance, that some degrees offer more job opportunities and that students may follow a herding behavior. If such students target more frequently some type of degree than others, then the mismatch between skills and motivation affects the target degrees more than the others. Among the target degree, the drop-out rate may result exceptionally higher due to the negative self-selection effect.

We use the sample selection approach to account for the correlation between unobserved heterogeneity in the enrolment decision (selection) and unexplained factors driving drop-out (outcome). We rely on 16 cohorts of students from the Università di Sassari enrolled in degrees supplied by ten departments to investigate this aspect. The cohorts allow monitoring students enrolled in the same year to determine who drops out from the Università di Sassari. Considering one university, a student leaving the degree between the first and the second year is a drop-out, although we cannot exclude that droppers enroll to other universities. Since we do not have any direct measure of popularity of departments, we label as *popular* the departments with relatively more students, as they attract an above-average number of students. The observations in each cohort are merged into a single pool of 57,974 observations. We choose the ten departments as the observation unit, and we use this criterion to cluster the data.

Across the 16 cohorts, five departments (Architecture, Agricultural Science, Biomedical Science, Chemistry, Pharmacy, and Veterinary Medicine) have an enrolment rate below the university average (10%). Architecture and Veterinary Medicine are one standard deviation below the average enrolment rate, while Agricultural Science, Biomedical Science, Chemistry, and Pharmacy are close to the university mean. The other five departments (Economics, History, Humanities, Law, and Medical Science) have an enrolment rate above average. Law and History are one standard deviation above the mean.

We hypothesize that drop-out is affected by the mismatch between student ability and motivation and those required in each degree. Popular degrees, with an above-average number of students, may attract relatively more individuals with low motivation. As motivation is unobservable, this determines a negative self-selection effect because in these degrees, less motivated students are overrepresented.

We estimate the probability of drop-out, i.e., to leave a degree course between the first and the second year of enrolment, by estimating Eq. 2, where the variable  $drop_i$  is a dummy variable that takes value 1 for droppers.

$$\begin{aligned}
 drop_i = & \beta_1(\text{final high school grade})_i + \beta_2(\text{ECTS credits})_i \\
 & + \beta_3(\text{exempted from tuition})_i \\
 & + \beta_4(\text{years from college graduation to university enrollment})_i \\
 & + \beta_5(\text{tuition fees})_i \\
 & + \beta_6(\text{lyceum})_i + \beta_7(\text{technical vocational high school})_i \\
 & + \beta_8(\text{training high school})_i + \beta_9(\text{woman})_i \\
 & + \beta_{10-35}(\text{cohorts fixed effects})_i + \epsilon_i
 \end{aligned} \tag{2}$$

After estimating Eq. 2 we compute the marginal predicted probability of drop-out for the whole university sample and separately for each of the ten departments. In each cohort, we average the individual marginal probability to obtain a mean by cohorts and departments. We compare these probabilities with the average marginal predicted probabilities of drop-out obtained estimating the following bivariate probit with selection defined by Eqs. 3 and 4. Equation 3 is the selection equation (choice of the department), while Eq. 4 is the drop-out equation.

$$\begin{aligned}
 department_i^k & \\
 = & \alpha_1(\text{final high school grade})_i + \alpha_2(\text{beneficiary of scholarship})_i \\
 & + \alpha_3(\text{enrolled first time})_i \\
 & + \alpha_4(\text{years from college graduation to university enrollment})_i \\
 & + \alpha_5(\text{year of birth})_i + \alpha_6(\text{woman})_i + \alpha_7(\text{lyceum})_i \\
 & + \alpha_8(\text{technical vocational high school})_i + \alpha_9(\text{training high school})_i \\
 & + \alpha_{10}(\text{number of enrollments})_i + \alpha_{11}(\text{tuition fees})_i + \epsilon_{1,i},
 \end{aligned} \tag{3}$$



$$\begin{aligned}
drop_j^k = & \gamma_1(\text{final high school grade})_j + \gamma_2(\text{ECTS credits})_j \\
& + \gamma_3(\text{exempted from tuition})_j \\
& + \gamma_4(\text{years from college graduation to university enrollment})_j \\
& + \gamma_5(\text{tuition fees})_j \\
& + \beta_6(\text{lyceum})_j + \beta_7(\text{technical vocational high school})_j \\
& + \gamma_8(\text{training high school})_j + \gamma_{9-34}(\text{cohorts fixed effects})_j + \epsilon_i.
\end{aligned}
\tag{4}$$

The model is estimated employing maximum likelihood.<sup>a</sup> The outcome Eq. 4 is estimated for all the  $k = 1, \dots, 10$  departments. The estimation results for each of the ten departments (not reported) show that, once we account for selection, the average marginal predicted probability of drop-out in the five departments with above-average enrolment rate is systematically below the one computed employing the standard probit. For the least *popular* departments, Architecture, and Veterinary Medicine, results are as expected, i.e., that predicted probability considering selection is way above the one resulting from the standard probit. Biomedical Science, Chemistry and Pharmacy, and Agricultural Science (see Fig. 2 in the Appendix), which have an enrolment rate close to the university average, follow a pattern similar to the *popular* departments.

Remarkably, predicted probabilities with and without selection tend to be similar for Medical Science. Note that this is the only department during the sample period in which students have to pass a national-based test to enroll. It seems that the selection process prevents students with below-average motivation and skills from enrolling in this department.

The selection of the degree may also affect the magnitude, significance, and sign of estimated parameters. In some cases, it helps to uncover effects that are confounded because one variable may positively affect the department's selection and negatively the drop-out, or vice versa. This is particularly interesting for the case of gender. In our estimation, the parameter of the dummy woman ( $\alpha_9$  in Eq. 2) is negative and significant for the whole sample and for all departments but Architecture (positive but not significant). We cannot say that this result depends on the fact women choose more likely departments with lower drop-out rates or that women are better students, thereby reducing drop-out when they are numerous. Descriptive statistics do not suggest a clear-cut. Indeed, women are relatively underrepresented in the department where drop-out rate is higher (Economics and Law), but they are also overrepresented in departments where drop-out rate is still high (History and Humanities). We cannot say whether is drop-out that causes the gender

mix in a department or the opposite. However, we can compute the marginal contribution of a gender on the selection and that of the selection on drop-out.

Selection estimation is used to compute the marginal contribution on drop-out of an additional woman who decides to enroll in a department. To this purpose, we compute the marginal effect for the dummy woman on the conditional probability of drop-out. The change in conditional probability due to women is the change of the ratio between the joint probability and the marginal probability due to a discrete change of the dummy woman included in the selection equation:

$$\frac{\partial \text{Prob}(\text{drop}_k = 1 | k = 1)}{\partial \text{woman}} = \frac{\partial [\text{Prob}(\text{drop}_k = 1, k = 1) / \text{Prob}(k = 1)]}{\partial \text{woman}} \quad (5)$$

for all  $k = 1, \dots, 10$  departments.

Note that we include the dummy woman in the selection equation only. A positive sign of the dummy woman means positive selection and positive effect on the marginal probability of choosing department  $k$ , the denominator of conditional probability. If we obtain a positive marginal effect on the conditional probability of drop-out, the joint probability is positive, i.e., women contribute positively to the joint event drop-out and department  $k$  selected. We interpret this as a positive contribution of women to the probability of drop-out in that department. A negative marginal effect on the conditional probability suggests the opposite.

In case of negative selection, results are reversed. A positive marginal effect on the conditional probability of drop-out means that the joint probability is negative. On the contrary, a negative marginal effect on the conditional probability means women contribute positively to the joint event.

We classify the above results as follows. For the cases of positive selection:

- i.  $\frac{\partial \text{Prob}(\text{drop}_k=1|k=1)}{\partial \text{woman}} > 0$ , more women, more drop-out
- ii.  $\frac{\partial \text{Prob}(\text{drop}_k=1|k=1)}{\partial \text{woman}} < 0$ , more women, less drop-out

for the cases of negative selection

- iii.  $\frac{\partial \text{Prob}(\text{drop}_k=1|k=1)}{\partial \text{woman}} > 0$ , less women, less drop-out
- iv.  $\frac{\partial \text{Prob}(\text{drop}_k=1|k=1)}{\partial \text{woman}} < 0$ , less women, more drop-out

selection:

Our dataset has 6 departments with positive selection (Humanities, History, Veterinary Medicine, Medical Science, Biomedical Science, Chemistry, and Pharmacy). In Veterinary Medicine, the marginal probability and the conditional of drop-out for women are not significant. Medical science is an

example of case i. Although the dummy woman is not significant in the single probit, we uncover a positive contribution of women on drop-out. The other five departments fall in case ii., excluding Chemistry and Pharmacy, for which the dummy woman is not significant in the selection equation.

The remaining four departments (Economics, Agricultural Science, Law, and Architecture) exhibit a negative selection. The first three fall in case iii., while Architecture is an example of case iv., although in the single probit, the dummy woman is not significant.

We conclude that women contribute to increasing the drop-out rate in Medical Science and Architecture, although both marginal effects are very small. On the contrary, women reduce drop-out rates in Humanities, History, Economics, Agricultural Science, Biomedical Science, and Law. There is no evidence of any contribution to drop-out of women in Veterinary science, Chemistry, and Pharmacy.

---

<sup>a</sup> Notice that the set of regressors differs between Eqs. 3 and 4, and our seemingly unrelated probit captures the correlations between the choice of the course and the drop-out behavior, allowing us to compute the marginal effect relevant for our analysis. The SUR approach prevents us to incur in the identification issues raised in Maddala (1983) and Li et al. (2019).

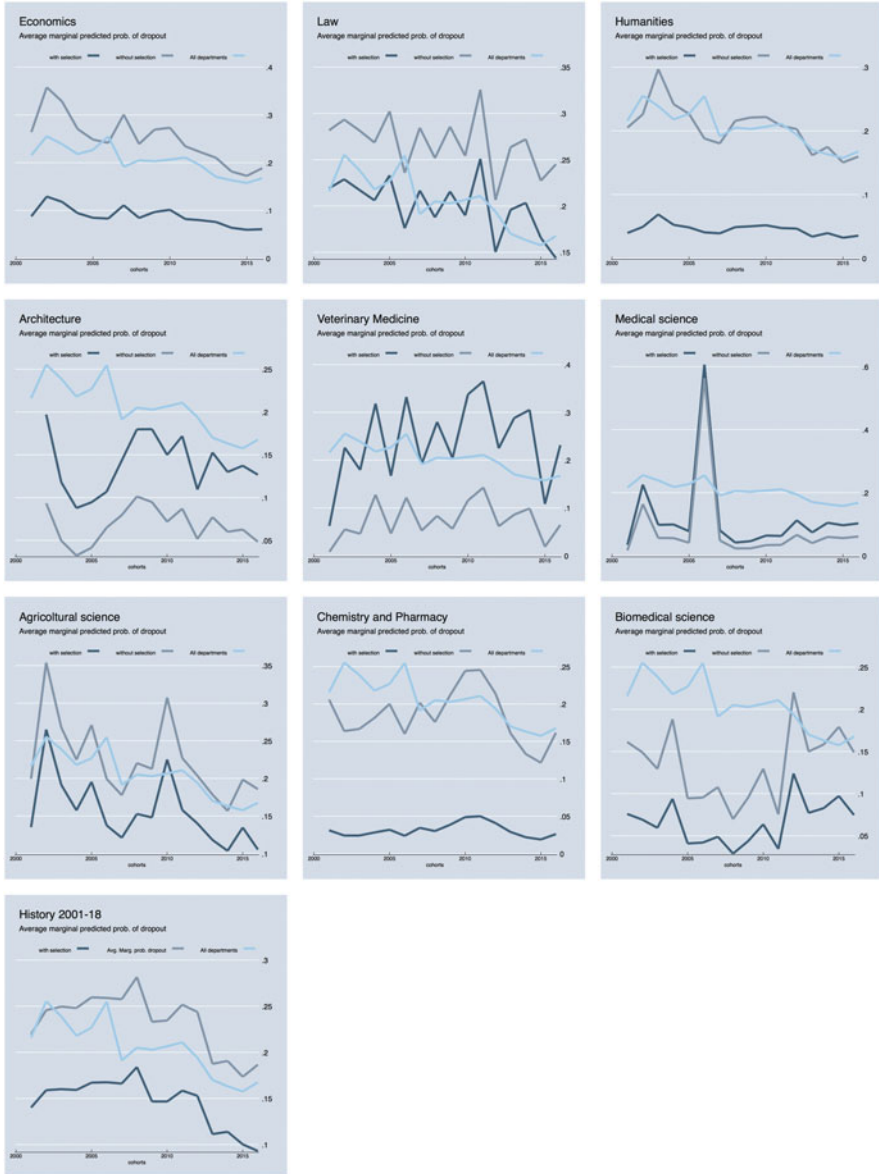


Fig. 2 Predicted probabilities of drop-out by departments

## 6 Conclusions

In this chapter, we investigated the determinants of the drop-out decision in the population of students enrolled at the public university system in Italy. We document that off-site students, who left home to pursue university, are a self-selected population for various characteristics that are candidate determinants of the drop-out decision. Then, we show a robust and strong negative correlation between the likelihood of dropping out of university and the off-site status of students. To go beyond correlation and assess the causal link between the off-site status and the decision to drop out, we employ an instrumental variable approach. The estimates provide strong evidence that off-site status reduces the likelihood of dropping out among southern students, who typically study in universities in the Center-North of Italy. The negative effect is still present considering the whole population of students, although lower in magnitude and barely significant. Our findings have relevant policy implications.

First, due to the documented sizable self-selection, our estimates suggest that it is not fair to rank university quality through a naive comparison of drop-out rates. We produce abundant evidence that a significant fraction of the best southern students moves to complete higher education at institutions located in the Center-North of Italy. On the contrary, the flow of students from the Center-North to the south is negligible. Our empirical results suggest that self-selection among off-site status explains part of the sizable difference in drop-out rates between northern and southern institutions. Second, our result suggests that universities aiming to improve the quality of their students' pool shall set policies to attract off-site students.

We address whether there is any causal relationship between off-site status and drop-out behavior. We conduct our analysis taking advantage of the instrumental variable approach. We employ as an instrument of our off-site indicators variables capturing the proximity from the closest university. Our results show that, especially for off-site students originating from the south, there is substantial evidence that going off-site reduces the likelihood of dropping out of university. In line with Checchi (2000) we argue that studying off-site by requiring substantial investments (not only monetary ones), eventually positively impact the students' effort.

However, we are aware of some shortcomings of our IV approach. Although our sample is large, our IV estimates provide strong evidence for an effect of the off-site status on drop-out rates only for the subset of southern students. To conclude, we acknowledge the limitation of our IV exercise, calling for further research to determine better both the magnitude and significance of the relationship between off-site and drop-out status.

Our analysis that exploits detailed data from the Università di Sassari highlights that, without taking into account selection, it is not sensible to naively compare drop-out rates among different departments. In addition, we shed light on the marginal contribution of women on drop-out rates, showing that the estimated parameter for women in a univariate probit model is not informing on this issue.

## Appendix

The table below provides a detailed description of each variable employed in the main analysis:

**Table A.1** Data sources and definitions

Variable	Definition	Source	Remarks
Drop-out ( $D_{(i,u,c,o)}$ )	dummy variable that takes one when the student drop out from the course/university and zero otherwise	ANS data, our computation	$i$ identifies the individual, $u$ , the university, $c$ the field of study, $o$ the origin of the students
$HG_i$	variable capturing the High school grade of student $i$	ANS data	The minimum grade to obtain a high school title in Italy is equal to 60 with the maximum equal to 100 (however, students may obtain a mention). We scale subtracting 60 to each vote
$AGE_i$	$AGE_i =$ $-1 (Yearofbirth_i - 1995)$	ANS data, our computation	Notice that in Italy students usually finish high school at the age of 19
$HT_i$	dummy variable that captures the type of the high school attended by student $i$	ANS data	The variable takes value equal to one only if the high school is a <i>Liceo</i> of the traditional type, either <i>Classico</i> or <i>Scientifico</i> . For all the rest of high schools, the variable is set equal to zero
$G_i$	dummy variable that captures the gender of the student $i$ . Takes value 1 for man and 0 otherwise	ANS data	
$OD_{i,u,o}$	dummy variable that takes value 1 when the students enrolls in a university not located in his/her district of residence	ANS data	
$OR_{i,u,o}$	dummy variable that takes value 1 when the students enrolls in a university not located in his/her region of residence	ANS data	
$TD_{i,u,o}$	measures the distance between the student $i$ place of residence, $o$ and the university of destination $u$	ANS data, our computation	Our computation employing the routine developed by Weber and Péclat (2017), one unit is equal to 100 km

(continued)

**Table A.1** (continued)

Variable	Definition	Source	Remarks
$OFF_{150\ i,u,o}$	dummy variable that takes value 1 when the student enrolls in a university distant more than 150 km, in term of travel distance, from his/place of origin	ANS data, our computation	
$OFF_{200\ i,u,o}$	dummy variable that takes value 1 when the student enrolls in a university distant more than 150 km, in term of travel distance, from his place of origin	ANS data, our computation	

## References

- Acemoglu, D. (2002). Directed technical change. *The Review of Economic Studies*, 69(4), 781–809.
- Aina, C., Baici, E., Casalone, G., & Pastore, F. (2018). The Economics of University Dropouts and Delayed Graduation: A Survey. GLO Discussion Paper Series 189, Global Labor Organization (GLO).
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Number 8769 in Economics Books. Princeton University Press.
- Beine, M., Delogu, M., & Ragot, L. (2020). The Role of Fees in Foreign Education: Evidence from Italy. *Journal of Economic Geography*, 20(2), 571–600.
- Belloc, F., Maruotti, A., & Petrella, L. (2010). University drop-out: An Italian experience. *Higher Education*, 60(2), 127–138.
- Borjas, G. (1987). Self-selection and the earnings of immigrants. *American Economic Review*, 77(4), 531–553.
- Bratti, M., Barbato, G., Biancardi, D., Conti, C., & Turri, M. (2021). Degree-programme determinants of university student performance. In D. Checchi, T. Jannelli & F. Uricchio (Eds.), *Teaching, research and academics Careers*. Springer.
- Bussu, A., Detotto, C., & Serra, L. (2019). Indicators to prevent university drop-out and delayed graduation: An Italian case. *Journal of Applied Research in Higher Education*, 12(2), 230–249.
- Card, D. (1993). Using Geographic Variation in College Proximity to Estimate the Return to Schooling. NBER Working Papers 4483, National Bureau of Economic Research, Inc.
- Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5), 1127–1160.
- Checchi, D. (2000). University education in Italy. *International Journal of Manpower*, 21(3–4), 177–205.
- Clarke, D., & Matta, B. (2018). Practical considerations for questionable IVs. *The Stata Journal*, 18(3), 663–691.
- Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly exogenous. *Review of Economics and Statistics*, 94(1), 260–272.
- Contini, D., & Zotti, R. (2021). Do financial conditions play a role in university dropout? New evidence from administrative data. In D. Checchi, T. Jannelli & F. Uricchio (Eds.), *Teaching, research and academics careers*. Springer.

- Garibaldi, P., Giavazzi, F., Ichino, A., & Rettore, E. (2012). College cost and time to complete a degree: Evidence from tuition discontinuities. *The Review of Economics and Statistics*, 94(3), 699–711.
- Johnes, G., & McNabb, R. (2004). Never give up on the good times: Student attrition in the UK. *Oxford Bulletin of Economics and Statistics*, 66(1), 23–47.
- Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics*, 107(1), 35–78.
- Lee, D. S., McCrary, J., Moreira, M. J., & Porter, J. (2020). Valid t-ratio inference for IV. Papers 2010.05058, arXiv.org.
- Li, C., Poskitt, D. S., & Zhao, X. (2019). The bivariate probit model, maximum likelihood estimation, pseudo true parameters and partial identification. *Journal of Econometrics*, 209(1), 94–113.
- Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Econometric Society Monographs. Cambridge University Press.
- Modena, F., Tanzi, G. M., & Rettore, E. (2018). The effect of grants on university drop-out rates: Evidence on the Italian case. Temi di discussione (Economic working papers) 1193, Bank of Italy, Economic Research and International Relations Area.
- OECD. (2019). *Education at a glance 2015: OECD indicators*. Paris: OECD Publishing.
- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2), 187–204.
- Rosenzweig, M. R., Irwin, D. A., & Williamson, J. G. (2006). Global wage differences and international student flows [with comments and discussion]. *Brookings Trade Forum*, 57–96.
- Weber, S., & Péclat, M. (2017). A simple command to calculate travel distance and travel time. *The Stata Journal*, 17(4), 962–971.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Number 0262232588 in MIT Press Books. The MIT Press.
- Zotti, R. (2015). Should I stay or should I go? Dropping out from university: An empirical analysis of students' performances. Working Papers 70, AlmaLaurea Inter-University Consortium.

**Gianfranco Atzeni** (Ph.D. University of Sassari) is Associate Professor of Economics at the University of Sassari. His research interests are in applied econometrics, financial economics, environmental economics, education economics, and economics of innovation.

**Luca Deidda** (Ph.D. SOAS) is Professor of Economics at the University of Sassari. His research interests are in the areas of migration, education and the labor market, finance and macroeconomics, and economics of information.

**Marco Delogu** (Ph.D. Université Catholique de Louvain and University of Luxembourg) is Assistant Professor of Economics at the Università di Sassari. His research interests are in the effects of migration, education economics, and sports economics.

**Dimitri Paolini** (Ph.D. Université Catholique de Louvain) is Professor of Economics at the University of Sassari. His research interests are in education and labor economics, cultural economics, and industrial organization.



**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.



**Part III**  
**Recruiting and Academic Careers**

# Is Entering Italian Academia Getting Harder?



Daniele Checchi and Tindaro Cicero

**Abstract** While a PhD degree is often considered the first necessary step to an academic career, since 2010 only a small fraction (less than 10%) of doctoral graduates obtained a position in academia within six years of the award of their degree. While we do not have information on their labour market outcomes, we can examine the determinants of this transition in order to study whether entry to an academic job is becoming more difficult. We merge three national administrative data archives covering completed doctoral degrees, postdoc collaborations and new hirings to academia (mostly assistant professor level). We find a decline in appointment probability after 2010, due to the hiring freeze imposed by fiscal austerity. We find, also, that a PhD degree and postdoc experience have a positive effect on the probability of obtaining a position in academia, while being a woman or being a foreign-born candidate has a negative effect. We found no evidence of career disadvantages for candidates from Southern universities.

**Keywords** Academic career · PhD graduates · Survival analysis

## 1 Introduction

Traditionally, in Italy, academics are respected. In a famous book, entitled *Baroni e burocrati. Il ceto accademico italiano* (Barons and bureaucrats: The Italian

---

Paper presented at the first Scuola Democratica conference (Cagliari, June 2019) and a University of Turin seminar (October 2021). We thank Gabriele Ballarino, Maria DePaola, Tullio Jappelli and Raffaella Rumiati for helpful comments. All remaining errors are our responsibility.

---

D. Checchi (✉)

University of Milan and INPS, Milano, MI, Italy

e-mail: [daniele.checchi@gmail.com](mailto:daniele.checchi@gmail.com)

T. Cicero

ANVUR, Rome, Italy

e-mail: [tindaro.cicero@anvur.it](mailto:tindaro.cicero@anvur.it)

© The Author(s) 2022

D. Checchi et al. (eds.), *Teaching, Research and Academic Careers*,

[https://doi.org/10.1007/978-3-031-07438-7\\_5](https://doi.org/10.1007/978-3-031-07438-7_5)

academic class) Giglioli (1982) compares academics to *mandarins* whose power over society was legitimized by selective access. In the occupational rankings proposed originally by DeLillo and Schizzerotto (1985), academics (*docenti universitari*) scored 80.91 out of 100 and in a recent update scored similarly highly (81.13) (Meraviglia & Accornero, 2007), equivalent to a manager (*dirigente*) in the public administration. However, almost contemporaneous scandals related to public competition for university professorships raised doubts among the public over the fairness of the selection process for entry to an academic profession (Perotti, 2008).

International comparisons (Janger et al., 2013), accounting salary levels, quality of life, doctoral degree, career prospects, research organization, balance between teaching and research, funding and probability of working with high-quality peers, suggest that Italy is one of the least attractive countries at entry to an academic career.<sup>1</sup> It is difficult to compare data on salaries at entry: a survey conducted on behalf of the European Commission (2007) reports that the average remuneration of an Italian researcher is 34.120 euro in PPP, against the EU-25 average of 40.126 euro and an equivalent US entry salary of 62.793 euro.<sup>2</sup>

Such a large wage differential, combined with the increasing career uncertainty revealed by our empirical analysis, might explain the brain drain among Italian researchers. Nascia et al. (2021) show that Italian researchers working abroad achieve faster career progression than those researchers who remain in the Italian system and they provide evidence of low confidence among Italian researchers of career advancement in Italy. The authors document how the decline in the number of university positions (20% over the period 2009-16) has translated into delayed career advancement and an increase in the average age of university staff. The main driver of migration is the perception that promotion abroad is based more on merit than on seniority-based progressions. This increases the salary differential and works against a domestic career.<sup>3</sup> The initial transition from doctoral graduate

<sup>1</sup> “With the exceptions of salaries and the teaching load, Italy shows elements of job attractiveness below average, in particular, the quality of peers, funding and career perspectives as well as the quality of life” (Janger et al., 2013, p. 17).

<sup>2</sup> However, the corresponding figures for Italy at entry level seem underestimated: the European Commission (2007, Table 12) reports 12.648 euro for a researcher with 0–4 years of experience, against an EU-25 average of 20.374 euro. However, the European University Institute Academic Career Observatory in Florence (<https://www.eui.eu/en/academic-units/max-weber-programme-for-postdoctoral-studies/aco-academic-careers-observatory>) reports a monthly entry salary for assistant professors in 2007 of 1500 euro (gross of tax, not corrected for PPP) against a corresponding value of 3708 euros in the USA and 3810 euros in the UK.

<sup>3</sup> “A drastic divide emerges between researchers in Italy and abroad with regards to the mechanisms of hiring and in terms of remuneration and career prospects. Recruitment in the home institution is considered to be transparent and merit-based by 57% of researchers in Italy and 80% of those abroad. PhDs and younger researchers in Italy have the most critical view of recruitment mechanisms in place. Considering the criteria for career progression, the same gap emerges. Merit is considered as the operating criteria by 54% of researchers in Italy against 75% of those abroad. Tenured positions are considered to be assigned on the basis of merit by 43% of researchers in Italy and by 62% of those abroad . . . The examination of remuneration shows that the share of

to assistant professor (from R1 to R2 in the OECD ranking) takes 5.5 years in Italian universities and 4.3 in universities abroad.

This evidence is supported by the increasing proportion of doctoral graduates who migrate abroad after finishing their degree (Istat, 2018): in 2018, 15.9% of graduates who obtained their PhD degree in 2012 were living and working abroad, and the percentage for those who were awarded their degree in 2014 was 18.5. Among those doctoral graduates who chose to remain in Italy, only 10.2% were employed as academics six years after the award of their degree, whereas among those who moved abroad, 25.9% achieved a position in academia.<sup>4</sup>

If doctoral graduates decide not to follow an academic career, where do they end up?

Passaretta et al. (2019) use Istat survey data on two cohorts (PhDs obtained in 2004 and 2008)<sup>5</sup> to show that academic reforms<sup>6</sup> and the 2009 economic crisis coincided with decreasing employment in academia and increasing chances of having a fixed-term contract, being employed abroad and working in research-related occupations outside academia. In particular, they show that five years after graduation, the proportion of doctoral graduates with tenured positions in academia (i.e. excluding postdocs and temporary assistant professors—*ricercatore di tipo A*) declined from 36% in the 2004 cohort to 29% in the 2008 cohort.

Ballarino (2020) provides new evidence about the social origins and occupational outcomes of doctoral graduates in Italy. He takes master's (MA) degree holders as the reference and shows that, after controlling for an equivalent time from degree award, doctoral graduates do not achieve higher incomes and have no greater employment probability. Almalaurea (an agency that interviews graduates on behalf of universities) data for 1999–2009, show that the doctoral graduates' academic employment probability declines by 0.7% per year, but increases by 1.4% for employment in the private sector. This is consistent with the reduced employment opportunities in academia over that period and, especially, in the social science and humanities disciplines.

---

researchers reporting to be badly paid or paid just to make ends meet is 47% in Italy and 15% for Italians abroad". (Nascia et al., 2021, p. 6).

<sup>4</sup> All these comparisons are potentially biased by self-selection: the best Italian PhDs could migrate abroad where they achieve faster career progression, simply because they are more productive. This is confirmed by Coda and Geuna (2018), who provide evidence that internationally mobile doctoral graduates perform better and have stronger international networks than their domestic peers.

<sup>5</sup> Istat conducted two surveys: the first, in 2009, covering the 2004 and 2006 cohorts and the second, in 2014, covering the 2008 and 2010 cohorts (Istat, 2010, 2015, respectively—descriptive evidence in Decataldo et al., 2016).

<sup>6</sup> "In a nutshell, in the second half of the 2000s, a set of academic reforms (1) cut the funding for the recruitment of new researchers (assistant professors) and for the promotion of academics (Berlusconi reforms—the so-called turnover block [2008]) and (2) abolished open-ended contracts at the start of the academic career (assistant professor) in favour of fixed term contracts, mostly without a tenure track, and put constraints on the renewals of temporary contracts in academia (Gelmini reform [2010])" (Passaretta et al., 2019, p. 545).

The existing evidence supports the conclusion that *Italian PhDs are gradually discouraged from entering academia in Italy* and are diverted towards foreign universities and/or the domestic private sector. However, the comparative decline in employment of doctoral graduates in universities could be due, also, to use of adjunct professors (*professori a contratto*) to cover teaching demand promoted by the increase in the courses offered by Italian universities, especially those less well-funded universities in the South of Italy (De Angelis & Grüning, 2020).

The decline in academic occupation opportunities is attributable mainly to the *recruitment restrictions imposed on Italian universities*: Italian budgetary law restricted new hirings to 20% of past retirements in 2012–2013 and then raised new hirings to 50% in 2014–2015, 60% in 2016, 80% in 2017 and 100% in 2018 (Corte dei Conti, 2021, p. 128). Table 1 presents employment in Italian universities and shows relative stability among teaching staff (+2.2% over 4 years), a marked decline in assistant professor entry level (−8.8%) and an increase in temporary positions (+10.9% for adjunct professor, +3.7% for postdoc, +12.9% for research assistants).<sup>7</sup> If we examine geographical variations (available in the original source) the picture is starker: based on the ratio of postdocs to assistant professors to proxy for the increased precariousness of academic jobs, this ratio changes from 0.92 to 1.0 in Northern universities and 0.31 to 0.35 in Southern universities. This implies that there are more (temporary) opportunities created in the North compared to the South, likely due to better availability of research funds in the former.

Within this general framework, we study entry to academic jobs since 2010. We address two questions: (1) whether gender, discipline and location affect the decreasing entry opportunities and (2) whether obtaining a postdoc position is advantageous for entry probability.

Our work extends the analysis conducted by Coda and Geuna (2020) who analyse the academic progression of graduates awarded their PhD degree by an Italian university, over the period 1983–2006 (before the hiring freeze and the reform to assistant professor positions). Doctoral graduates who pursued an academic career were identified by matching names to research fields using the list of academics active in Italian universities in the period 1990–2015. The most relevant conclusion is that, in the first 20 years (up to 2003) almost one third (33%) of Italian PhD degree holders were employed in academia (as assistant, associate or full professors). This excludes those awarded their PhD degree by a foreign university (this information is not included in the available databases) and includes the effects of the legal requirement for a PhD degree in order to apply for an assistant professorship (law 210/1998). Coda and Geuna provide various disaggregations (by gender, research field and geographical mobility): the share of PhDs pursuing an academic career is highest among men working in the fields of economics and statistics (52.2%) and lowest among women in medicine (19.2%).<sup>8</sup> The average time required for

<sup>7</sup> The gradual proletarianization of the academic profession has been described by the sociological literature: see Moscati (2020) and also Marini et al. (2019).

<sup>8</sup> See Table 12 in Coda and Geuna (2020).

**Table 1** Personnel in Italian universities, by role—2016–2019

	Public universities					Public and private universities						
	2016	2017	2018	2019	2016	2017	2018	2019	2016	2017	2018	2019
Tenured professors (full or associate)	31,100	31,160	31,979	33,827	32,896	33,034	33,969	35,968	32,896	33,034	33,969	35,968
Assistant professors (open or fixed)	19,737	19,224	18,824	18,055	21,339	20,759	20,193	19,458	21,339	20,759	20,193	19,458
Teaching staff	50,837	50,384	50,803	51,882	54,235	53,793	54,162	55,426	54,235	53,793	54,162	55,426
Adjunct professors	18,906	19,719	19,903	20,310	25,770	26,869	27,834	28,576	25,770	26,869	27,834	28,576
Postdocs ( <i>assegno di ricerca</i> )	13,484	13,550	13,480	13,751	13,946	14,124	14,105	14,459	13,946	14,124	14,105	14,459
RAs ( <i>collaboratori attività di ricerca</i> )	47,911	46,448	50,615	54,174	51,605	50,073	54,422	58,290	51,605	50,073	54,422	58,290
Linguistic collaborators ( <i>collaboratori linguistici</i> )	1544	1511	1476	1419	1713	1676	1587	1580	1713	1676	1587	1580
Temporary positions	81,845	81,228	85,474	89,654	93,034	92,742	97,948	102,905	93,034	92,742	97,948	102,905
Non-teaching staff ( <i>pers. tecnico-amministrativo</i> )	50,968	50,264	49,497	48,843	55,709	55,206	54,603	54,140	55,709	55,206	54,603	54,140
Total	183,650	181,876	185,774	190,379	202,978	201,741	206,853	212,471	202,978	201,741	206,853	212,471

Source: tavole 29–30 in Corte dei Conti, 2021

transition from PhD degree award to assistant professor is 4.5 years, with a declining trend between 1986 and 1996 and 1997 and 2006.

Our paper extends their analysis in several ways. First, we consider a more recent period: we use data on PhD degrees awarded during 2006–2017 and academic posts attained between 2010 (or earlier) and 2019. The collaboration between ANVUR and MIUR provided access to the list of doctoral graduates, which avoided having to parse them from the Italian National Library in Florence, where PhDs were required to materially deposit a copy of their dissertation.<sup>9</sup> We were able to exactly match the databases using social security identifiers, which minimized the risks caused by homonyms. The data also allow us to consider both direct transitions (from degree award to professorship, which became less usual) and indirect transitions (from PhD degree to postdocs, and postdocs to assistant professorships, distinguishing between permanent and temporary positions, but not between tenure track and purely temporary positions). We were able, also, to take account of age and citizenship, and a more precise definition of field of study.

## 2 The Data

The data for the analysis come from three administrative archives, which are not inter-connected, although they are managed by the same agency (CINECA) on behalf of the Ministry of University and Research (MIUR). Each database contains basic information on the individuals included, that is, gender, age, country of birth, research field and university of study/work. Our objective is to study the academic career paths of Italian doctoral graduates as the potential outcome of transitions within the national system: from PhD degree to professorship, possibly including some postdoc experience. The available data do not allow us to include those awarded their PhD degree from a foreign university or those in academic positions abroad; thus, we cannot assess what constitutes a “typical” academic career in Italy.<sup>10</sup>

---

<sup>9</sup> Our dataset contains more holders than Coda and Geuna (2020). Looking at their Table 1, for the period 2003–2006, they, respectively, count 6680, 8287, 9344 and 6795 PhDs. In our database for the same years, we count 10,665, 11,093, 11,291 and 11,395, suggesting that the parsing from the national library may be defective (for example, PhD schools—*istituti a statuto speciale*, like the Scuola Normale in Pisa—are excluded) or that a portion of PhDs did not comply with the legal obligation.

<sup>10</sup> The MIUR data do not contain information on PhD degrees awarded by foreign universities. If we observe a candidate in an assistant professor position who does not have a PhD degree awarded by an Italian university after 2010, we can assume that the individual was awarded the PhD degree from a foreign university. We still ignore the number of potential applicants and the proportion of candidates with two PhD degrees. Anecdotal evidence suggests that many doctoral students in Italian universities use their study abroad period to enrol in a foreign PhD programme (and obtained a second PhD degree after completing their studies in Italy).



However, the budget cuts imposed in 2009 and the abolition in 2010 of open-ended contracts for assistant professors (law n.240/2010) significantly modified internal career paths. The open-ended contracts (*ricercatore universitario a tempo indeterminato*) were replaced by three-year fixed contracts (with the possibility of one renewal for an additional 2 years—*ricercatore universitario a tempo determinato di tipo A*) and three-year tenure-track contracts (*ricercatore universitario a tempo determinato di tipo B*), which could be converted into open-ended contracts for associate professors upon attainment of the national qualification (*abilitazione scientifica nazionale*). Both types of vacancies were allocated according to open competitions at the local level.<sup>11</sup> The selection procedure was also modified: before the reform, candidates were required to pass two written exams to prove their knowledge of the discipline; after the reform, these exams were abolished and candidate were assessed only on their CVs (although many departments required shortlisted candidates to give a seminar, resembling job market paper interviews in the US system). The sequential nature of this reform, which was aimed at accelerating career progress for the most brilliant candidates, ultimately increased the queue for entry to academia.<sup>12</sup> Our analysis highlights the consequences of these policy changes.

## 2.1 The PhD Database

The first archive contains information on PhD students and graduates. At the moment of latest data retrieval (6/5/2019), the archive contained 175,423 individuals. After some data cleaning to exclude inaccurate repeat records, interrupted careers, dual entries for individuals with more than one doctoral degree,<sup>13</sup> we were

---

<sup>11</sup> The second type was open only to candidates who had already obtained the first type of contract or who had held a postdoc position for at least 3 years. To try to limit fixed-term employment, the law introduced a cap of five years on the cumulative duration of postdocs and fixed-term contracts for assistant professorships. Note that, for the first time, completion of a PhD degree became a prerequisite for application to assistant professor.

<sup>12</sup> This is openly recognized in official accounts as systemic failure: “*La ratio della riforma attuata dalla legge n. 240/2010 si basava sull’idea che la sostituzione delle figure a tempo indeterminato con quelle a tempo determinato avrebbe dovuto aumentare competitività e selezione basata sul merito, portando i ricercatori più meritevoli a transitare in poco tempo nel ruolo degli associati (tenure track). Tuttavia, il percorso per approdare a professore associato è costellato da una serie di posizioni a tempo determinato, partendo dall’assegno di ricerca (che deve essere preceduto da tre anni di dottorato), per una durata massima pari a quattro anni, cui segue un concorso per ricercatore a tempo determinato di tipo A (la cui durata massima è di cinque anni), per poi giungere al posto di ricercatore di tipo B, della durata di un triennio e suscettibile di conversione in professore associato, nel caso in cui sia stata conseguita l’Abilitazione Scientifica Nazionale*” (Corte dei Conti, 2021, p. 135).

<sup>13</sup> 2490 individuals have two PhD degrees and 75 have more than two PhD degrees. To preserve sample size, we consider the oldest degree.

**Table 2** Records in the PhD archive

Cycle	Median year of enrolment	Median year of defence	Interrupted	Still student	PhD obtained	Total
17	2003	2007	0	0	5	5
18	2003	2007	38	2	643	683
19	2003	2007	1173	114	<b>10, 489</b>	11, 776
20	2004	2008	1281	143	<b>11, 050</b>	12, 474
21	2005	2009	1498	271	<b>11, 313</b>	13, 082
22	2006	2010	1506	248	<b>11, 391</b>	13, 145
23	2007	2011	1431	313	<b>11, 385</b>	13, 129
24	2008	2012	1289	357	<b>11, 439</b>	13, 085
25	2009	2013	1363	447	<b>10, 438</b>	12, 248
26	2010	2014	1185	841	<b>9754</b>	11, 780
27	2011	2015	1238	789	<b>9380</b>	11, 407
28	2012	2016	933	3344	<b>7165</b>	11, 442
29	2013	2017	819	6401	<b>3997</b>	11, 217
30	2014	2018	393	8829	23	9245
31	2015		231	8631	0	8862
32	2016		104	9086	0	9190
33	2017		0	9782	0	9782
<b>Total</b>			<b>14, 482</b>	<b>49, 598</b>	<b>108, 472</b>	<b>172, 552</b>

left with 172,552 records. Table 2 groups PhD students by cohort of entry, which corresponds to national admission waves (*ciclo di dottorato*). It can be seen that more than 10% of *PhD students did not complete their study course or did not defend their thesis*. Since we are interested in the potential advantage from obtaining an Italian PhD for the probability of an academic career in an Italian university, we restrict our analysis to the 107,801 individuals (bold figures in Table 2) who were enrolled in PhD study programmes between the 19th and the 29th cycles, corresponding to completion years between 2006 and 2018.<sup>14</sup>

Table 3 shows the geographical distribution and research area of the PhD degrees. The share in the South declined by approximately 12 percentage points over a decade. In the case of discipline, the meanwhile STEM (science, technology, engineering and mathematics) and LIFE (biology, medicine, veterinary science) *have expanded by 10 percentage points in both the North and the South*, at the expenses of SSH (social science and humanities). Based on information on the labour market transitions of BA and MA graduates, STEM and LIFE PhDs confer a significant private sector employment advantages, which might account for the decline in

<sup>14</sup> Note that the number of PhD candidates has increased steadily since the year 2000 while the number of degrees awarded by Italian universities was less than 2000 up to 1992, increased to 4000 per year between 1992 and 2002, and increased to 12,000 a year in 2008. See Fig. 1 in Ballarino (2020).

**Table 3** PhD degrees by university location, gender and research area composition (%)

Cycle	North and Centre	South	Men	Women	STEM	LIFE	SSH
19	0.714	0.286	0.481	0.519	0.288	0.304	0.408
20	0.718	0.282	0.480	0.520	0.294	0.315	0.391
21	0.718	0.282	0.476	0.524	0.283	0.327	0.390
22	0.728	0.272	0.470	0.530	0.290	0.333	0.377
23	0.728	0.272	0.463	0.537	0.281	0.338	0.380
24	0.729	0.271	0.463	0.537	0.284	0.331	0.385
25	0.751	0.249	0.482	0.518	0.306	0.320	0.374
26	0.761	0.239	0.474	0.526	0.309	0.326	0.365
27	0.782	0.218	0.486	0.514	0.335	0.321	0.344
28	0.789	0.211	0.496	0.504	0.356	0.325	0.319
29	0.840	0.160	0.481	0.519	0.355	0.329	0.317
<b>Total</b>	<b>0.743</b>	<b>0.257</b>	<b>0.476</b>	<b>0.524</b>	<b>0.302</b>	<b>0.324</b>	<b>0.374</b>

Note: STEM includes CUN area 1, 2, 3, 4, 8b and 9; LIFE includes CUN area 5, 6 and 7; SSH includes CUN area 8a, 10, 11, 12, 13 and 14

transition to an academic career. Eventually, the gender partition fluctuates with a slight majority to the female component.

## 2.2 The Postdoc Database

The second archive includes postdoc positions (*assegni di ricerca*), which were introduced by the 1998 budget law.<sup>15</sup> The database is organized by events with 186,948 postdoc positions created between 1 September 1998 as starting date and expiration dates reaching 1 May 2022. These positions involved 80,659 scholars, which suggests that more than half obtained more than one position. In order to study the relative contribution of postdoc experience to the probability of an academic career, we retain only those positions that were still active or became active after 1 January 2006, when we start observing completed PhD degrees. Table 4 indicates that most postdoc positions (85%) were in Northern and Central universities, where their repeated use was also more frequent. The market seems segmented since the fraction of individuals who enjoyed a postdoc position in both macro-partitions is small.

<sup>15</sup> See item 6 art.51 in law no.447/1997, which sets a maximum of 8 years (reduced to 4 for PhDs students who benefit from a scholarship). Art.22 of law no. 240/2010 revised the maximum length to 4 years, making postdoc scholarships exempt from tax but liable for social contributions to pension schemes.

**Table 4** Postdocs by university location (positions and recipients)

#postdocs held by each individual	Full sample		Postdoc positions active after 1/1/2006				Postdoc individuals active after 1/1/2006				Total postdoc individuals
	Positions	Individuals	North and Centre	South and islands	Total postdoc positions	Men	Women	North and Centre	Entire country	South and islands	
1	36,690	36,690	26,051	5133	31,184	16,242	14,942	26,051	0	5133	31,184
2	37,970	18,985	26,955	5586	32,541	8598	8257	13,856	199	2800	16,855
3	29,328	9776	22,331	4296	26,627	4669	4656	7733	176	1416	9325
4	25,096	6274	19,146	3268	22,414	2874	3126	5060	136	804	6000
5	17,675	3535	14,454	2025	16,479	1719	1784	3048	63	392	3503
6	13,530	2255	11,401	1447	12,848	1088	1159	1972	49	226	2247
7	7861	1123	6524	857	7381	566	556	979	23	120	1122
8	6976	872	5789	561	6350	385	487	794	10	68	872
9	4185	465	3550	376	3926	230	235	418	6	41	465
10	3070	307	2538	368	2906	160	147	265	7	35	307
11	1980	180	1742	84	1826	86	94	169	4	7	180
12	1128	94	988	58	1046	40	54	89	0	5	94
13	559	43	470	27	497	20	23	40	1	2	43
14	378	27	308	0	308	9	18	27	0	0	27
15	300	20	231	42	273	12	8	17	1	2	20
16	80	5	74	0	74	4	1	5	0	0	5
17	85	5	79	0	79	1	4	5	0	0	5
18	18	1	18	0	18	1	0	1	0	0	1
19	19	1	19	0	19	1	0	1	0	0	1
20	20	1	19	0	19	0	1	1	0	0	1
<b>Total</b>	<b>186,948</b>	<b>80,659</b>	<b>142,687</b>	<b>24,128</b>	<b>166,815</b>	<b>36,705</b>	<b>35,552</b>	<b>60,531</b>	<b>675</b>	<b>11,051</b>	<b>72,257</b>

**Table 5** Postdocs duration by research area (positions)—active after 1/1/2006

	STEM	LIFE	SSH	total
0–6 months	7273	4053	2121	13,447
6–12 months	51,851	38,729	28,864	119,444
12–18 months	2868	2662	1904	7434
18–24 months	7429	6928	6335	20,692
24–30 months	437	695	415	1547
30–36 months	1073	814	729	2616
36–42 months	68	95	52	215
42–48 months	96	74	166	336
Beyond 48 months	14	20	28	62
Total	71,109	54,070	40,614	165,793

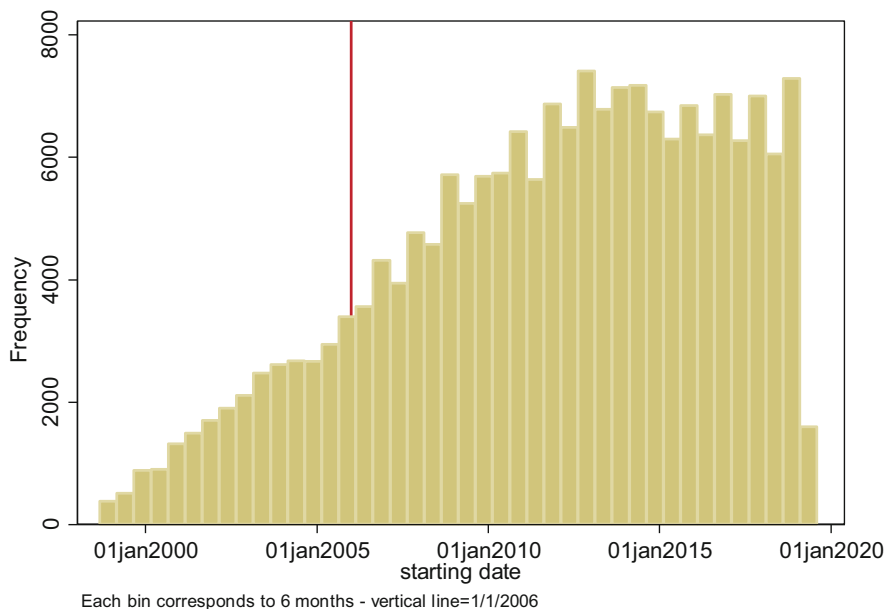
The duration of most of these postdocs is one year or less (72% in STEM, 71% in LIFE and SSH). Less than 3% of postdoc positions are for more than two years (see Table 5).<sup>16</sup>

Figure 1 shows that the offer of postdoc positions is a relatively recent phenomenon in Italian academia: they became more frequent after open-ended contracts for assistant professors were abolished in 2010. First, the increase in the assistant professor turnover rate reduced the number of teaching and/or research assistant jobs previously filled by tenured assistant professors, thus creating a demand for collaborators in research and teaching. Second, exempting postdoc scholarships from tax created an incentive to create postdocs rather than other forms of contractual arrangements. Figure 1 shows that, overall, Italy has offered an average of 14,000 new postdoc openings every year since the reform (in line with the aggregate figures in Table 1).

### 2.3 The Academics Database

The third archive includes administrative data on professors employed in Italian universities between 2010 and 2020. Table 6 shows a significant decline of around 8 percentage points, driven mostly by assistant professors. If we consider open-ended and fixed-term assistant professor contracts, we observe an overall decline of 22% (with an internal reallocation towards the temporary component, currently at 38%),

<sup>16</sup> While we have no information on the type of collaboration, recall that there are two types of postdoc position: *assegno di tipo A*, typically lasting 2 years, renewable for an additional 2 years, assigned based on open competitions, CVs and individual research projects; and *assegno di tipo B*, short-term collaborations for specific projects, often lasting only 6 months, based on discretionary hiring of principal investigators of larger research projects. Table 5 shows that this second type was frequent in STEM schools and does not necessarily reflect any academic aspirations, but rather temporary job opportunities for new graduates.



**Fig. 1** Time profile of new postdoc openings

followed by a similar (though smaller) decline of 10% in full professors. Over the same period, we observe a large wave of promotions from open-ended assistant professor to associate professor based on funds aimed at reducing the number of assistant professor positions (described as *ruolo ad esaurimento*, depletable position).

If we consider the final year (see Table 7), we see a new hiring pattern related to the newly created assistant professor position associated with a fixed-term contract: two-thirds of these individuals were awarded their PhD degree by and/or worked as a postdoc in an Italian university.<sup>17</sup> Unfortunately we ignore the date of entries to academia before 2010. However, to partially account for this, the right-hand panel in Table 7 includes only academics aged less than 40 years, which corresponds to 10% of the relevant population. For almost all positions, the fraction of doctoral graduates whose degree was awarded by an Italian institution jumps to 81%, confirming that among the most recent cohorts a PhD degree is required to obtain a position in Italian academia. Should we obtain data on PhD degrees awarded by foreign universities, this share would likely be closer to 100%. This is not surprising since law 240/2010 made a PhD degree a necessary requirement for the position of assistant professor with a fixed-term contract, however, it became enforceable only after 2016.<sup>18</sup>

<sup>17</sup> There is a caveat: we do not have information on Italian PhD degrees obtained before 2006, so it is likely that the shares indicated for full and associate professors constitute lower bound estimates of the actual shares.

**Table 6** Professors in Italian universities

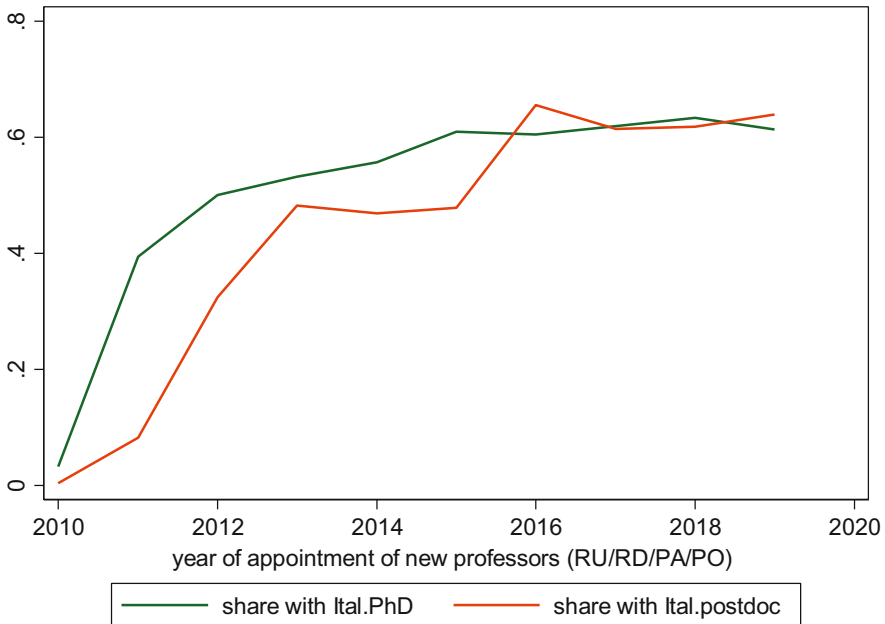
Stock at 31/12	Full professors	Associate professors	Assistant professor on open-ended contract	Assistant professor on fixed-term contract	Others	Total	% Full professors	% Associate professors	% Assistant professor on open-ended contract	% Assistant professor on fixed-term contract
2010	15,852	16,958	24,934	1139	242	59,125	0.27	0.29	0.42	0.02
2011	15,243	16,628	24,589	1577	175	58,212	0.26	0.29	0.42	0.03
2012	14,521	16,146	24,262	2378	146	57,453	0.25	0.28	0.42	0.04
2013	13,888	15,814	23,740	3019	171	56,632	0.25	0.28	0.42	0.05
2014	13,264	17,548	21,030	3560	231	55,633	0.24	0.32	0.38	0.06
2015	12,877	20,050	17,444	4105	326	54,802	0.23	0.37	0.32	0.07
2016	12,897	19,848	15,895	5285	299	54,224	0.24	0.37	0.29	0.10
2017 <sup>a</sup>	13,106	20,042	15,633	5699	285	54,765	0.24	0.37	0.29	0.10
2018 <sup>a</sup>	13,066	20,335	14,178	6415	289	54,283	0.24	0.37	0.26	0.12
2019 <sup>a</sup>	13,235	20,881	12,498	7743	410	54,767	0.24	0.38	0.23	0.14

<sup>a</sup> Recorded at 31/3

**Table 7** Professors in Italian universities by credentials—2019

	Full sample					Aged less than 40 years old (birth year >1978)						
	Number	Share with Italian PhD obtained since 2006	Share with postdoc since 2006 in Italian universities	Share with Italian PhD and postdoc position since 2006	Number	Share with Italian PhD obtained since 2006	Share with postdoc since 2006 in Italian universities	Share with Italian PhD and postdoc position since 2006	Number	Share with Italian PhD obtained since 2006	Share with postdoc since 2006 in Italian universities	Share with Italian PhD and postdoc position since 2006
Full professors	13, 235	0.01	0.00	0.00	39	0.61	0.00	0.00				0.00
Associate professors	20, 881	0.08	0.00	0.00	789	0.84	0.00	0.00				0.00
Assistant professor on open-ended contract	12, 498	0.09	0.00	0.00	363	0.82	0.00	0.00				0.00
Assistant professor on fixed-term contract	7743	0.66	0.66	0.51	4272	0.81	0.74	0.66				0.66
Others	410	0.09	0.05	0.03	26	0.65	0.35	0.27				0.27
Total	54, 767	0.15	0.09	0.07	5489	0.81	0.58	0.51				0.51





**Fig. 2** Credentials of newly hired assistant professors

For this reason we focus on new entrants (i.e. hired as assistant/associate/full professors) over the period 2010–19.<sup>19</sup> Thus, the temporal variations in the relevant share of Italian PhDs among newly hired professors (see Fig. 2) reflect both the limits of our administrative data and variations in the enforceability of the law. If we consider the most recent years as stable, we can argue that, currently, *newly hired professors have a PhD degree, two-thirds from an Italian university and (quite likely) one-third from a foreign university*. Also, in two-thirds of cases they have proof of research activity as postdocs. We also investigated whether there were

<sup>18</sup> Item 2b art.24 of law 240/2010 sets out the admission requirements for applying for a fixed-term assistant professor position: “*b) ammissione alle procedure dei possessori del titolo di dottore di ricerca o titolo equivalente, ovvero, per i settori interessati, del diploma di specializzazione medica, nonché di eventuali ulteriori requisiti definiti nel regolamento di ateneo, con esclusione dei soggetti già assunti a tempo indeterminato come professori universitari di prima o di seconda fascia o come ricercatori, ancorchè cessati dal servizio*”. The same law (para 13 of art.29) allows 5 years of derogation of this requirement: “*13. Fino all’anno 2015 la laurea magistrale o equivalente, unitamente ad un curriculum scientifico professionale idoneo allo svolgimento di attività di ricerca, è titolo valido per la partecipazione alle procedure pubbliche di selezione relative ai contratti di cui all’articolo 24*”.

<sup>19</sup> Since our academics data start in 2010, we can reconstruct new entries based on differences between 2011 and 2010 (and so on). For the earlier years (say, an assistant professor hired in 2009) we proxy new entries by restricting them to teaching staff existing in 2010 younger than 41, who were most likely hired in the previous decade.

variations in this dynamics by gender, but found no significant differences (see Table 8). However, we identified a clear declining trend in university hirings in the South, and a gradual substitutions of social science and humanities positions by LIFE sciences.

### 3 The Transition to Academic Careers

By merging the three datasets, we obtain a population of 144,446 individuals who completed their PhD degrees in an Italian university between 2006 and 2018 and/or held a (concluded) postdoc position in an Italian university between 2006 and 2019. This population is observed entering Italian academia during the period 2010–2019, and 10,104 had been appointed professor by the end of the sample period (see Table 9).

We first examine the reduced academic opportunities for the most recent cohorts. Given the structure of the data, if individuals enter the sample in different years, but are observed in the same year, older candidates have more time to obtain an academic post. To enable comparability, Table 10 presents the data in a moving window, recording eventual appointments in the six years following award of the PhD degree and/or completion of a postdoc period (choosing the earlier date in the case of both conditions being present).

If we consider all candidates, the probability of appointment declines by 2 percent points over eight years, but this hides a compositional effect: at the start of the period, all candidates awarded a PhD degree were headed towards an academic position and a postdoc position was a threat to promotion. At the end of the period, individuals with a PhD degree and postdoc experience were five times more likely to achieve an academic position compared to individuals with only a PhD degree or a postdoc experience. The disadvantage for the candidate is clear: if we observe the mean waiting period between degree completion and academic appointment (within the 6-year window), we see that it is around two years for only PhD degree and around four years for PhD + Postdoc candidates. The changing composition of the pool of newly appointed professors seems to keep an almost constant age at first appointment: the increasing age for PhD-only candidates is mostly counterbalanced by the declining age for the PhD + Postdoc group. Thus, over the sample period the age of first appointment across a six-year window, increases by one year only.<sup>20</sup>

The worsening conditions for the most recent cohorts who faced the hiring freeze and the reform of the assistant professorship after 2010, are confirmed by applying survival analysis for the risk of being appointed professor. The Kaplan–Meier failure functions (i.e. the share of promoted by year of entry in the sample) reported in Fig. 3

---

<sup>20</sup> Conditions worsen over time, such that candidates completing their degrees and/or postdoc experience in 2006 could be observed until 2019: if we take the average age among all appointees by year of appointment, we observe 34.7 years of age in 2010 increasing to 38.9 years in 2019.

**Table 8** Newly hired professors (assistant/associate/full) in Italian academia—2011–2019

Year of entry	Assistant professors (open-ended)	Assistant professors (fixed term)	Associate professors	Full professors	Total	Share with Italian PhD	Share with postdocs	Share with PhD+ Postdoc	% women	% South	% STEM	% LIFE	% SSH
2010 or earlier <sup>a</sup>	9072	585	1017	94	10,768	0.13	0.02	0.01	0.43	0.34	0.33	0.21	0.46
2011	1254	446	86	12	1798	0.39	0.08	0.04	0.42	0.30	0.30	0.25	0.44
2012	483	784	36	3	1306	0.50	0.32	0.20	0.43	0.30	0.30	0.28	0.43
2013	42	755	56	14	867	0.53	0.48	0.30	0.43	0.25	0.30	0.29	0.41
2014	12	750	119	15	896	0.56	0.47	0.33	0.38	0.22	0.33	0.30	0.37
2015	17	952	230	27	1226	0.61	0.48	0.37	0.40	0.28	0.31	0.31	0.38
2016	11	1464	154	39	1668	0.60	0.66	0.46	0.37	0.14	0.17	0.29	0.54
2017	1	624	82	28	735	0.62	0.61	0.50	0.39	0.11	0.33	0.33	0.34
2018	2	1218	195	37	1452	0.63	0.62	0.48	0.41	0.25	0.35	0.29	0.36
2019	1	2040	274	54	2369	0.61	0.64	0.48	0.41	0.23	0.30	0.33	0.37

<sup>a</sup> restricted to people aged 40 or less

**Table 9** PhDs/Postdocs working sample and their fraction hired as professors (assistant/associate/full) in Italian academia—2011–2019

	Absolute values				Hired as academic		
	no Postdoc	Postdoc	Total		no Postdoc	Postdoc	Total
No PhD		36,625	36,625	No PhD		4.8%	4.8%
With PhD	75,337	32,484	107,821	With PhD	5.3%	13.4%	7.7%
Total	75,337	69,109	144,446	Total	5.3%	8.9%	7.0%

suggest that if we abstract from the small initial cohort (made up of late completers, thus not fully representative of the quality of the candidates), the cohorts completing their degree in the period 2007–2009 benefited from a higher chance of obtaining a position in Italian academia, compared to later cohorts. The convexity of the lines associated to these cohorts is consistent with the effect of reopening vacancies in recent years.

The compositional effects presented in Table 10 can be represented by a different disaggregation of Kaplan–Meier analysis. Figure 4 depicts the failure function by type of credentials: candidates with just one credential (either a PhD degree or a postdoc experience) are at a disadvantage with respect to candidates with both types of credentials.<sup>21</sup>

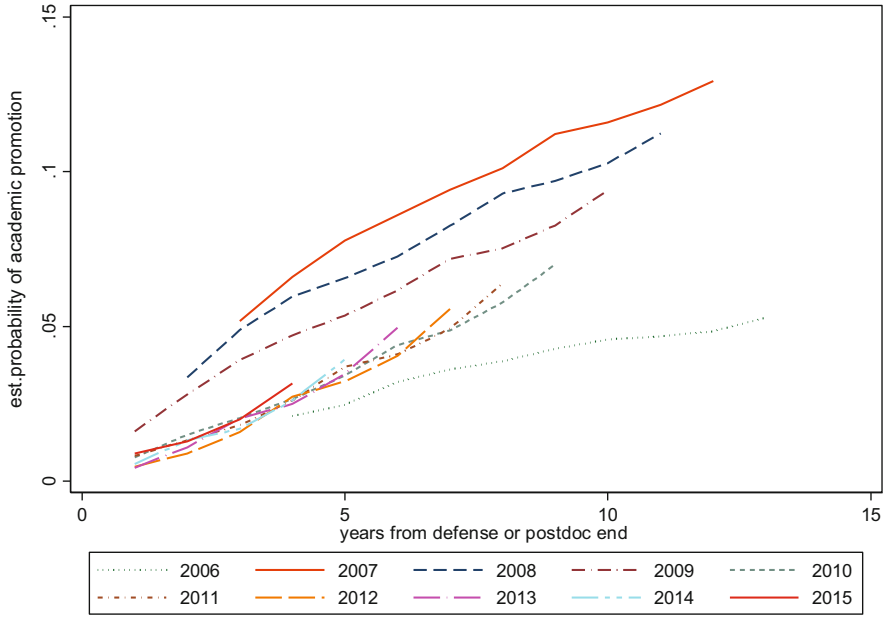
We are interested, also, in potential heterogeneity in academic prospects. We have shown that there are differences associated to period of completion and the type of credentials obtained by candidates, but need to examine the effects of gender, location and discipline differences. Figure 5 plots failure functions by gender and provides clear evidence of gender discrimination against women. The horizontal distance between the two lines describes the slower queue for females: at five years after completion, a man has a 5.6 point probability of an academic appointment, whereas it takes 7.5 years for women. At 13 years after graduation or postdoc completion, a women has the same chance of obtaining an academic position that a man has after 8 years.

This may be related, in part, to disciplinary differences, since academic career progression is faster in the social sciences and STEM disciplines and significantly shorter than in LIFE science disciplines (graph not shown). At 5 years after completion, a social science scholar has a 5.2 points chance of promotion; for STEM scholars reaching the same probability requires 5.3 years and for LIFE sciences scholars it takes 7 years. Since women are underrepresented in STEM and

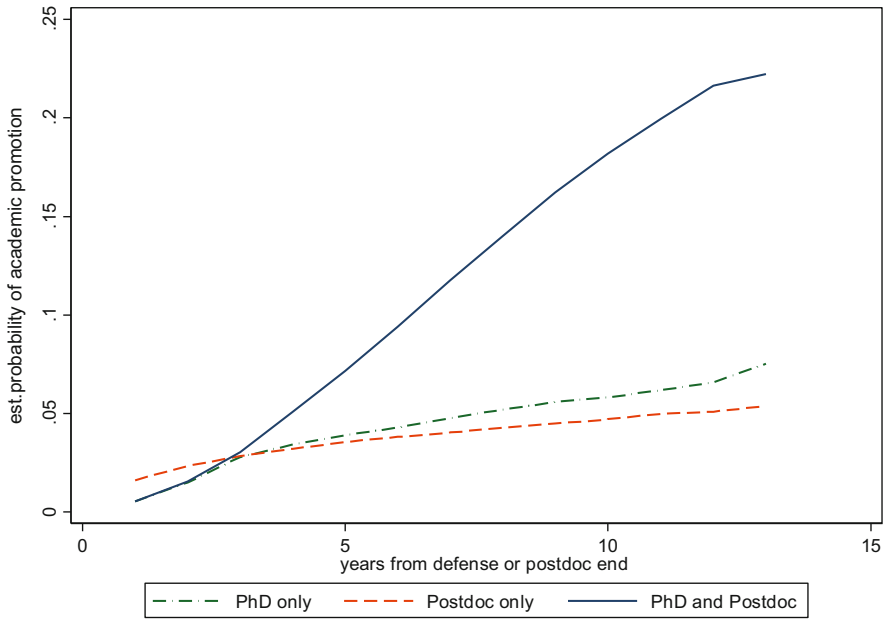
<sup>21</sup> Candidates with just postdoc experience could conceal scholars who were awarded their PhD degree by a foreign university and are trying to (re)enter the domestic academic market. However, the flat line in Fig. 4 is not in line with this interpretation, suggesting that candidates with a foreign PhD are also more likely to get postdoc experience abroad and to use this experience to apply for a higher academic position (such as associate professors) in Italy. Table 8 shows that half of new appointees have neither a PhD degree from nor postdoc experience in an Italian university. This can be taken as indirect evidence of the brain drain that has afflicted the Italian highly skilled scholars market.

**Table 10** PhDs/Postdocs appointments in a 6-year window from completion

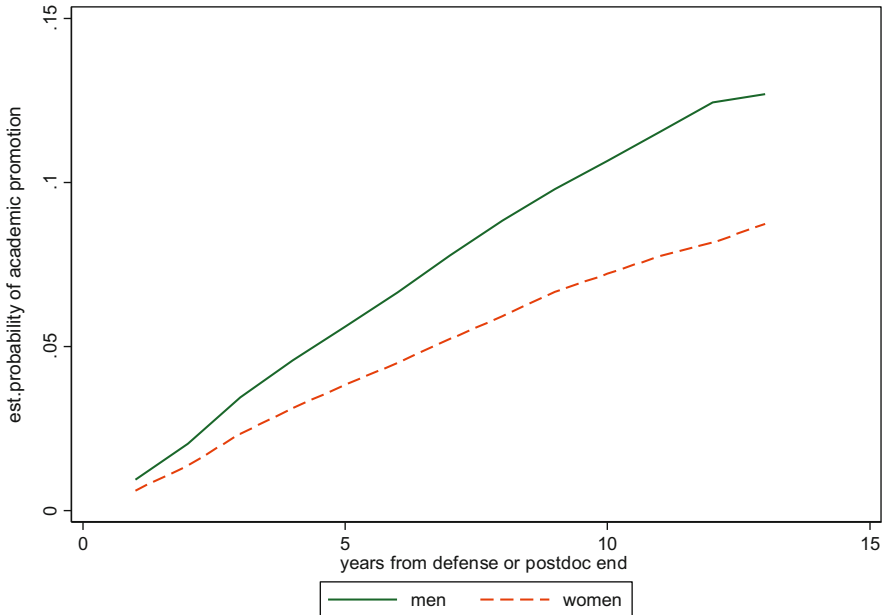
Year of conclusion (defence or postdoc termination)	PhD and/or Postdoc concluded in the year	Appointed professor in 6 years window	% appointed professor in 6 years window	Mean waiting (years) for appointment	Mean age at appointment
<i>All candidates</i>					
2006	2990	96	3.21%	4.57	35.30
2007	10,643	915	8.60%	3.72	35.05
2008	13,351	970	7.27%	3.14	34.70
2009	13,290	820	6.17%	3.02	35.03
2010	13,677	723	5.29%	3.00	35.21
2011	13,067	646	4.94%	2.79	35.52
2012	13,202	689	5.22%	2.82	35.78
2013	12,396	750	6.05%	3.17	36.22
<i>Candidates with PhD only</i>					
2006	486	74	15.23%	4.41	34.68
2007	5672	775	13.66%	3.52	34.90
2008	7607	731	9.61%	2.76	34.36
2009	7496	506	6.75%	2.45	34.82
2010	7734	317	4.10%	2.00	34.36
2011	7831	235	3.00%	1.48	35.14
2012	7922	185	2.34%	1.76	35.90
2013	7142	167	2.34%	2.18	36.32
<i>Candidates with PhD and postdoc experience</i>					
2006	442	10	2.26%	5.70	36.10
2007	2744	120	4.37%	4.94	35.46
2008	3434	207	6.03%	4.41	35.21
2009	3575	259	7.24%	4.17	34.68
2010	3537	309	8.74%	4.26	34.84
2011	3370	328	9.73%	3.97	34.93
2012	3450	373	10.81%	3.94	34.47
2013	3237	429	13.25%	4.19	34.95



**Fig. 3** Transition to academic professorship by cohort of PhD/postdoc completion



**Fig. 4** Transition to academic professorship by PhD/postdoc position



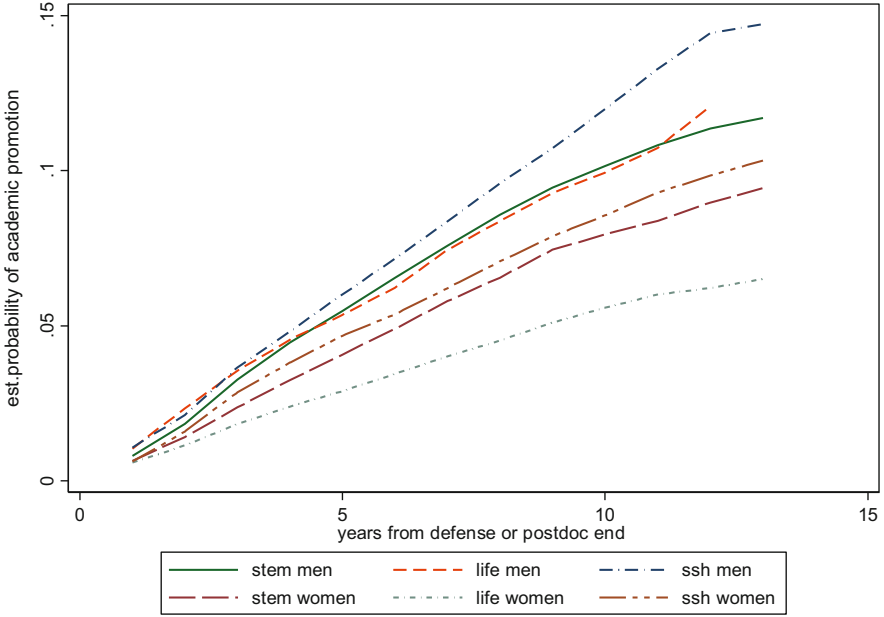
**Fig. 5** Transition to academic professorship by gender of the candidate

overrepresented in the other two disciplinary groups,<sup>22</sup> we can conclude that the disciplinary divide partly explains gender differences as shown in Fig. 6 which depicts both gender and discipline. Figure 6 shows that women in LIFE sciences experience slower career progression, attributable mostly to the long medical study period where the need for both academic and hospital experience imposes a double penalty.

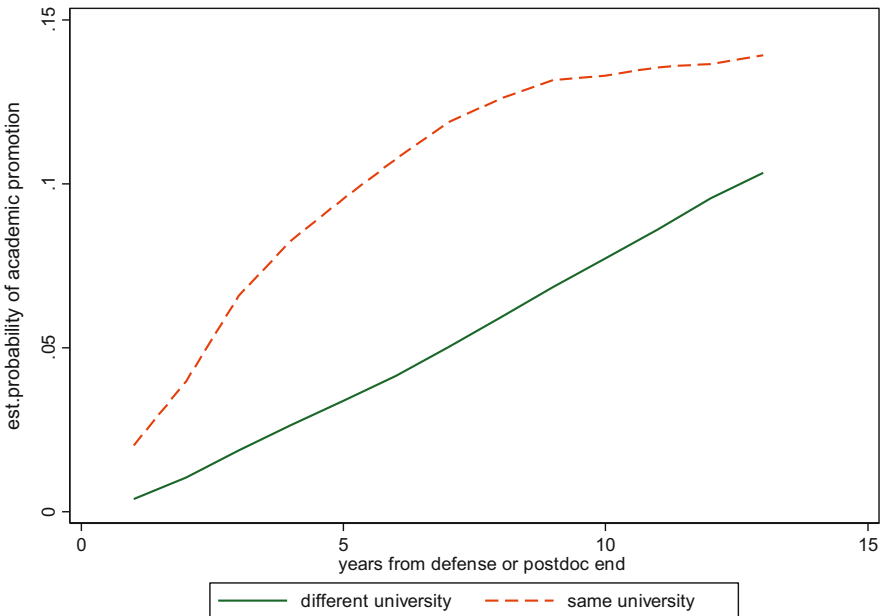
We explored another compositional effect related to the problem of “inbreeding” in Italian academia. In Italy, but not in other countries, universities are allowed to recruit their own PhD graduates: as a consequence candidates are more likely to be co-opted within the faculty if they are PhDs graduates from the same departments (as shown by Fig. 7, where local candidates almost double external ones in the appointment probability between year 3 and year 10 from degree defence).

A final and quite surprising result is the finding that there are no academic career differences between the North-Centre and the South of Italy (see Fig. 8). In Fig. 8, geographical location is the location of the university awarding the PhD degree and/or the location of the postdoc collaboration.<sup>23</sup> We observe a

<sup>22</sup> In the present working sample (based on disciplinary allocation of PhDs’ programs and/or postdoc disciplinary requirements), where the gender distribution is 48.8% of men and 51.2% of women, the male distribution is 46.1% in STEM, 23.8% in LIFE and 30.6% in SSH, while the corresponding figures for females are 22.8%, 39.9% and 37.3%.

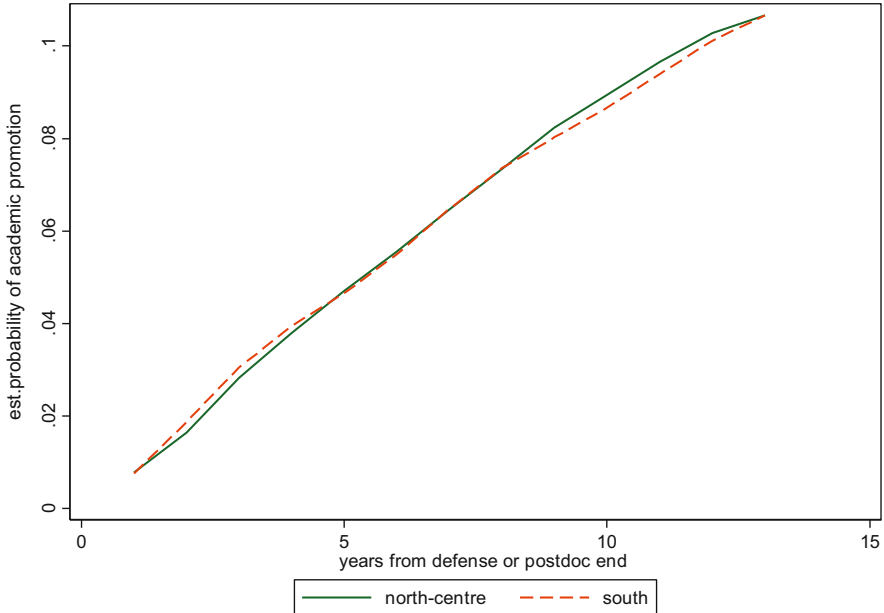


**Fig. 6** Transition to academic professorship by gender and disciplinary group of the PhD programme or the disciplinary requirement of the postdoc position



**Fig. 7** Transition to academic professorship by university of origin





**Fig. 8** Transition to academic professorship by location of the university awarding the degree or offering the postdoc position

fraction of “movers”, but only among the most successful graduates: 14.9% of applicants educated in Southern universities were appointed to positions in Northern universities and 8.6% of PhDs and/or postdocs from Northern universities were hired by Southern universities. However, given the non-observability of intended moves, we cannot assess whether moving to a different macro-region increases the probability of appointment and/or speeds up the promotion process. The similar career patterns in both macro-regions and the observed decline in the number of PhD degrees awarded by Southern universities (Table 3) combined with the reduced number of vacancies posted by Southern universities (Table 8) and the reduced mobility noted above, suggest that the Italian academic market is significantly segmented and spill-overs across universities tend to be minimal (consistent with the dynamics in Fig. 7).

We employ a Cox proportional hazard modelling to career dynamics and heterogeneous effects to estimate the relative risk of being appointed professor associated to different covariates; the reference case is a male aged 24 with a social sciences PhD awarded by a Southern university in the initial year (2007). Table 11 presents the estimates for two complementary models in a multivariate context, to

<sup>23</sup> Some candidates occupied postdoc positions in both regions: in these cases, location is attributed according to the prevailing length of the postdoc collaboration.

**Table 11** Transition from PhD/Postdoc to academic appointment

	Cox (coefficients)	Logit	Logit marginal effects
Female	-0.328 [0.021]***	-0.404 [0.022]***	-0.021 [0.001]***
Foreign born	-0.983 [0.061]***	-1.057 [0.058]***	-0.037 [0.001]***
Birth year	0.111 [0.027]***	0.102 [0.030]***	-0.000 [0.000]***
Research field = stem	0.352 [0.027]***	0.362 [0.029]***	0.005 [0.002]***
Research field = ssh	0.003 [0.024]	0.039 [0.026]	0.018 [0.001]***
Southern university	-0.07 [0.004]***	-0.011 [0.002]***	0.002 [0.001]
Postdoc only	-3.755 [0.059]***	-3.26 [0.090]***	-0.093 [0.002]***
PhD + postdoc	0.767 [0.023]***	1.09 [0.024]***	0.072 [0.002]***
Same university	3.383 [0.033]***	3.382 [0.069]***	0.421 [0.012]***
2007 cohort	0.596 [0.076]***	0.903 [0.085]***	0.064 [0.008]***
2008 cohort	0.579 [0.077]***	0.797 [0.085]***	0.054 [0.007]***
2009 cohort	0.542 [0.078]***	0.645 [0.086]***	0.041 [0.007]***
2010 cohort	0.417 [0.079]***	0.516 [0.086]***	0.031 [0.006]***
2011 cohort	0.499 [0.081]***	0.458 [0.088]***	0.027 [0.006]***
2012 cohort	0.586 [0.082]***	0.429 [0.089]***	0.025 [0.006]**
2013 cohort	0.763 [0.084]***	0.356 [0.090]***	0.020 [0.006]
2014 cohort	0.874 [0.088]***	0.183 [0.092]**	0.009 [0.005]*
2015 cohort	1.012 [0.093]***	0.098 [0.094]	0.005 [0.005]
2016 cohort	1.249 [0.100]***	0.306 [0.097]***	0.017 [0.006]***
2017 cohort	2.265 [0.107]***	0.183 [0.105]*	0.009 [0.006]*
Observations	133,481	134,565	134,565
Log Lik	-93,754	-30,091	-30,091
Pseudo R-squared		0.14	0.14

Robust z statistics in brackets - \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

The reference (excluded) case is a man, native, with a PhD only in the LIFE area obtained in 2006 in a university located in the North or Centre of the country, who (possibly) obtained a position in a different university from PhD one

study the correlation of each variable keeping into account the effects of the others. The first model is a Cox proportional hazard model of the risk of entering academia and, to enable comparison, we report the coefficients rather than the hazards. The second model is a standard probit, where the positive outcome is being appointed to an academic position. Column 3 reports the marginal effects of the probit model, evaluated at the sample mean, to capture the size of the effect.

Both models show that women, young individuals and candidates born abroad are at disadvantaged in terms of entering an academic career. Working in a STEM or social science field increases the likelihood of an academic career compared to working in LIFE sciences. University geographical location in the country shows no differences.<sup>24</sup> Being an incumbent candidate increases the probability of appointment by 42.1 percentage points vis a vis an external candidate.

In terms of qualifications and experience, candidates with a PhD degree and a postdoc experience have a 7.2 percentage points higher probability of an academic hiring vis a vis candidates with a PhD degree. Postdoc experience without an Italian PhD is associated to a negative premium: this reflects the predominance of short-term collaborations, but no long-term academic aspiration (*assegno di ricerca di tipo B*). When we consider the cohort dummies, we find that the cohorts between 2010 and 2014 are disadvantaged by the hiring freeze. The time profiles based on the two models differ: the Cox model shows that the cohorts that completed their degrees/experience after 2010 retained their positive advantage with respect to the initial cohort, with the best candidates being hired soonest. The probit model suggests that these cohorts are indistinguishable in the probability of academic appointment with respect to the excluded case.

## 4 Concluding Remarks

This chapter provides new evidence regarding the initial steps towards an academic career in Italy. Contrary to the literature, which suggests that one in three individuals who complete their PhD degree programmes in an Italian university is appointed to the position of professor, we find that this probability has more than halved. Among those awarded their PhD degrees or completed postdoc experience in 2007, only 12.9% were in an academic position in 2019. This is due mostly to the temporary hiring freeze which affected the whole of Italy's public administration in the period 2010–14. However, there is also evidence of increased competition from abroad, since half of newly hired candidates do not have an Italian PhD degree (and, therefore, were likely awarded one by a foreign university since a PhD degree is a precondition for applying for an assistant professor vacancy).

---

<sup>24</sup> This is confirmed if we split the sample into discipline subgroups and re-estimate the model based on subsamples. The estimated marginal impact of being a woman is  $-0.015$  for STEM and social sciences and  $-0.025$  for LIFE sciences.

We found also that rather than a transition from PhD completion to academic position, the sequence has gradually become doctoral graduate, postdoc experience and then an academic position. This has increased the time from study completion to appointment. If we consider that the average duration of a postdoc position is 2.5 years in our dataset and it gives access at best to a two-year temporary assistant professor contract, we can expect the following sequence. Based on the averages for the individuals in our sample, PhD defence occurs at 32.7 years of age, followed by a transition period of 4 years (including a 2.5 year postdoc position) and for one in ten candidates, at 36.7 years of age, a temporary three-year contract (*ricercatore a tempo determinato di tipo A*). This equates reaching almost the age of 40 years of age without a permanent position. For the best researchers who have obtained the appropriate qualification (*abilitazione scientifica nazionale*), application for a tenure track position is possible and, after three years, at around age 43, an associate professor position. Since in 2010, the average age of an assistant professor on an open end contract was 34 years (44 years for associate professor) we can see how much more difficult it has become to obtain a permanent position in academia in Italy.

We found heterogeneity in transition probabilities. Women, young individuals and foreign-born candidates are negatively discriminated, regardless of research field. Internal candidates receive an undue advantage. This should concern policy-makers, since there are no reasons for these selection biases. In the case of gender, it could be argued that the length of transition impinges on and causes conflicts between childbearing and academic aspirations.<sup>25</sup> Unfortunately, we have no information on the scientific productivity of these candidates since most have few entries in the Web of Science and Scopus databases. If we could identify a proxy for their academic impact, we could investigate whether this negative discrimination is based on productivity or is purely statistical discrimination.

We found an absence of geographical differences in the transition probability, which contradicts claims that Southern universities suffered more than Northern universities from the cuts imposed during the period of financial austerity. However, we found that their share of PhD degrees and vacancies/promotions declined in the period analysed, suggesting the presence of a consistent brain drain. Finally, we found that research field matters: careers in social sciences and STEM progress more rapidly than those in LIFE science.

Our study has two main limitations, which suggest caution when interpreting our results. The first limitation is that we do not observe the alternative careers of candidates who emigrate, have experience at a foreign university, and then return to Italy to take up an academic position. Anecdotal evidence suggests that their share is increasing since Italian doctoral graduates are found in postdoc positions in many departments of foreign universities. It would be useful to gather data on this (temporary?) brain drain. The only current source of administrative information on this aspect is AIRE (the registry of Italians residing abroad), but this would

---

<sup>25</sup> See chapter “Academic Careers and Fertility Decisions” in this volume.

underestimate their number. Data could be obtained by parsing the curricula of the applicants to the competition for assistant professorships, but this would be rather time consuming since the information is not organized in a consistent way.

The second limitation is related to outside options for the doctoral graduates. Nine out of ten abandon any academic aspirations and apply for positions in the public administration or the private sector. We do not have information on the monetary return associated with a PhD degree, but it is likely that the quality of the job obtained compensates for the lack of monetary incentives. Otherwise, it would be puzzling that more than 10,000 Italian graduates embark on a PhD degree with no expectation of future outcomes.

Evidence of a brain drain to foreign universities and companies can be seen as confirming the quality of the training offered by Italian universities. Being a net exporter of PhD graduates is the joint outcome of excess supply (recall the doubling of PhD positions over the last decades) and lack of demand (the number of posts in academia has reduced over the same time period), combined with a low price (represented by the options outside of academia). In the absence of indirect proxies for candidate abilities (such as scientific productivity), we are unable to assess whether self-sorting of candidates deprives the country of the brightest individuals.

## References

- Ballarino, G. (2020). *Come cambia il dottorato di ricerca. Organizzazione istituzionale e sbocchi occupazionali in comparazione internazionale*. Rapporto di ricerca per UNIRES.
- Coda Zabetta, M., & Geuna, A. (2018). *International postdoctoral mobility and career path: Evidence from Italian academia*. Mimeo.
- Coda Zabetta, M., & Geuna, A. (2020). *Italian Doctorate holders and academic career: progression in the period 1986–2015*. Carlo Alberto Notebooks n. 629.
- Corte dei Conti. (2021). *Referto sul Sistema Universitario*, Rome.
- De Angelis, G., & Grüning, B. (2020). Gender inequality in precarious academic work: Female adjunct professors in Italy. *Frontiers in Sociology*, 4, 87. <https://doi.org/10.3389/fsoc.2019.00087>
- Decataldo, A., Giancola, O., & Trivellato, P. (2016). *L'occupazione dei PhD in ambito accademico e nel mercato esterno. La mobilità geografica come vincolo ed opportunità*. Comunicazione convegno AIS-Verona, mimeo.
- DeLillo, A., & Schizzerotto, A. (1985). *La valutazione sociale delle occupazioni*. il Mulino.
- European Commission. (2007). *Remuneration of researchers in the public and private sectors*. Carsa.
- Giglioli, P. P. (1982). *Baroni e burocrati. Il ceto accademico italiano*. il Mulino.
- Istat. (2010). *L'inserimento professionale dei dottori di ricerca. Anno 2009–2010*, Statistiche in breve. 14/12/2010.
- Istat. (2015). *L'inserimento professionale dei dottori di ricerca - Anno 2014*, Statistiche Report, 21/1/2015.
- Istat. (2018). *L'inserimento professionale dei dottori di ricerca – Anno 2018*. Statistiche Report 26/11/2018.
- Janger, J., Strauss, A., & Campbell, D. F. J. (2013). *Academic careers: a cross-country perspective*. WWW for Europe Working Paper no 37.

- Marini, G., Locke, W., & Whitchurch, C. (2019, March). *The future higher education workforce in locally and globally engaged higher education institutions: A review of the literature on the topic of “the academic workforce”*. Centre for global higher education working paper n. 43.
- Meraviglia, C., & Accornero, L. (2007). La valutazione sociale delle occupazioni nell'Italia contemporanea: una nuova scala per vecchie ipotesi. *Quaderni di Sociologia*, 45, 19–73.
- Moscato, R. (2020). Academics facing unpredictable changes. *Scuola Democratica*, 3, 405–416.
- Nascia, L., Pianta, M., & Zacharewicz, T. (2021). Staying or leaving? Patterns and determinants of Italian researchers' migration. *Science and Public Policy*, 2021, 1–12.
- Passaretta, G., Trivellato, P., & Triventi, M. (2019). Between academia and labour market. The occupational outcomes of PhD graduates in a period of academic reforms and economic crisis. *Higher Education*, 77, 541–559.
- Perotti, R. (2008). *L'università truccata*. Einaudi.

**Daniele Checchi** (Ph.D. University of Siena) is Professor of Economics at the University of Milan and a former member of the Board of Directors of ANVUR, currently on leave at the Research Department of INPS (Italian Social Security Administration). He studies academic careers and research evaluation.

**Tindaro Cicero** (Ph.D. University of Rome Tor Vergata) is a researcher at ANVUR, where he supports the organization of the national qualifications for professorship.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.



# Academic Careers and Fertility Decisions



Maria De Paola, Roberto Nisticò, and Vincenzo Scoppa

**Abstract** We investigate how academic promotions affect the propensity of women to have a child. We use administrative data on the universe of female assistant professors employed in Italian universities from 2001 to 2018. We estimate a model with individual fixed effects and find that promotion to associate professor increases the probability of having a child by 0.6 percentage points, which translates into an increase by 12.5% of the mean. This result is robust to employing a Regression Discontinuity Design in which we exploit the eligibility requirements in terms of research productivity introduced since 2012 by the Italian National Scientific Qualification (NSQ) as an instrument for qualification (and therefore promotion) to associate professor. Our finding provides important policy implications in that reducing uncertainty on career prospects may lead to an increase in fertility.

**Keywords** Fertility · Promotion · Academic Career · Career uncertainty

**JEL Classification:** J13, J65, J41, M51, C31

---

M. De Paola (✉) · V. Scoppa

Department of Economics, Statistics and Finance “Giovanni Anania”, University of Calabria, Rende, CS, Italy

Institute for the Study of Labor (IZA), Bonn, Germany  
e-mail: [m.depaola@unical.it](mailto:m.depaola@unical.it); [v.scoppa@unical.it](mailto:v.scoppa@unical.it)

R. Nisticò  
Institute for the Study of Labor (IZA), Bonn, Germany

Department of Economics and Statistics, University of Naples “Federico II”, Napoli, NA, Italy  
Center for Studies in Economics and Finance (CSEF), Naples, Italy  
e-mail: [roberto.nistico@unina.it](mailto:roberto.nistico@unina.it)

© The Author(s) 2022

D. Checchi et al. (eds.), *Teaching, Research and Academic Careers*,  
[https://doi.org/10.1007/978-3-031-07438-7\\_6](https://doi.org/10.1007/978-3-031-07438-7_6)

## Used Acronyms

ANVUR	National Agency for the Evaluation of the University and Research Systems (“AGENZIA NAZIONALE DI VALUTAZIONE DEL SISTEMA UNIVERSITARIO E DELLA RICERCA”)
LPM	Linear Probability Model
MIUR	Italian Ministry of Education, Universities and Research (“Ministero dell’Istruzione, Università e Ricerca”)
NSQ	National Scientific Qualification
OECD	Organization for Economic Cooperation and Development
OLS	Ordinary Least Squares
RDD	Regression Discontinuity Design
RTD-A	Assistant Professor (Type A) (“Ricercatore a Tempo Determinato Tipo A”)
RTD-B	Assistant Professor (Type B) (“Ricercatore a Tempo Determinato Tipo B”)
RU	Assistant professor (permanent position) under the old regime (“Ricercatore Universitario”)

## 1 Introduction

Do career prospects affect fertility choices? Researchers have long been concerned with the economic factors driving the decision to have a child, typically looking at such decision as the result of a utility maximization process that takes into account costs and benefits of children, subject to income constraints and individual’s preferences (Becker, 1981). Women’s fertility decisions interact with those regarding employment as they are the solution of a common constrained maximization problem (Del Boca & Sauer, 2009; Francesconi, 2002; Cigno, 1991). On the one hand, better employment prospects, by increasing opportunity costs, reduce fertility. On the other hand, higher income may lead to an increase in fertility. The ambiguity of this relationship (depending on whether the income effect prevails over the substitution effect) is confirmed by the changing correlation between fertility and female labour market participation observed in recent years.

An important aspect that has been attracting greater attention, especially in explaining the persistently low fertility rates of many advanced countries, is the increased labour market insecurity (Sobotka et al., 2011). As individuals are typically risk averse, higher economic insecurity and more uncertain career prospects might push them to decrease the number of children in order to reduce risk. There is growing empirical evidence on how economic uncertainty affects fertility decisions. Prior studies have shown a negative impact of aggregate unemployment on fertility (Currie & Schwandt, 2014). Other studies have investigated the impact of unemployment at the individual level, providing evidence of a strong negative effect



that is mostly caused by the career shock rather than the income shock induced by unemployment (Del Bono et al., 2015). Some other works have looked at the fertility consequences of job instability focusing on temporary contracts (De La Rica & Iza, 2005) or on employment protection (De Paola et al. 2021; Prifti & Vuri, 2013; Bratti et al., 2005).

An alternative reason why the increased economic insecurity may affect fertility is that women might decide to postpone childbearing due to their desire to pursue a career: a higher economic instability might induce people, in particular the young, to defer family formation until they achieve full integration into the labour market. Unsurprisingly, the mean age of women at birth of first child has increased remarkably in most OECD countries, rising from an average of 24 in 1970 to 30 in 2017.<sup>1</sup> A number of recent papers find very relevant child penalties and women might consider these costs in their fertility decisions (see Bertrand, 2018 for a survey). While previous research has documented negative effects of fertility on a woman's career, little is known on the extent to which promotion affects fertility. The present research aims at filling this gap by addressing how career advancements within academic positions of women employed in the Italian University system affect fertility decisions. As explained above, the effect might be driven by different channels including income effects, reduced insecurity and desire for recognition on the workplace.

Academic career in Italy remains markedly characterized by strong vertical segregation: only 21% of full professors are women, while the proportion of women among associate and assistant professors is 36% and 47%, respectively. The low representation of women at the top of the hierarchical ladder can be due to many factors, such as differences in productivity, but they may also be related to the fact that promotion procedures favour men rather than women. For example, some previous works examining gender differences in the academic labour market show that women suffer a disadvantage in promotions and a within-rank pay gap (Blackaby et al., 2005; Ginther & Kahn, 2004; McDowell et al., 1999). Moreover, a number of papers looking at gender differences in career prospects in Italian academia provide evidence of a lower success probability of women compared to men in career advancement (Bagues et al., 2017; De Paola et al., 2017; Jappelli et al., 2017; De Paola & Scoppa, 2015). There is also evidence that the average number of years required for the transition from researcher to associate professor is greater for women (SIE gender commission, 2016).

Due to domestic responsibilities, which include among others child-rearing and household keeping, women might have less time to perform the research and teaching necessary for advancement. Many studies show, in fact, that women do much more household labour than men and that this extends to academics (Ward & Wolf-Wendel, 2004). These delays and difficulties might induce women who want to consolidate their professional position to postpone motherhood with negative consequences on their total fertility rate. This can also lead to involuntary

---

<sup>1</sup> In Italy, this figure has reached 32 years in 2018.

childlessness, also because of the health-related risks associated with delaying entry into motherhood (te Velde et al., 2012). The proportion of childlessness among women at the end of their reproductive period has increased dramatically in many OECD countries, especially in Italy where, the fraction of childlessness for those born in 1978 has doubled up (22.5%) with respect to that for women born in 1950 (11.1%).<sup>2</sup>

This chapter contributes to the existing research on economic uncertainty and fertility decisions by focusing on the impact of improvements in career prospects, which has been so far overlooked. More specifically, we analyse how the transition from the entering position in the Italian academia (“Researcher”) to the position of Associate professor changes the propensity to have children. We use administrative data gathered by the Italian Ministry of Education and from the National Agency for the Evaluation of the University and Research Systems (ANVUR) providing information on both fertility decisions and career advancement. Our investigation relies on two different estimation strategies. The first one considers the whole sample of women hired by an Italian University as researchers starting from 2001 to 2018. For these women we have yearly information both on Compulsory Maternity Leave (which we use as a proxy of fertility decision) and on their career advancements: exploiting the panel structure of our dataset we estimate an individual fixed-effect model that allows us to control for time-invariant individual characteristics to investigate the impact of promotion to the position of associate professor on the probability of having a child. The second estimation strategy exploits the eligibility requirements in terms of research productivity imposed by Italian National Scientific Qualification (NSQ) to advance in the academic ladder from assistant to associate professor. This institutional feature allows us to adopt a Fuzzy Regression Discontinuity Design and estimate the causal effect of career advancement on fertility by comparing the propensity to have a child for women who just got the Qualification with that of women who just missed it.

Our empirical analysis shows that women who experience career advancements have a higher probability of having a child. More specifically, we document that promotion to associate professor positions increases the likelihood of child birth of about 0.6 percentage points, which translates into an increase by 12.5% of the mean. This finding is robust to a battery of checks, including a specification that allows either age or years of experience to enter non-linearly in order to flexibly control for the fact that both promotion and maternity could be related to age or seniority, respectively. More importantly, the size of the impact is fairly stable across the two alternative estimation strategies used in the empirical analysis. Moreover, as promotions from not tenured to tenured positions are those expected to have the

---

<sup>2</sup> The number of women working in the Italian academia who do not have children is particularly high. According to the report produced by the Gender Commission of the Italian Economic Society (2016), based on a survey proposed in 2014 and 2016 to the members of the Italian Economic Society, about 33.9% of female economists aged above 50 have no children, while this percentage is only 13% for their male counterparts. A similar gap is found also for individuals aged 40–50, with this figure being 32% and 23.4% for women and men, respectively.

highest impact on fertility, the effect we estimate focusing on promotions between tenured positions can be considered a sort of lower bound of the impact deriving from increased job security.

The estimated positive effect of promotion can be explained either by the higher income associated to promotion or by the fact that women who have obtained a career recognition feel more comfortable devoting time and energy to childbearing without fear of negative career repercussions.

We also find that the estimated effect is highly heterogeneous by age and geographic area of the individual's university. In addition, we document that the effect is more salient immediately after promotion and gradually vanishes with the number of years from promotion. Furthermore, we find heterogeneous effects depending on whether promotion occurs before or after 2012, i.e. when the recruiting system changed due to the Gelmini Reform, with the effect being larger in the latter case (i.e. an increase by 20% of the mean). This differentiated effect can be due to the fact that the reform, by increasing the minimum standard required to obtain promotion, has made it more difficult for women who want to pursue a career to have children before career advancement.

Importantly, we find similar results when we apply a Fuzzy Regression Discontinuity Design and estimate the impact of obtaining the National Scientific Qualification on the probability of having a child. The size of the impact is consistent in magnitude with that obtained using the individual fixed-effects model, though the estimates become relatively imprecise.

The chapter is organized as follows. In Sect. 2, we discuss the institutional background. In Sect. 3, we describe the data we use in the analysis and show some descriptive statistics. In Sect. 4, we present the results of the effects of promotion on fertility using the individual fixed-effects model. In Sect. 5, we illustrate the estimates obtained from our alternative estimation strategy based on a Fuzzy Regression Discontinuity Design. Section 6 offers some concluding remarks.

## 2 Institutional Background

In this section, we provide some information on the institutional setting of the Italian Academia. The rules governing careers in Italian universities have changed over time. The sample of individuals we consider was interested by two different systems. According to the first system, before 2012, there were three academic positions: Assistant Professor or Researcher ("Ricercatore"), i.e., the entry level; Associate Professor ("Professore Associato"); Full Professor ("Professore Ordinario"). All three were permanent positions: formally there was a probationary period of 3 years, but tenure was very rarely denied.

Since all were permanent positions, the key differences between associate professors and researchers were the annual income (about 35% higher for associate professor position) and teaching duties, that were more intense for associate professors.

In the first system, a university willing to fill a vacancy initiated a competition, and a committee of five members was selected to choose a shortlist of candidates (the so-called *idonei*).<sup>3</sup>

Once the process was concluded, the university that initiated the competition could decide to appoint one of the winning candidates as professor, while the other could be appointed by another university within three years. This mechanism remained in place until 2011.

In 2012, a new system was introduced, following a major reform of the university system in 2010 (the so-called Gelmini Law). The reform was aimed at increasing transparency and meritocracy through a two-stage procedure: a first stage, in which candidates aiming for promotion to associate or full professor positions are required to qualify in a centralized national competition held at the field level, the so-called National Scientific Qualification (NSQ), in which candidates' publications and CVs are evaluated in relation to a field-specific minimum standard; and a second stage, in which effective promotions (or new hiring) are managed at the local level by each university.

Obtaining the NSQ is only the first step to get a promotion. In fact, university departments can autonomously choose full and associate professors to hire among individuals who have obtained the NSQ, through an open competition for both internal and external candidates or, alternatively, through a competition limited to internal candidates. Then, the probability of being effectively promoted for individuals who gained the NSQ depends on the number of vacancies opened by university departments, which in turn depends on resources obtained from the central government.

As a consequence of the "Gelmini Reform", since 2012, the entering positions have become temporary with two main types of contracts, "Ricercatore di tipo A" and "Ricercatore di tipo B" with different contractual length but similar teaching duties. The position of "Ricercatore di tipo A" (type-A Researcher) may last for up to 3 years and is temporary, with no career path. The position of "Ricercatore di tipo B" (type-B Researcher) lasts for three years and is a tenure-track position towards Associate Professorship, conditional on the researcher obtaining the National Scientific Qualification as Associate Professor.<sup>4</sup> Therefore, while the position of Researcher in the pre-reform system was permanent, the new positions of type-A and type-B Researcher are temporary positions.

In our analysis we do not consider individuals in these temporary positions because we do not have data on their fertility decisions, so we focus on individuals

---

<sup>3</sup> One member was appointed by the university Department opening the vacancy while four members out of five were before 2008 elected by all the professors in the field (but de facto nominated by the same Department) and—after 2008—randomly selected (among all the full professors in each field). The number of winning candidates ("idonei") changed along time (3 in some periods, 2 in some others, 1 for a period of 2 years).

<sup>4</sup> Even if the contractual length of "Ricercatore di tipo A" and "Ricercatore di tipo B" is not different (three years for both), the length of the former can be—in some cases—prolonged for two additional years.

who were hired by an Italian University with a permanent contract in the position of assistant professors starting since 2001. Since we observe these individuals for a period of about 20 years (from 2001 to 2018), we are able to detect any change in their position in the hierarchical ladder.

In this study, we consider the procedure for obtaining the NSQ launched in 2012. We focus on individuals who have already a permanent position as Assistant Professor who apply in the NSQ for the Qualification of Associate Professor. In case of failure to obtain the promotion, they remain at the same level of the academic ladder. Promotion is quite relevant in terms of salary: the yearly gross salary for assistant professors is about €41,000, while it rises to €54,000 for associate professors and about €72,000 for full professors. As regards other aspects, Italian academics have similar obligations and constraints at all the hierarchical levels and carry out similar tasks. However, prestigious positions such as rector, dean, head of department are open only to full professors.

Italian academia is organized into 14 different scientific areas (e.g. physics, medicine, economics and statistics); each area is in turn divided into different scientific fields (e.g. applied physics, econometrics, private law), for a total of 184 fields. The NSQ is awarded by a committee (specific to each field) of five members, randomly selected from the full professors in each field who have reached some scientific productivity standards and volunteered for the task. Committee members evaluate candidates for both associate and full professor positions and award the NSQ. There are no limits to the number of qualifications awarded in each field. Committees have full autonomy on the criteria to be used in the evaluation, but some criteria were suggested by the Italian Ministry of Education, Universities and Research (MIUR) in relation to the research productivity of candidates in the previous 10 years, as measured by some bibliometric indicators (see Sect. 4 for details).

### 3 Data

We use administrative data from the universe of women working in Italian universities since 2001 until 2018. The dataset is collected by ANVUR, the Italian National Agency for Evaluation of University and Research, and provides detailed information on the academic position covered by each woman in each year, her *Age*, the years since hiring (*Experience*), Compulsory Maternity Leave, the geographical area of the University in which the individual is employed. Data are structured as an individual-year panel data set.<sup>5</sup> Due to the features of the dataset that only provides information on maternity leaves, we will focus exclusively on women aged up to 46.

---

<sup>5</sup> The dataset was provided to us in anonymized form for our empirical analysis at the Laboratory of ANVUR headquarters.

We build the dependent variable  $ChildBirth_{it}$  a dummy equal to one for woman  $i$  who have a birth at year  $t$  (and zero otherwise). The data at hand also provide information on: age, years of experience (years since hiring in the University), academic position, academic fields (84 “macro-settori”) and university’s geographic areas (North-West, North-East, Centre, South, Islands).

The main explanatory variable is *Promotion to Associate Professor*, a dummy equal to one if an individual obtains a promotion in year  $t$  (or has obtained a promotion in the past  $k$  years).<sup>6</sup> In the second part of our analysis we will use as alternative explanatory variable *Qualification*, a dummy variable taking value one if the individual has been awarded the NSQ in the previous year or earlier (and zero otherwise). The NSQ introduces explicit thresholds in three productivity indicators (e.g. # of publications, # of citations, h-index, etc.) which vary across academic sectors, and scholars have to meet at least two out of three indicators in order to gain the eligibility for career advancements. We will exploit data for the NSQ in 2012 that provide for each candidate the score in each of the three productivity indicators and the outcome of the qualification procedure.

In Table 10 in the Appendix, we report some descriptive statistics for the universe of women who were hired by an Italian University as assistant professor in the period from 2001 to 2018 (excluding women who were already Associate and Full Professors in 2001)—whom our individual fixed-effects estimates are based on (descriptive statistics for the sample used in the RDD analysis are provided in Sect. 5). Women included in our sample are on average 40.17 years old (with a minimum of 24 years and a maximum of 46). The vast majority of them have an age ranging from 36 to 46. The probability of having a child is of 4%. About 15% of women who started their career as assistant professor have been promoted to associate professor in the period covered by our data.

## 4 The Effect of Promotion on Fertility: An Individual Fixed-Effects Approach

In this section, we investigate the impact of being promoted to the position of associate professor on the fertility decision of women working as researchers (i.e. assistant professors) in Italian Universities. In order to try to handle confounding factors deriving from unobserved heterogeneity, we exploit the panel structure of our dataset (with about 12,000 individuals observed on average for nearly 9 years) and estimate the following model including individual fixed effects:

$$ChildBirth_{it} = \beta_0 + \beta_1 Promotion_{it-k} + \beta_2 X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

---

<sup>6</sup> We experiment with different values of  $k$ .

where the dependent variable is a binary variable  $ChildBirth_{it}$  which takes the value of one if the researcher  $i$  in the year  $t$  had a child and zero otherwise. Among the independent variables, we consider the step variable  $Promotion_{it-k}$  which takes the value of one starting from year  $t$  in which the researcher has been promoted to a higher position and zero otherwise (for  $k$  years).  $X_{it}$  is a vector of the candidate's characteristics including age and years of experience.  $\mu_j$  and  $\lambda_t$  are individual and year fixed effects, respectively. In some specifications we also include scientific field- and geographic area-specific trends. In all the regressions, standard errors are robust to heteroskedasticity and allowed for clustering at the individual level. By estimating our model with individual fixed effects, we are able to take into account time-invariant heterogeneity in productivity across individuals, even if we are not able to control for variation of productivity occurring over time.

We estimate our model on the sample of women who were hired by an Italian University as assistant professor in the period from 2001 to 2018. We exclude women already in a position of Associate or Full Professor in 2001. We also restrict the sample to women under the age of 46, therefore ending up with 11,897 individuals and a total of 101,774 observations, one for each year since hiring.

Results from individual fixed-effects regressions are reported in Table 1. Reading across columns of Table 1, our estimates indicate that promotion to the position

**Table 1** The effect of promotion on child birth. LPM with individual fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Promotion to Associate Prof.	0.006** (0.003)	0.006** (0.003)	0.006** (0.003)	0.006** (0.003)	0.006** (0.003)	0.005* (0.003)
Age	-0.000 (0.000)	-0.002 (0.003)	-0.004 (0.003)	-0.004 (0.003)	-0.004 (0.003)	-0.001 (0.004)
Age Sq.		0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Experience (yrs.)			0.002** (0.001)	0.002* (0.001)	0.002* (0.001)	0.002* (0.001)
Individual FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	No	No	No	Yes	Yes	Yes
Geographic area-specific trends	No	No	No	No	Yes	Yes
Scientific field-specific trends	No	No	No	No	No	Yes
Observations	101,774	101,774	101,774	101,774	101,774	101,774
Clusters (individuals)	11,897	11,897	11,897	11,897	11,897	11,897

Notes: Estimates from OLS regressions are reported in each column. The dependent variable is *Child Birth*. We include individual fixed effects. Sample includes all female hired as assistant professors (RU) in Italian Universities after 2001 who are aged up to 46 followed until 2018. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR

of associate professor leads to an increase in the probability of having a child of about 0.6 percentage points, which is statistically significant at the 5% level. Results are quite stable across specifications. In column 1, we only control for *Age*, while in column 2 we also include *Age Squared*. The effect remains the same both in terms of magnitude and statistical significance also when we include *Years of Experience* (column 3), year dummies (column 4), geographical area-specific trends (column 5) and scientific field-specific trends (column 6). The size of the effect is slightly reduced to 0.5 percentage points when we include both scientific field- and geographic area-specific trends.

In the interest of comparing our main results with the OLS estimates, in Table 12 in the Appendix of the chapter, we report results from a Linear Probability Model in which we do not control for individual fixed effects. Results obtained when estimating a simple pooled OLS model might be biased due to the fact that women who are more likely to be promoted have peculiar features (for instance, they are characterized by a higher scientific productivity) which might also affect their probability of having a child. If these unobserved features positively affect the probability of both being promoted and having a child, we would expect an upward bias. In contrast, if these unobserved features are negatively correlated to fertility decisions, we will end up with a downward bias. A downward bias would emerge, for instance, if more productive women tend to postpone fertility or decide to have no children at all. This could well be the case if their higher productivity depends on the fact that, being free from duties related to childbearing, they devote more time to research. Estimates reported in Table 12 show that failing to account for individual unobserved heterogeneity leads to an insignificant effect of promotion on fertility when we control for years of experiences (column 3), year dummies (column 4), geographic area dummies (column 5) and scientific field dummies (column 6).

Next, we test the robustness of our main results, reported in Table 1, to several checks. To begin, column 1 in Table 2 shows the results obtained from a specification in which instead of controlling for age we include a saturated set of age dummies to flexibly control for the fact that both promotion and maternity could be related to age in a complex non-linear form. Notwithstanding the inclusion of age dummies, the impact of promotion on fertility is unchanged.

Second, we redo the same exercise and test whether results hold when replacing the variable *Experience* with the full set of dummies for the number of years of experience. This allows accounting for the fact that both promotion and maternity could be related to seniority in a complex way. Reassuringly, the estimates reported in column 2 of Table 2 are in line with our baseline results both in terms of magnitude and statistical significance.

Next, we carry out a robustness check excluding from our sample women too young or too old (respectively, bottom 1% in *Age* and top 1% in *Experience*). The estimates in columns 3 to 4 of Table 2 do not change qualitatively with respect to those shown in Table 1. Finally, in column 5 we exclude both at the same time and the results are unchanged.

In Table 3, we investigate whether the effect of promotion on fertility takes place immediately after promotion or in the subsequent years. In column 1, we restrict the



**Table 2** The effect of promotion on child birth. Robustness checks

	(1)	(2)	(3)	(4)	(5)
	Using age dummies	Using experience dummies	Excluding bottom 1% in age	Excluding top 1% in seniority	Excluding both at the same time
Promotion to Associate Prof.	0.005* (0.003)	0.005* (0.003)	0.005* (0.003)	0.005* (0.003)	0.005* (0.003)
Age		-0.000 (0.004)	-0.000 (0.004)	-0.001 (0.004)	-0.000 (0.004)
Age Sq.		0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Experience (yrs.)	0.002* (0.001)		0.002* (0.001)	0.002* (0.001)	0.002* (0.001)
Individual FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Geographic area- specific trends	Yes	Yes	Yes	Yes	Yes
Scientific field-specific trends	Yes	Yes	Yes	Yes	Yes
Observations	101,774	101,774	101,499	100,733	100,458

Notes: Estimates from OLS regressions are reported in each column. The dependent variable is *Child Birth*. We include individual fixed effects. Sample includes all female hired as assistant professors (RU) in Italian Universities after 2001 who are aged up to 46 followed until 2018. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR

**Table 3** The effect of promotion on child birth. Short versus long run effects

	(1)	(2)	(3)	(4)	(5)
	Up to 1 year from promotion	Up to 3 years from promotion	Up to 7 years from promotion	2 to 3 years after promotion	4 to 7 years after promotion
Promotion to Associate Prof.	0.010** (0.004)	0.006** (0.003)	0.005* (0.003)	0.003 (0.003)	0.003 (0.004)
Age	-0.001 (0.005)	-0.000 (0.004)	-0.001 (0.004)	-0.001 (0.004)	-0.002 (0.004)
Age Sq.	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Experience (yrs.)	0.002* (0.001)	0.002* (0.001)	0.002* (0.001)	0.002* (0.001)	0.002* (0.001)
Individual FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Geographic area-specific trends	Yes	Yes	Yes	Yes	Yes
Scientific field-specific trends	Yes	Yes	Yes	Yes	Yes
Observations	89,995	95,193	100,100	91,931	91,640

Notes: Estimates from OLS regressions are reported in each column. The dependent variable is *Child Birth*. We include individual fixed effects. Sample includes all female hired as assistant professors (RU) in Italian Universities after 2001 who are aged up to 46 followed until 2018. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR

**Table 4** The effect of promotion on child birth. Heterogeneous effects by age (above/below median)

	(1)	(2)
	Age < =40	Age > 40
Promotion to Associate Prof.	0.010* (0.006)	0.001 (0.004)
Experience (yrs.)	0.005** (0.002)	0.001 (0.001)
Individual FEs	Yes	Yes
Year FEs	Yes	Yes
Geographic area-specific trends	Yes	Yes
Scientific field-specific trends	Yes	Yes
Observations	49,184	52,590
Clusters (individuals)	8472	11,260
Mean of dependent variable	0.052	0.029

Notes: Estimates from OLS regressions are reported in each column. The dependent variable is *Child Birth*. We include individual fixed effects. Sample includes all female hired as assistant professors (RU) in Italian Universities after 2001 who are aged up to 46 followed until 2018. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR

sample in order to include only observations within a year after promotion. We find that the probability of having a child increases by 1 percentage point immediately after promotion. In column 2, we consider a period up to three years after promotion and the impact reduces to an increase of 0.6 percentage points, while it reaches a lower bound of 0.5 percentage points when considering a period up to seven years after promotion (column 3).

Finally, in columns 4 and 5, we test whether the impact of promotion on the probability of having a child extends to the two-three years following promotion (column 4) or to the period between four to seven years after promotion (column 5), respectively. As expected, the estimates in columns 4–5 are lower in magnitude than the ones in column 1, though they are not statistically significant at conventional levels. The result showing that the impact of promotion reduces over time might depend on the fact that, since women in our sample are on average 37 years old, the time left for childbearing is limited.

In Table 4, we estimate our model (the specification with the full set of controls) separately for women aged below and above 40 (i.e. the median age). This permits to compare women who are more similar in terms of age and then as regards the probability of having a child. We find that the effect of promotion on fertility is mainly driven by younger women. This is consistent with the hypothesis that younger women, by facing lower time pressure to have a child, have greater incentives to postpone childbearing in the interest of pursuing a professional career.

**Table 5** The effect of promotion on child birth. Heterogeneous effects by geographic area

	(1)	(2)	(3)
	North	Centre	South
Promotion to Associate Prof.	0.010** (0.004)	-0.002 (0.005)	0.001 (0.004)
Age	0.000 (0.006)	-0.003 (0.006)	-0.014** (0.005)
Age Sq.	-0.000 (0.000)	0.000 (0.000)	0.000* (0.000)
Experience (yrs.)	0.002 (0.002)	0.001 (0.001)	0.006** (0.003)
Individual FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	46,892	22,149	32,734
Clusters (individuals)	5335	2767	3795
Mean of dependent variable	0.055	0.025	0.029

Notes: Estimates from OLS regressions are reported in each column. The dependent variable is *Child Birth*. We include individual fixed effects. Sample includes all female hired as assistant professors (RU) in Italian Universities after 2001 who are aged up to 46 followed until 2018. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR

Thus, according to this prediction, the fertility response to promotion for younger women should be larger than that for older women.<sup>7</sup>

In Table 5, we investigate whether the effects are heterogeneous according to geographic areas in which university are located. We find that women affected by promotion in their fertility decisions are mainly those working in the North of the country. This might be due to the fact that the income constraint is more binding in a more developed area, where the support coming from grandparents is weaker (due to better employment conditions).

However, the higher supply of nurseries and kindergartens is likely to work in the opposite direction. It is well known that Southern regions are characterized by low availability of child care services: the percentage of places available in nurseries with respect to resident children (up to the age of 2) is approximately 30% in the North, 33% in the Centre and 13% in the South. Therefore, in Southern regions, even when a promotion (and a higher income) is obtained, couples can be discouraged by the lack of child care services, whereas the latter problem is less binding in Northern regions.

Furthermore, the North-South difference could also be explained by the fact that, due to different social norms, the pursuit of professional advancement is more relevant for women living in the northern part of the country (as documented

<sup>7</sup> Note that in our sample the average age at promotion is 43.

**Table 6** The effect of promotion on child birth. Heterogeneous effects by university regulation

	(1)	(2)
	Promotion under old regime (pre-2012)	Promotion under new regime (post-2012)
Promotion to Associate Professor	0.003 (0.003)	0.009** (0.005)
Age	-0.003 (0.005)	-0.001 (0.004)
Age Sq.	0.000 (0.000)	0.000 (0.000)
Experience (yrs.)	0.002 (0.001)	0.000 (0.001)
Individual FEs	Yes	Yes
Year FEs	Yes	Yes
Geographic area-specific trends	Yes	Yes
Scientific field-specific trends	Yes	Yes
Observations	63,740	81,000
Clusters (individuals)	10,629	9631
Mean of dependent variable	0.032	0.044

Notes: Estimates from OLS regressions are reported in each column. The dependent variable is *Child Birth*. We include individual fixed effects. Sample includes all female hired as assistant professors (RU) in Italian Universities after 2001 who are aged up to 46 followed until 2018. Standard errors clustered by individual are reported in parentheses. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . Source: Administrative data provided by ANVUR

by Istat, 2018, southern Italian regions are still characterized by strong gender stereotypes).

Table 6 reports the estimated effect of promotion on child birth depending on whether promotion took place in the period before 2012, i.e., during the old recruiting system, or in the period starting from 2012 when promotion is regulated by the Gelmini Reform. According to our data, the percentage of women who got promoted during the old regime (i.e. before 2012) is 11%, while this figure increases to 21% during the new regime (i.e. since 2012).

Results in Table 6 highlight that the effect of promotion is larger in the more recent period, i.e., under the new regime regulating career advancements in the Italian academia (i.e. an increase by 20% of the mean). This finding likely stems from the fact that the reform, through the National Qualification System, has increased the relevance of scientific productivity for promotion and therefore has made it more difficult for women who want to pursue a career to have children, as they would typically have less time to reach the minimum standards in terms of scientific productivity required for career advancement. On the other hand, once promoted, women increase their propensity to have children.

Finally, we verify if the impact of promotion on fertility is heterogeneous across macro-fields (Natural Sciences, Medicine, Engineering, Social Sciences and

Humanities) and we find a quite similar impact (results, not reported, are available upon request).

## 5 The Effect of Qualification on Fertility: A Fuzzy Regression Discontinuity Approach

In this section, we investigate the impact of improved career prospects on women fertility decisions using an alternative identification strategy which exploits the eligibility requirements in terms of research productivity imposed by the Italian National Scientific Qualification (NSQ) to advance in the academic ladder to positions for associate and full professor. As explained in Sect. 2, currently in order to get promoted to associate and full professors, candidates need first to obtain a National Scientific Qualification (NSQ), awarded by a national committee who consider candidates' publications and CVs in relation to a field-specific minimum standard.

Obtaining the NSQ is only the first step to get promotion. In fact, university departments can autonomously choose full and associate professors to hire among individuals who have obtained the NSQ, through an open competition for both internal and external candidates or, alternatively, through a competition limited to internal candidates. Then, the probability of being effectively promoted for individuals who gained the NSQ depends on the number of vacancies opened by university departments, which in turn depends on resources obtained from the central government.

To award the qualification, the committee members in each scientific field first consider three measures of candidates' scientific productivity (in the 10 years preceding the evaluation) in relation to some field-specific cutoffs (defined on the basis of the median values of these measures in the target position). In bibliometric (mainly scientific) fields,<sup>8</sup> the productivity indicators used are: (1) the number of articles published in scientific journals, (2) the total number of citations and (3) the h-index. In non-bibliometric fields (social sciences and humanities), the indicators are: (1) the number of articles published in scientific journals, (2) the number of articles published in high-quality journals and (3) the number of books.

The fact that Italian researchers have to meet at least two out of three productivity thresholds to qualify for associate and full professor allows us to employ a Fuzzy Regression Discontinuity Design and exploit the discontinuity in the likelihood of being awarded the qualification when two out of three indicators are equal or above the relative thresholds.

---

<sup>8</sup> Bibliometric fields include mathematics, physics, chemistry, earth sciences, biology, medicine, agricultural and veterinary sciences, civil engineering and architecture, industrial and information engineering, and psychology.

The Regression Discontinuity Design (RDD), compared to the individual fixed-effect model adopted before, has the advantage of exploiting variation that is arguably exogenous around a given threshold. In fact, in a RDD framework individuals are characterized by some variable  $X$  over which they do not have full control, and that has to reach a given threshold for them to receive a certain treatment. In this way, focusing on individuals near the threshold (above and below) allows to compare individuals that are very similar in terms of observable and unobservable characteristics, but some of them are “treated” and some others are not. This enhances the credibility of the estimation results, since there are less concerns that treated and control individuals differ for factors other than the treatment.

Then, we estimate the causal effect of *Qualification* (and so the prospect of career advancement) on fertility by comparing the likelihood to have a child for women who just achieve and just miss the qualification. In this way, any jump in fertility in proximity of the cutoff point of productivity indicators can be interpreted as evidence of a treatment effect.<sup>9</sup>

Following most of the papers in the literature, we use a parametric approach. Formally, we estimate the following first-stage equation:

$$\text{Qualification}_{it} = \alpha_0 + \alpha_1 \text{Above}_{it} + \sum_{m=1}^3 \delta_m f(\text{distance}_{itm}) + \alpha_2 X_{it} + \mu_j + \gamma_g + \lambda_t + \varepsilon_{it} \quad (2)$$

where  $\text{Above}_{it}$  is a dummy variable equal to one when at least 2 of the 3 indicators are above (or equal) the relative thresholds,  $f(\text{distance}_{itm})$  are three flexible functions of the distance of each  $m$  running variable (individual productivity indicator) from its respective cutoff,  $X_{it}$  is a vector of individual characteristics (e.g. age, seniority) and  $\mu_j$ ,  $\gamma_g$ ,  $\lambda_t$  are dummies for scientific fields, university’s geographic areas and year, respectively.  $\varepsilon_{it}$  is an error term. We will allow standard errors for clustering at the individual level.

Then, we use the discontinuity in the probability of achieving the qualification as an instrumental variable in the following second-stage equation:

$$\text{ChildBirth}_{it} = \beta_0 + \beta_1 \widehat{\text{Qualification}}_{it-k} + \sum_{m=1}^3 \delta_m f(\text{distance}_{itm}) + \beta_2 X_{it} + \mu_j + \gamma_g + \lambda_t + \varepsilon_{it} \quad (3)$$

where  $\beta_1$  is the local average treatment effect (LATE) of being awarded the NSQ on the subsequent propensity to have a child.

We estimate our model on the universe of female assistant professors who have applied for the Associate Professor Qualification at the NSQ in 2012. We apply the same restrictions discussed in Sect. 3 and focus exclusively on women aged up to 46, thus ending up with a sample of 3986 individuals and 19,407

<sup>9</sup> For a similar strategy exploiting the discontinuity in productivity indicators from NSQ to analyse a different outcome (productivity after promotion), see Nieddu and Pandolfi (2018).

observations. As shown in Table 11 in the Appendix, women included in this sample are on average 41.74 years old with the majority of them being older than 41 (67.2%). The probability of having a child is of 5.2% (higher compared to the probability of 4% observed on the full sample). About 87% of them have a scientific productivity above the cutoff point for being considered for the National Scientific Qualification, about 69% of them have obtained the Qualification as Associate Professor (this percentage rises to 79.6% among those whose productivity is above the threshold). On the other hand, only 22.6% of women applying for the NSQ have been effectively promoted to the position of Associate Professor.

In Table 7, we report first-stage estimation results in which the dummy *Qualification* is used as a dependent variable in relation to the dummy *Above 2/3 cutoffs* for passing 2 out of 3 productivity thresholds. Controlling for the distance from the three different cutoffs, having met at least two of them strongly determines the probability of obtaining the NSQ. More precisely, individuals who met at least two of the three productivity thresholds have a higher probability of obtaining the

**Table 7** First-stage results. The probability to obtain the qualification and the above 2/3 cutoffs

	(1)	(2)	(3)	(4)	(5)	(6)
Above 2/3 cutoffs	0.485*** (0.022)	0.484*** (0.023)	0.483*** (0.023)	0.483*** (0.023)	0.479*** (0.022)	0.476*** (0.023)
Distance from cutoff (indicator 1)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.001 (0.001)	0.001 (0.001)
Distance from cutoff (indicator 2)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Distance from cutoff (indicator 3)	-0.002* (0.001)	-0.002* (0.001)	-0.002* (0.001)	-0.002* (0.001)	-0.002*** (0.001)	-0.002** (0.001)
Age		-0.002 (0.002)	0.079** (0.033)	0.082** (0.034)	0.085** (0.034)	0.076** (0.034)
Age Sq.			-0.001** (0.000)	-0.001** (0.000)	-0.001*** (0.000)	-0.001** (0.000)
Experience (yrs.)				-0.004* (0.002)	-0.002 (0.002)	-0.004* (0.002)
Field dummies	Yes	Yes	Yes	Yes	Yes	Yes
Geographic area dummies	No	No	No	No	Yes	Yes
Year dummies	No	No	No	No	No	Yes
First-stage F-statistics	462.831	458.907	456.205	457.065	453.138	443.343
Observations	19,407	19,407	19,407	19,407	19,407	19,407
Clusters (individuals)	3986	3986	3986	3986	3986	3986

Notes: Estimates from first-stage regressions are reported in each column. The dependent variable is a dummy indicating whether the individual qualified as associate professor in the NSQ 2012. Sample includes all female assistant professors (RU) in Italian Universities as for 2012 who are aged up to 46. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR



NSQ of about 48 percentage points (the first-stage F-statistics is 443.343 in our most demanding specification in column 6).<sup>10</sup> The distance from indicator 1 is not statistically significant, while the distance from the second indicator is positive and statistically significant. The negative sign of the coefficient attracted by the third indicator might be driven by individuals in non-bibliometric fields as in these fields this indicator is represented by the number of books published in the last 10 years and often it happens that individuals with low research productivity publish their work as books.

In Table 8, we report results from the Two-Stage Least Squares estimation approach. In column 1, we do not control for individual covariates and find that having obtained the NSQ leads to an increase in the probability of having a child of about 1.6 percentage points. The effect, however, is not statistically significant at conventional levels.

Adding age and field dummies (column 2), age squared (column 3) and experience (column 4) reduces the effect to 1.2 percentage points. When adding geographic areas fixed effects (column 5) and years fixed effects (column 6) the magnitude of the estimated coefficient further reduces to 0.8 and to 0.6 percentage points, respectively. Importantly, this effect is in line with that in Table 6 for women who got promoted under the new regime that considers the NSQ as a pre-requisite, though it is not statistically significant at conventional levels, likely because of the reduced sample size (19,407 versus 80,100 observations). The estimated effect was of 0.9 percentage points when employing the individual fixed-effects model, while it becomes smaller (0.6 percentage points) when using the fuzzy RDD approach, which is reasonable considered that having acquired the NSQ is only the first step for promotion.

Finally, with the aim of investigating the impact of promotion on fertility using the Fuzzy Regression Discontinuity Approach described above, we have also experimented by instrumenting  $Promotion_{it}$ , instead of  $Qualification_{it}$ , with the dummy variable  $Above_{it}$  (see Table 9). As expected, first-stage estimation results confirm that individuals with a scientific productivity above the cutoffs are more likely to be promoted. More precisely, individuals who met at least two of the three productivity thresholds have a higher probability of being promoted of about 14 to 17 percentage points (the first-stage F-statistics is 133.13 in our most demanding specification in column 6). As regards the second-stage results, we again find a

---

<sup>10</sup> In Table 13 in the Appendix, we report reduced-form estimates. Also in this case results are in line with those discussed above. As expected the magnitude of the effects is smaller.

**Table 8** The effect of scientific qualification on child birth. Two-stages-least-squares results

	(1)	(2)	(3)	(4)	(5)	(6)
Qualified as Associate Professor	0.016 (0.012)	0.013 (0.012)	0.012 (0.012)	0.012 (0.012)	0.008 (0.012)	0.006 (0.012)
Distance from cutoff (indicator 1)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance from cutoff (indicator 2)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)
Distance from cutoff (indicator 3)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Age		-0.003*** (0.001)	0.015 (0.012)	0.013 (0.012)	0.013 (0.012)	0.011 (0.012)
Age Sq.			-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Experience (yrs.)				0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)
Field dummies	Yes	Yes	Yes	Yes	Yes	Yes
Geographic area dummies	No	No	No	No	Yes	Yes
Year dummies	No	No	No	No	No	Yes
Observations	19,407	19,407	19,407	19,407	19,407	19,407
Clusters (individuals)	3986	3986	3986	3986	3986	3986

Notes: Estimates from two-stages-least-squares regressions are reported in each column. Dependent variable is *Child Birth*. The endogenous variable Qualified as Associate Professor is instrumented with the Above 2/3 Cutoffs. Sample includes all female assistant professors (RU) in Italian Universities as for 2012 who are aged up to 46. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR

positive effect of promotion on the probability of having a child. The effect, even if imprecisely estimated, is larger in magnitude compared to the one obtained in Table 8, consistent with the fact that we are now looking at the effective promotion to associate professor, while before we were considering the effect of the qualification for an assistant professor position.

**Table 9** The effect of promotion on child birth. Two-stages-least-squares results

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Second-stage estimation results</i>						
Promotion to Associate Professor	0.049 (0.037)	0.035 (0.034)	0.034 (0.034)	0.034 (0.034)	0.022 (0.034)	0.021 (0.040)
Distance from cutoff (indicator 1)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance from cutoff (indicator 2)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)
Distance from cutoff (indicator 3)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Age		-0.004*** (0.001)	0.012 (0.012)	0.011 (0.012)	0.011 (0.012)	0.011 (0.012)
Age Sq.			-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Experience (yrs.)				0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
<i>Panel B: First-stage estimation results</i>						
Above 2/3 cutoffs	0.161*** (0.012)	0.173*** (0.013)	0.172*** (0.013)	0.172*** (0.013)	0.168*** (0.013)	0.142*** (0.012)
First-stage F-statistics	175.707	174.973	172.713	171.856	165.398	133.126
Field dummies	Yes	Yes	Yes	Yes	Yes	Yes
Geographic area dummies	No	No	No	No	Yes	Yes
Year dummies	No	No	No	No	No	Yes
Observations	19,407	19,407	19,407	19,407	19,407	19,407
Clusters (individuals)	3986	3986	3986	3986	3986	3986

Notes: Estimates from two-stages-least-squares regressions are reported in each column. In the second stage the dependent variable is *Child Birth*. In the first stage the dependent variable is *Promotion*. Sample includes all female assistant professors (RU) in Italian Universities as for 2012 who are aged up to 46. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR

## 6 Concluding Remarks

It is well documented that the shortfall of women in top academic positions is at least partially due to a family–work conflict since these jobs entail high effort and time which are incompatible with family related necessities. This conflict seems to induce many women to either sacrifice family or career.

While many papers have documented negative effects of fertility on a woman's career, we took a different approach by looking at the impact of improved career prospects on the decision to have a child. To this purpose, we use administrative data on the universe of female Assistant Professors employed in Italian universities from 2001 to 2018 and estimate an individual fixed-effects model to capture the effect of promotion to Associate Professor on fertility. Our results document that promotion to associate professor increases the probability of having a child by 0.6 percentage points, which translates into an increase by 12.5% of the mean.

The effect of promotion on fertility could be determined by a higher income available (the promotion determines an increase of about 35% of the disposable income): this is in line with the findings of Modena et al. (2013) that show that 40–50% of Italian couples were discouraged to have (more) children because of an insufficient income.

An alternative explanation for the positive effect of promotion we find in the present analysis could be the following: during the phase of assistant professorship, women postpone fertility—since having children is very time consuming—and devote their time and energy to scientific research in order to increase their scientific productivity and raise their probability of promotion. Once they are promoted, they have the possibility to have (more) children minimizing negative effects on their careers.

The effect of promotion on fertility we estimate is robust to various sensitivity tests, including a specification that allows either age or years of experience to enter non-linearly in order to flexibly control for the fact that both promotion and maternity could be related to age or seniority, respectively. In addition, we document that the effect mainly occurs immediately after promotion and gradually vanishes with the number of years from promotion. Also, we find that the impact of promotion is higher for women aged below 40, and for those who work in a university which is located in the North of Italy. Furthermore, we find that the impact is stronger under the new university regulation that, starting from 2012, considers the NSQ as a pre-requisite for career advancements. In particular, we show that promotion to associate professor under the new regime increases the likelihood of having a child by almost 1 percentage point, that implies an increase by 20% of the mean.

Our empirical analysis shows positive effects of promotion also when using a Fuzzy Regression Discontinuity Design in which we exploit the eligibility requirements in terms of research productivity introduced since 2012 in the system regulating career advancement in Italian academia. In this econometric framework the credibility of our identification strategy is increased since we are able to compare the fertility behaviour of very similar women: those who just pass the NSQ productivity thresholds and those who just miss them. We find that women who obtain the NSQ—and therefore increase substantially the probability of being

promoted to Associate Professor in the near future—have a 0.6 percentage point higher probability of having a child, though the effect is imprecisely estimated due to the reduced sample size. This effect is similar in magnitude to the one obtained when looking at promotion since 2012 using individual fixed-effects regression analysis, suggesting that our main results are unlikely driven by omitted variable bias.

Our findings suggest that policies aimed at improving women career prospects are important not only to increase productivity and enhance equal opportunities but also to help increasing fertility. This could be very important for all OECD countries currently plagued by very low fertility rates.

**Acknowledgements** We are grateful to ANVUR (National Agency for the Evaluation of the University and Research Systems) for allowing us to use the dataset (in anonymized form) at its Laboratory. We would also like to thank Daniele Checchi, Massimiliano Bratti and seminar participants at ANVUR conference (“III Concorso Pubblico Idee di Ricerca”, Workshop Finale, Roma, 12 November 2019) for useful comments and suggestions. We acknowledge funding from ANVUR (III Concorso Pubblico di Idee di Ricerca). The usual disclaimer applies.

## Appendix

**Table 10** Descriptive statistics—individual fixed-effects approach

	Mean	SD	Min	Max
Child birth	0.040	0.196	0	1
Promotion to Associate Professor	0.148	0.355	0	1
Age	40.175	4.145	24	46
Aged 24 to 30	0.015	0.120	0	1
Aged 31 to 35	0.138	0.345	0	1
Aged 36 to 40	0.331	0.470	0	1
Aged 41+	0.517	0.500	0	1
Experience (yrs.)	6.051	4.165	0	25
Years from promotion	0.561	1.703	0	16
North	0.461	0.498	0	1
Centre	0.218	0.413	0	1
South	0.322	0.467	0	1

Source: Administrative data provided by ANVUR

Sample: only women; aged $\leq$ 46; Assistant Professor in 2001 or later (we exclude women who were already Associate and Full Professors in 2001). Total Observations (individual\*year): 101,774; Individuals: 11,897

**Table 11** Descriptive statistics—Fuzzy regression discontinuity approach

	Mean	SD	Min	Max
Child birth	0.053	0.224	0	1
Promotion to Associate Professor	0.226	0.175	0	1
Qualified as Associate Professor	0.693	0.461	0	1
Above 2/3 cutoffs	0.870	0.336	0	1
Distance from cutoff (indicator 1)	8.760	249.103	−51	503.080
Distance from cutoff (indicator 2)	17.340	129.021	−69.708	959.164
Distance from cutoff (indicator 3)	3.621	124.342	−12	39
Age	41.744	3.217	24	46
Aged 24 to 30	0.0005	0.021	0	1
Aged 31 to 35	0.042	0.201	0	1
Aged 36 to 40	0.285	0.451	0	1
Aged 41+	0.673	0.469	0	1
Experience (yrs.)	7.943	3.539	0	24
North	0.491	0.500	0	1
Centre	0.202	0.402	0	1
South	0.306	0.461	0	1

Source: Administrative data provided by ANVUR

Sample: only women, aged  $\leq 46$ , Assistant Professor in 2012 followed in subsequent years (we exclude women who were already Associate and Full Professors in 2012). Total Observations (individual\*year): 19,407; Individuals: 3, 986

**Table 12** The effect of promotion on child birth. Linear probability model (LPM)

	(1)	(2)	(3)	(4)	(5)	(6)
Promotion to Associate Prof.	0.010*** (0.002)	0.010*** (0.002)	0.002 (0.002)	0.001 (0.002)	−0.001 (0.002)	−0.000 (0.002)
Age	−0.004*** (0.000)	−0.001 (0.003)	−0.001 (0.003)	−0.004 (0.003)	−0.005* (0.003)	−0.005* (0.003)
Age Sq.		−0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)
Experience (yrs.)			0.002*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.000* (0.000)
Year dummies	No	No	No	Yes	Yes	Yes
Geographic area dummies	No	No	No	No	Yes	Yes
Field dummies	No	No	No	No	No	Yes
Observations	101,774	101,774	101,774	101,774	101,774	101,774
Clusters (individuals)	11,897	11,897	11,897	11,897	11,897	11,897

Notes: Estimates from OLS regressions are reported in each column. The dependent variable is *Child Birth*. Sample includes all female hired as assistant professors (RU) in Italian Universities after 2001 who are aged up to 46 followed until 2018. Standard errors clustered by individual are reported in parentheses. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . Source: Administrative data provided by ANVUR

**Table 13** The effect of scientific qualification on child birth. Reduced-form results

	(1)	(2)	(3)	(4)	(5)	(6)
Above 2/3 cutoffs	0.008 (0.006)	0.006 (0.006)	0.006 (0.006)	0.006 (0.006)	0.004 (0.006)	0.003 (0.006)
Distance from cutoff (indicator 1)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance from cutoff (indicator 2)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)
Distance from cutoff (indicator 3)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Age		-0.003*** (0.001)	0.016 (0.012)	0.014 (0.012)	0.013 (0.012)	0.011 (0.012)
Age Sq.			-0.000* (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Experience (yrs.)				0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)
Field dummies	Yes	Yes	Yes	Yes	Yes	Yes
Geographic area dummies	No	No	No	No	Yes	Yes
Year dummies	No	No	No	No	No	Yes
Observations	19,407	19,407	19,407	19,407	19,407	19,407
Clusters (individuals)	3986	3986	3986	3986	3986	3986

Notes: Estimates from reduced-form regressions are reported in each column. Sample includes all female assistant professors (RU) in Italian Universities as for 2016 who are aged up to 46. Dependent variable is a dummy indicating whether the individual has a child birth. Standard errors clustered by individual are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Source: Administrative data provided by ANVUR

## References

Bagues, M., Sylos-Labini, M., & Zinovyeva, N. (2017). Does the gender composition of scientific committee matter? *American Economic Review*, 107(4), 1207–1238.

Becker, G. (1981). *A treatise on the family*. Harvard University Press.

Bertrand, M. (2018). Coase lecture—the glass ceiling. *Economica*, 85(338), 205–231.

Blackaby, D., Booth, A., & Frank, J. (2005). Outside offers and the gender pay gap: Empirical evidence from the UK academic labour market. *The Economic Journal*, 115(501), F81–F107.

Bratti, M., Del Bono, E., & Vuri, D. (2005). New mothers’labour force participation in Italy: The role of job characteristics. *Labour*, 19(1), 79–121.

Cigno, A. (1991). *Economics of the family*. Oxford University Press.

Currie, J., & Schwandt, H. (2014). Short- and long-term effects of unemployment on fertility. *Proceedings of the National Academy of Sciences*, 111(41), 14734–14739.

De la Rica, S., & Iza, A. (2005). Career planning in Spain: Do fixed-term contracts delay marriage and parenthood? *Review of the Economics of the Household*, 3, 49–73.

De Paola, M., Nisticò, R., & Scoppa, V. (2021). Employment protection and fertility decisions: The unintended consequences of the Italian Jobs Act. *Economic Policy*, 36(108) 735–773. <https://doi.org/10.1093/epolic/eiab015>

- De Paola, M., Ponzo, M., & Scoppa, V. (2017). Gender differences in the propensity to apply for promotion: Evidence from the Italian scientific qualification. *Oxford Economic Papers*, 69(4), 986–1009.
- De Paola, M., & Scoppa, V. (2015). Gender discrimination and evaluators' gender: Evidence from Italian academia. *Economica*, 82(325), 162–188.
- Del Boca, D., & Sauer, R. (2009). Life cycle employment and fertility across institutional environments. *European Economic Review*, 53(3), 274–292.
- Del Bono, E., Weber, A., & Winter-Ebmer, R. (2015). Fertility and economic instability: The role of unemployment and job displacement. *Journal of Population Economics*, 28(2), 463–478.
- Francesconi, M. (2002). A joint dynamic model of fertility and work of married women. *Journal of Labor Economics*, 20, 336–380.
- Ginther, D. K., & Kahn, S. (2004). Women in economics: Moving up or falling off the academic career ladder? *Journal of Economic Perspectives*, 18(3), 193–214.
- Jappelli, T., Nappi, C. A., & Torrini, R. (2017). Gender effects in research evaluation. *Research Policy*, 46, 911–924.
- McDowell, J. M., Singell, L. D., & Ziliak, J. P. (1999). Cracks in the glass ceiling: Gender and promotion in the economics profession. *American Economic Review*, 89(2), 392–396.
- Modena, F., Rondinelli, C., & Sabatini, F. (2013). Economic insecurity and fertility intentions. *The Case of Italy, Review of Income and Wealth*, 60(1), 233–255.
- Nieddu, M., & Pandolfi, L. (2018). *The effectiveness of promotion incentives for public employees: Evidence from Italian academia*. CSEF Working Paper, 507.
- Prifti, E., & Vuri, D. (2013). Employment protection and fertility: Evidence from the 1990 Italian reform. *Labour Economics*, 23, 77–88.
- Sobotka, T., Skirbekk, V., & Philipov, D. (2011). Economic recession and fertility in the developed world. *Population and Development Review*, 37(2), 267–306.
- te Velde, E., Habbema, D., Leridon, H., & Eijkemans, M. (2012). The effect of postponement of first motherhood on permanent involuntary childlessness and total fertility rate in six European countries since the 1970s. *Human Reproduction*, 27(4), 1179–1183.
- Ward, K., & Wolf-Wendel, L. (2004). Academic motherhood: Managing complex roles in research universities. *Review of Higher Education*, 27, 233–257.

**Maria De Paola** (Ph.D. University of Rome) is Professor of Economics at the University of Calabria and currently on leave at the Research Department of INPS (Italian Social Security Administration). Her research interests include labor economics, education economics, gender, experimental economics, political economy, and evaluation of public policies.

**Roberto Nisticò** (Ph.D. University of Essex) is Associate Professor of Economics at the University of Naples Federico II. His research interests are in development, labor and political economics.

**Vincenzo Scoppa** (Ph.D. University of Siena) is Professor of Economics at the University of Calabria. His research interests are in labor economics, education economics, policy evaluations, and economics of sports.



**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made. The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.



**Part IV**  
**Conformism in Research**

# Social Network Tools for the Evaluation of Individual and Group Scientific Performance



**Domenico De Stefano, Luka Kronegger, Valerio Leone Sciabolazza, Maria Prosperina Vitale, and Susanna Zaccarin**

**Abstract** Over the past few decades, scientific collaboration has been widely considered an important driver of research innovation. By collaborating together, scientists can benefit from both methodological and technological complementarities and synergy, improving the quality and quantity of their research outputs. As evidence of this, collaboration among scientists is increasing in all disciplines and government policies in international exchange programs are aimed at promoting collaboration among researchers. Collaboration among scientists can be represented as a network, usually adopting co-authorship as linkages. In this view, Social Network Analysis provides a useful theoretical and methodological approach because collaboration features can be related to the topological characteristics of the network. Recently, several empirical studies have found positive associations between researchers' position in the co-authorship network and their productivity, although the results can be different depending on the discipline, scientific performance measure, and data source retrieved to construct the co-authorship networks. In this contribution, we propose the use of SNA tools for scientific evaluation purposes. Network indices at the individual and subgroup levels will be introduced

---

D. De Stefano (✉)

Department of Political and Social Sciences, University of Trieste, Trieste, Italy  
e-mail: [ddestefano@units.it](mailto:ddestefano@units.it)

L. Kronegger

Faculty of Social Sciences, University of Ljubljana, Ljubljana, Slovenia  
e-mail: [luka.kronegger@fdv.uni-lj.si](mailto:luka.kronegger@fdv.uni-lj.si)

V. Leone Sciabolazza

Department of Economics and Law, Sapienza University of Rome, Rome, Italy  
e-mail: [valerio.leonesciabolazza@uniroma1.it](mailto:valerio.leonesciabolazza@uniroma1.it)

M. P. Vitale

Department of Political and Social Studies, University of Salerno, Fisciano, Italy  
e-mail: [mvitale@unisa.it](mailto:mvitale@unisa.it)

S. Zaccarin

Department of Economics, Business, Mathematics and Statistics 'B. de Finetti', University of Trieste, Trieste, Italy  
e-mail: [susanna.zaccarin@deams.units.it](mailto:susanna.zaccarin@deams.units.it)

© The Author(s) 2022

D. Checchi et al. (eds.), *Teaching, Research and Academic Careers*,  
[https://doi.org/10.1007/978-3-031-07438-7\\_7](https://doi.org/10.1007/978-3-031-07438-7_7)

to analyze the relation with both the individual research productivity and scientific output quality measure provided by the Italian academic researchers involved in VQR from the period 2011–2014.

## 1 Introduction

Over the past few decades, scientific collaboration has been considered an important driver of research progress that supports researchers in generating novel ideas (see, among others, Beaver 2001). The role of scientific collaboration has been emphasized in recent government policies and international exchange programs that aim at stimulating the mobility of researchers and fostering scientific collaboration and productivity (Wuchty et al. 2007; Defazio et al. 2009; Leone Sciabolazza et al. 2020). Recently, university administrations and research funders have explored a variety of programs and policies to stimulate interdisciplinary collaboration. Among them, it is worth recalling the funding initiatives targeting: interdisciplinary projects, such as the INSPIRE program of the US National Science Foundation (NSF), the Interdisciplinary Research Consortia program of the US National Institutes of Health (National Institute of Health 2007), the EU funding research network (Commission of European Communities 2006), and the national Spanish Ingenio 2010 Program (Ministry of Education and Science 2006); interdisciplinary training programs such as the NSF Integrative Graduate Education and Research Traineeship (IGERT); and interdisciplinary university fellowship programs (Sà 2008). Scientific collaboration has also been recognized as a key factor in measuring and evaluating scholars' scientific performance (Ferligoj et al. 2015; De Stefano & Zaccarin 2016).

Moving from this perspective, this chapter aims at presenting the main results of the SnEval (Social Network tools for the Evaluation of individual and group scientific performance) research project. The main contribution of the project was to show novel results based on a network analysis on the ANVUR VQR data. The proposed methodology can be adopted in future research evaluation exercises.

The analysis focuses on the co-authorship networks among academic scholars in two research areas of the Italian university system, namely Area 2—Physics and Area 13—Economics and Statistics. These areas have different characteristics in the evaluation exercise. In particular, Area 2 is classified as a fully “bibliometric” area,<sup>1</sup> that is, the majority of scientific products in the area are published in international journals and bibliometric indicators (journal metrics and citations indicators) are commonly used for evaluation purposes. Conversely, Area 13 is classified as a “non-bibliometric” area. Although a few disciplines in this area are characterized by bibliometric-like publication behavior, the evaluation of scientific products is performed mainly by a peer-review process (or informed peer review where

---

<sup>1</sup> In the Italian Evaluation exercise, scientific disciplines are divided into bibliometric and non-bibliometric areas (however, each SSD has its own evaluation committee that can choose the criteria on which the evaluation is performed).

reviewers additionally take into account a number of bibliometric indicators). Co-authorship information has been derived from the scientific products scholars in the two areas submitted for the VQR exercise in the period 2011–2014. Co-authorship networks were built at different levels of the official aggregations (macro-sectors and meso-sectors or *Settore concorsuale*) of the disciplines<sup>2</sup> belonging to the two abovementioned areas in the VQR exercise period (see Sect. 4).

We then selected some of the most appropriate network indices—at the individual and subgroup network levels—and some useful techniques (as described in Sect. 3) to disentangle the different publication and collaboration styles characterizing the two areas. These indices are used both to characterize the structure of the disciplines and to look for their effect on the quality of the research outputs. More specifically, we compared the different co-authorship networks by considering their topology and authors' position. This analysis is crucial to understand how the authors are related and how the collaboration patterns change across time and between disciplines. In particular, we considered the structural properties of the observed networks and their local characteristics. Furthermore, we fitted a regression model to provide empirical evidence of the relation between the network results, here at the author and network levels, and the VQR scores at the individual level, the latter representing the “dependent variables” in the model.

The results suggest that even in the Italian scenario, it would be worth fostering intra and interdisciplinary collaboration to improve group and individual productivity. We show how the proposed analytical tools can provide useful insights on the co-authorship network topology and detect those researchers in certain structural positions who can be the target of some network-based interventions (for instance, in scale-free networks, few important nodes act as hubs). Furthermore, the fitted models affirm that the researchers in a central position in the co-authorship network are also those scholars whose performance is significantly higher than the researchers in a more peripheral position.

## 2 Related Literature

Collaboration in science is a complex phenomenon that affects scientific productivity in various ways, as well as knowledge diffusion within and between disciplines.

It is straightforward to represent collaboration among scientists as a network, in which the nodes are scholars tied by the various forms of scientific collaboration among them. In this view, Social Network Analysis (SNA) (Wasserman & Faust 1994) provides a useful theoretical and methodological approach for studying collaboration among these individuals. Because collaboration features can be related to the network properties, this approach can help in the understanding of the structure and the evolution of research collaboration over topics and time (Yan &

---

<sup>2</sup> <https://www.miur.gov.it/settori-concorsuali-e-settori-scientifico-disciplinari>.

Guns 2014), as well as to cluster researchers and determine research groups (see, among others, Mali et al. 2012 and their related references).

Most of the empirical studies on scientific collaboration mainly refer to the analysis of co-authorship networks, with co-authorship ties being used as a proxy of scholars' collaborative behavior (Ponomariov & Boardman 2016). The increasing availability of electronic databases allows for good-quality data on co-authorship to be collected in a relatively inexpensive way. Over the past few decades, several SNA co-authorship studies have been carried out in various fields. Among them, seminal papers can be found in Albert and Barabási (2002) and Newman (2004) for physics and biomedical research, in Goyal et al. (2006) for economics, and in Moody (2004) for sociology. More recently, Abel et al. (2019) investigated the driving factors behind co-authorship both within and across institutions among demographers. The common aims of these network-based studies were understanding of the topological properties of networks and their implications for the evolution of topics and methods. For instance, a "small-world" pattern (Watts 1999) can support disciplinary fragmentation and specialty areas that are clustered into distinct groups of scientists, mainly because of scientists' research group membership, university affiliations, or geographic proximity. On the contrary, a broad connectivity among a large proportion of scientists can suggest theoretical integration, while more centralized structures that are driven by few highly connected scientists (usually called "stars") can imply the existence of a peculiar tie formation mechanism named "preferential attachment" (Albert & Barabási 2002). Clear evidence of the presence of small-world properties has been observed in the fields of economics (Goyal et al. 2006; Maggioni & Uberti 2011) and physics (Newman 2004). Physics, mathematics, neurosciences (Albert & Barabási 2002), and economics (Goyal et al. 2006) have also shown statistical properties consistent with a preferential attachment mechanism. Sociology is the one exception because it is better represented by an integrated (cohesive) collaboration network structure resembling a random network (Moody 2004).

Co-authorship networks can also be exploited to predict the scientific performance of researchers, that is, evaluating the effect of actors' embeddedness in co-authorship networks and their individual research outputs (Abbasi et al. 2011). Several empirical studies found positive correlations between researchers' position in the co-authorship network and their productivity (e.g., see Fischbach et al. 2011; Abbasi et al. 2012; Uddin et al. 2013; Ferligoj et al. 2015), even if the results depend on disciplines and by the measures used for scientific productivity or scientific performance (Melin 2000; Lee & Bozeman 2005), as well as by the characteristics of the data sources retrieved to construct the co-authorship networks (De Stefano & Zaccarin 2016).

A myriad of studies also focuses on specific scientific communities at the country level. Among them, see, for example, the contributions of Kronegger et al. (2012) on Slovenian scientists, Digiampietri et al. (2017) on Brazilian PhDs working in probability and statistics field, and Leone Sciabolazza et al. (2017) on researchers hired at the University of Florida. In Italy, Maggioni and Uberti (2011) analyzed co-authorship networks among academic economists, while De Stefano et al. (2013)

and Fuccella et al. (2016) studied academic statisticians. Bellotti (2012) considered the links among Italian physicists participating in funded national projects, and Bellotti et al. (2016) extended the analysis to several disciplines in Italian academia. Abramo et al. (2018) examined the collaboration behavior of stars and top scientists among Italian academic scientists, while gender and academic rank differences in collaboration were analyzed, respectively, in Abramo et al. (2014) and Abramo et al. (2019).

### 3 Basic Concepts on Networks

The basic notations and concepts to formally describe a co-authorship network in the SNA context are presented below. Co-authorship data are extracted from a set of authors and their papers and are arranged in an affiliation matrix that represents a bipartite network (i.e., two-mode network).

Let  $\mathcal{N} = \{1, 2, \dots, n\}$  be the set of  $n$  authors and  $\mathcal{P} = \{1, 2, \dots, p\}$  the set of  $p$  papers observed on  $n$  authors. An affiliation matrix  $\mathbf{A}(n \times p)$  *author-by-paper* is defined with the elements  $a_{ik}$ , assuming a value of 1 if  $i \in \mathcal{N}$  authored the paper  $k \in \mathcal{P}$ , and 0 otherwise. The co-authorship network is derived from the matrix product  $\mathbf{Y} = \mathbf{A}\mathbf{A}'$ , which helps in defining the undirected and valued  $n \times n$  adjacency matrix (i.e., one-mode network)  $\mathbf{Y}$  *author-by-author*. The element  $y_{ij}$  of  $\mathbf{Y}$  is greater than 0 if  $i, j \in \mathcal{N}$  co-authored one or more papers in  $\mathcal{P}$ , and  $y_{ij} = 0$  otherwise. The relations embedded in  $\mathbf{Y}$  can be represented by a graph  $G(\mathcal{N}, \mathcal{L})$ , that is, a collection of a set  $\mathcal{N}$  of nodes (authors in our case) connected by the set  $\mathcal{L}$  of their links (co-authorship relationships). The cardinality of  $\mathcal{L}$  is  $l = |\mathcal{L}| = 1/2 \sum_i \sum_j y_{ij}$ ,  $\forall i \neq j$ .

Several network statistics at the global and individual levels have been defined both to describe the structural characteristics of  $G$  and to test the consistency of  $G$  with theoretical network structures that have well-known topological features and properties.

The most basic network statistics at the global level is the density  $\Delta(G) = 2l/n(n-1)$ , which measures the cohesion of  $G$ . When  $\Delta(G) \approx 0$ ,  $G$  is said to be sparse. The network connectivity is described by the average path length  $\ell(G)$ , which is defined as the average number of links along the shortest paths (geodesic distance  $\ell(i, j)$ ) for all possible  $\binom{n}{2}$  pairs of nodes (Watts 1999). The largest  $\ell(i, j)$  over all pairs of nodes is called the diameter of  $G$ . In the presence of disconnected graphs,  $\ell(G)$  is computed on the so-called giant component, which is the largest subgraph in terms of the number of reachable connected nodes (that is a path connecting two randomly selected nodes).

Besides global network statistics, node-level centrality indices refer to the position of each node (or actor) in the network according to various definitions of “centrality”. The most used centrality measures are: degree, closeness, and betweenness (Freeman 1979). The degree  $d_i$  of the  $i$ -th node is the basic one among these measures. It expresses the number of links that  $i$  has with the other nodes in the network. If  $d_i = 0$ , the node  $i$  is isolated; on the contrary, if  $d_i = n - 1$

(the maximum value for degree), the node is the most central one in terms of its overall connectivity. In co-authorship networks,  $d_i$  indicates the number of distinct co-authors of the  $i$ -th author. Denoting with  $\sigma_{j,k} = \sigma_{k,j}$  the number of shortest paths from node  $j$  to  $k$  and with  $\sigma_{j,k}(i)$  the number of those shortest paths passing through  $i$ , a further centrality measure is the so-called betweenness  $b_i$ . This is related to the bridging role of an actor and his/her potential to control the flow of information or the exchange of resources (e.g., knowledge). The authors with large betweenness values denote the propensity to connect otherwise disconnected groups of researchers (e.g., connecting members of different labs or departments). In general, in network studies, high betweenness is observed for the authors with high interdisciplinary behavior.

Let  $\mathcal{N}_i = \{v : \ell(i, v) = 1\}$  be the neighborhood of the  $i$ -th node, so that  $i \notin \mathcal{N}_i$ . A measure of the overlap between the links of distinct nodes in  $\mathcal{N}_i$  is the (local) clustering coefficients (also called transitivity) (Fronczak et al. 2003) of the  $i$ -th node:  $\Gamma_i = |\mathcal{L}_i| / \binom{d_i}{2}$ , where  $|\mathcal{L}_i|$  is the total number of links in the subnetwork  $\mathcal{N}_i$ . The network clustering coefficient  $\Gamma(G)$  is defined as the average of the  $\Gamma_i$ ,  $\forall i \in \mathcal{N}$ .  $\Gamma(G)$  represents the average number of closed triplets of nodes (triangles) in the network out of the total number of triads, that is, arbitrary connected (or disconnected) triplets of nodes. Hence, this measure captures the extent to which authors are embedded in cohesive clusters characterized by high collaboration. High  $\Gamma(G)$  is a characteristic associated with the so-called small-world behavior in networks.

Furthermore, following the procedure proposed by Albert et al. (2000), degree centrality can also be used to analyze the extent to which most connected authors (i.e., scientists with the highest degree centrality) are crucial for the connectivity in the network. To this purpose, the consequences of deleting nodes at random and nodes that are highly connected could also be investigated.

The interest in the analysis of co-authorship networks lies in the fact that collaborative behavior within a scientific community closely depends on the network topological features. In particular, a frequent finding in co-authorship networks is that they are consistent with some theoretical network models with well-defined topological and relational properties, which have a meaningful interpretation in terms of knowledge diffusion in specific discipline. The simplest network models start with the idea that the connections between actors occur at random, as in the Erdos–Renyi (ER) random graphs, a family of networks in which the probability of a tie between actors' pairs is equal to  $p$ , independently of the rest of the network and actor neighborhood (i.e., actors do not have any preference to connect with other nodes). This model represents the baseline model for assessing evidence of non-random behaviors in the observed co-authorship networks.

Empirical evidence shows that co-authorship networks are usually nonrandom because they tend to exhibit distinctive statistical properties deriving from peculiar attachment mechanisms among authors. In particular, scale-free (Albert & Barabási 2002) and small-world (Watts 1999) configurations are the theoretical models that most frequently emerge in a co-authorship analysis.



Looking at the degree distribution, that is, the frequency distribution of the number of co-authors per author, if a power law distribution is observed, then there is evidence for the emergence of a *scale-free* structure in the network. This implies the existence of a peculiar tie formation mechanism named preferential attachment, which formally accounts for the tendency to interact with the best connected authors (i.e., the authors with the highest degree, usually called stars or hubs). A strategy to test if the degree distribution of the network is consistent with a power law distribution is provided in Clauset et al. (2009). This strategy generates ER random networks equivalent to the observed network to check the departure from the pure randomness of the co-authorship network under study.

The *small-world* configuration, instead, describes the simultaneous presence of dense local clustering (i.e., high value of a clustering coefficient) with short network distances (i.e., shortest path length) that can facilitate knowledge flows inside a network. In a co-authorship network, this means that there exist small cohesive groups of researchers with few connections between them that strategically reduce the overall distance among actors. Specifically, networks consistent with this topology have high node connectivity with a low average distance among regions of the network, that is, the average path length is not greater than the value observed in random networks of equal size together with a high tendency toward author clustering. Also, in small-world structures, the diameter is lower than the one observed in ER graphs.

## 4 The VQR 2011–2014 Data

The data used in this analysis are bibliographic information (authors, co-authors, and the quality of the paper according to the assigned VQR scores) derived from the publications submitted by the academic researchers for the evaluation exercise VQR 2011–2014, which assesses the quality of the scientific products published in the period 2011–2014. According to the official governmental Italian classification, scientific disciplines are classified in several research areas. For the VQR, these areas are divided into bibliometric and non-bibliometric classes, depending on the use of bibliometric indicators for the research quality assessment or the use of peer-review mechanism, respectively). In particular, we analyze Area 2 (Physics) and Area 13 (Economics and Statistics) scientific areas.

Area 2 comprises four macro-sectors: Physics of Fundamental Interactions (02/A), Physics of Matter (02/B), Astronomy, Astrophysics Earth and Planetary Physics, Applied Physics (02/C), and Physics Teaching and History of Physics (02/D). Each macro-sector encompasses one or two meso-sectors. Meso-sector 02/C1 is associated with a unique micro-sector, while each of the remaining

**Table 1** Scientific areas 2 (Physics) and 13 (Economics and Statistics) and macro and meso-sectors according to the official Italian classification

Area	Macro-sector	Meso-sector
Area 02	02/A	Experimental Physics of Fundamental Interactions (02/A1)
		Theoretical Physics of Fundamental Interactions (02/A2)
	02/B	Experimental Physics of Matter (02/B1)
		Theoretical Physics of Matter (02/B2)
	02/C	Astronomy, Astrophysics Earth and Planetary Physics, Applied Physics (02/C1)
02/D	Physics Teaching and History of Physics (02/D1)	
Area 13	13/A	Economics (13/A1)
		Economic Policy (13/A2)
		Public Economics (13/A3)
		Applied Economics (13/A4)
		Econometrics (13/A5)
	13/B	Business Administration and Accounting Studies (13/B1)
		Management (13/B2)
		Organization and Human Resource Management (13/B3)
		Financial Markets, Financial Institutions, and Corporate Finance (13/B4)
		Commodity Sciences (13/B5)
	13/C	Economic History (13/C1)
	13/D	Statistics (13/D1)
		Economic Statistics (13/D2)
		Demography and Social Statistics (13/D3)
Mathematical Methods of Economics, Finance and Actuarial Studies (13/D4)		

meso-sectors comprises two micro-sectors. This categorization is non-mutually exclusive, meaning that a researcher can be affiliated with multiple micro- and meso-sectors.<sup>3</sup>

Similarly, Area 13 is composed of four macro-sectors: Economics (13/A), Business Administration and Management (13/B), Economic History (13/C), and Statistics and Mathematical Methods for Decisions (13/D). Also, in this case, each macro-sector consists of one or more meso-sectors. The details of the classification of both scientific areas and the corresponding macro and meso-sectors are reported in Table 1

As a result of the evaluation, a VQR score on a 5-point scale, as shown in Table 2, was assigned to each product submitted by the academic researchers. In

<sup>3</sup> In particular, the researchers belonging to the SSD FIS/01 (experimental physics micro-sector) are associated either to the 02/A1 or to the 02/B1 meso-sector. Since we have anonymized data, we choose to allow for the 02/A1 (macro-sector 02/A) to include all members of the experimental physics micro-sector.

**Table 2** Labels and associated numerical scores in the VQR 2011–2014 evaluation exercise

Label	Score
Eccellente (excellent)	1
Elevato (high)	0.7
Discreto (fairly good)	0.4
Accettabile (acceptable)	0.1
Limitato o non valutabile (limited or not evaluable)	0

**Table 3** Average VQR scores at the macro-sector level

Macro-sector	VQR score (mean)
02/A	0.85
02/B	0.83
02/C	0.81
02/D	0.72
13/A	0.63
13/B	0.41
13/C	0.42
13/D	0.57

*Note:* For each macro-sector, we report the average VQR scores (see Table 2). The macro-sector relative to the code is indicated in Sect. 4

Table 3, the average VQR scores for each analyzed macro-sectors are reported. We can notice a slightly higher overall VQR performance of the physics macro-sectors with respect to the economics and statistics macro-sectors. A transformation of such scores, representing the “excellence” of the individual research outputs, will be used as the dependent variable in the regression model illustrated in Sect. 5.4.

The co-authorship networks at different levels of aggregation (macro- and meso-sectors) are built by retrieving all co-authors from the scientific production submitted for the evaluation exercise. To this end, it is worth noting that, on average, we observe two publications per author. In fact, in the VQR, each evaluated researcher should submit at most two scientific products (however, some researchers can appear as a co-author in a paper submitted by someone else). For this reason, the co-authorship network under analysis is a sample of the overall co-authorship networks among Area 2 and Area 13 researchers. Despite this limitation, we can consider these co-authorship networks as determined by the most significant production according to each researcher’s auto-evaluation.

#### 4.1 Co-authorship Networks

For each macro- and meso-sector considered, we create a co-authorship network, where each node indicates a researcher involved in the VQR exercise, and a link registers the presence and intensity of the collaboration between the two of them.

**Table 4** Areas 2 and 13 macro-sectors—descriptive table

Macro-sector	02/A	02/B	02/C	02/D	13/A	13/B	13/C	13/D
Number of nodes	389	523	580	87	227	552	542	36
Number of edges	194	278	329	35	3	110	184	2
Density*1000	1.3	1.0	1.0	4.6	0.1	0.4	0.6	1.5
Nodes making up to 50% of collaborations	21	28	36	9	1	25	24	1
Nodes making up to 90% of collaborations	46	60	78	16	2	53	56	2
Degree—average	1	1.06	1.13	0.8	0.03	0.4	0.68	0.11
Degree—standard deviation	1.29	1.42	1.35	0.82	0.19	0.65	1.02	0.32
Giant component—size	12	10	12	5	3	5	14	2
Giant component—percentage	3%	2%	2%	6%	1%	1%	3%	6%
Isolated authors—number	182	258	247	38	222	378	311	32
Isolated authors—percentage	47%	49%	43%	44%	98%	68%	57%	89%
Clustering coefficient in giant component	0.56	0.58	0.56	0	0	0.6	0.57	0
Distance in giant component: average	4.5	3.76	3.85	3.2	1.33	2.2	3.15	2
Distance in giant component: standard deviation	2.44	2.04	2.07	1.82	0.52	1.01	1.86	0

*Note:* For each macro-sector network, indicated with its relative code (in columns), we report a number of descriptive metrics. For a precise definition of metrics, refer to Sect. 3. The macro-sector relative to a code is indicated in Sect. 4

Intensity is proxied by the number of times two researchers co-authored a paper together.

Table 4 summarizes the main characteristics of the co-authorship networks at the macro-sector level. Networks are often composed of more than 500 authors (the number of nodes), who are linked by rare collaborations among them. The average number of co-authors of a scientist (i.e., degree centrality) is either zero or one, and the density of the network is in the order of  $10^{-3}$ .

The number of scientists involved in a collaboration is extremely small in each macro-sector; hence, it is plausible to expect that the circulation of knowledge and information is relatively limited. In all networks, 90% of the collaborations are activated by no more than 80 researchers, and at least 40% of the researchers are isolated. Moreover, most of the collaborations occur within small components, that is, a set of authors directly or indirectly connected among them. In fact, the density  $\times 1000$  is barely larger than 1 for most cases. The denser network is the macro-sector 02/D (physics teaching and history of physics). The giant component (largest component) in each network never comprises more than 6% of the total number of nodes. This suggests that the diffusion of information among connected scientists is likely to become rapidly redundant. Notably, however, the low propensity to collaboration is higher in Area 13 than in Area 2.

A similar picture emerges when considering the statistics relative to the networks of collaborations at the meso-level sectors, which are reported in Table 5. Density and degree centrality are extremely low in each meso-sector, and most of the authors

**Table 5** Areas 2 and 13 meso-sectors—descriptive table

Meso-sector	02/A1	02/A2	02/B1	02/B2	02/C1	02/D1	13/A1
Number of nodes	389	392	165	358	222	87	89
Number of edges	194	259	25	226	53	35	3
Density*1000	1.28	1.69	0.92	1.76	1.08	4.62	0.38
Nodes making up to 50% of collaborations	21	25	6	23	9	9	1
Nodes making up to 90% of collaborations	46	55	15	51	21	16	2
Degree—average	1	1.32	0.3	1.26	0.48	0.8	0.07
Degree—standard deviation	1.29	1.52	0.57	1.44	0.77	0.82	0.29
Giant component—size	12	10	5	10	6	5	3
Giant component—percentage	3%	3%	3%	3%	3%	6%	3%
Isolated authors—number	182	156	123	143	143	38	84
Isolated authors—percentage	47%	40%	75%	40%	64%	44%	94%
Clustering coefficient in giant component	0.56	0.58	0.5	0.58	0.56	0	0
Distance in giant component: average	4.5	3.76	2.1	3.76	2.4	3.2	1.33
Distance in giant component: standard deviation	2.44	2.04	0.97	2.04	1.1	1.82	0.52
Meso-sector	13/A2	13/A3	13/A4	13/A5	13/B1	13/B2	13/B3
Number of nodes	317	121	50	51	13	214	198
Number of edges	47	13	7	9	1	49	79
Density*1000	0.47	0.89	2.80	3.46	5.92	1.07	2.02
Nodes making up to 50% of collaborations	12	4	3	5	1	9	12
Nodes making up to 90% of collaborations	29	8	6	9	1	19	26
Degree—average	0.3	0.21	0.28	0.35	0.15	0.46	0.8
Degree—standard deviation	0.55	0.49	0.5	0.48	0.38	0.77	0.99
Giant component—size	4	3	3	2	2	9	6
Giant component—percentage	1%	2%	6%	4%	15%	4%	3%
Isolated authors—number	237	99	37	33	11	142	94
Isolated authors—percentage	75%	82%	74%	65%	85%	66%	47%
Clustering coefficient in giant component	0.6	1	0	0	0	0.19	0.75
Distance in giant component: average	1.33	5.33	2.67	1	1	2.67	2.2
Distance in giant component: standard deviation	0.49	0.52	1.37	0	0	1.03	1.3
Meso-sector	13/B4	13/B5	13/C1	13/D1	13/D2	13/D3	13/D4
Number of nodes	50	55	25	36	172	32	17
Number of edges	17	9	9	2	56	7	2
Density*1000	6.80	2.98	14.40	1.54	1.89	6.84	6.92
Nodes making up to 50% of collaborations	2	3	3	1	12	2	1
Nodes making up to 90% of collaborations	4	5	5	2	23	4	2
Degree—average	0.68	0.33	0.72	0.11	0.65	0.44	0.24
Degree—standard deviation	1.17	0.67	0.79	0.32	0.81	0.67	0.44
Giant component—size	6	3	3	2	5	5	2
Giant component—percentage	12%	5%	12%	6%	3%	16%	12%
Isolated authors—number	30	43	12	32	91	21	13
Isolated authors—percentage	60%	78%	48%	89%	53%	66%	76%
Clustering coefficient in giant component	0.75	1	0	0	0	0	0
Distance in giant component: average	1.93	1.33	1.33	2	2.6	2.4	1
Distance in giant component: standard deviation	0.78	0.52	0.52	0	1.39	1.31	0

*Note:* For each meso-sector network, indicated with its relative code (in columns), we report a number of descriptive metrics. For a precise definition of metrics refer to Sect. 3. The meso-sector relative to a code is indicated in Sect. 4

are either isolated or embedded in a very small number of components. Also, in this case, scientists from Area 13 show the smallest propensity to collaborate.

The comparison between the largest components of the networks of area 13 and area 2 shed some light on the different behaviors of the researchers in these two areas. At the macro level, scientists in Area 2 tend to share a higher number of collaborators with respect to the scientists in Area 13, as shown by their highest clustering coefficient. However, we observe the opposite when considering networks at the meso-level: scientists from Area 13 feature a higher number of co-authors with respect to their colleagues from Area 2. This suggests that the scientists from Area 2 are more inclined to activate collaborations across meso-sectors, while those from Area 13 are more prone to work with those belonging to their own meso-sector.

## 5 Network Analysis Results

In this section, we investigate the main features of the largest components of the networks, both at the macro- and at the meso-levels, to infer some relevant insights into the co-authorship behavior of the scientists considered in the current study.

### 5.1 Analysis at the Global Level

We begin by investigating the overall architecture of the networks' largest component, with the aim of finding evidence of specific model of interactions among scientists.

First, we find that scientists tend to form dense collaborations in the largest components, and many of them share one or more collaborators: that is, they feature a relatively high clustering coefficient. In Table 6, we compare this metric with that obtained from equivalent random (ER) networks, where macro-sector collaborations are formed by chance. We find that the clustering coefficient registered in the actual networks ( $C_{gC}$ ) is always higher than that observed in simulated networks ( $C_{rand}$ ). This is not surprising, and it is consistent with the fact that the scientists did not activate collaborations at random. On the contrary, scientists tend to choose a new collaborator among those already in contact with one of their co-authors, thus creating groups of collaborations presumably focused on a specific field of research, where skills are likely to be compatible.<sup>4</sup> The same behavior is observed when considering the networks at the meso- level, as reported in Table 7.

Second, we observe that the transmission of information among groups of scientists in the same component tends to be rather inefficient. This can be inferred

---

<sup>4</sup> It is worth noting that this behavior improves the chances to find new trusted collaborators, and it decreases screening costs.

**Table 6** Areas 2 and 13 macro-sectors—model checking

Macro-sector		02/A	02/B	02/C	02/D	13/B	13/C
Model behavior	Small world ( $C_{gC} \geq C_{rand}$ , and $L_{gC} \cong L_{rand}$ )	No	No	No	No	Yes	Yes
	$C_{gC} \geq C_{rand}$	Yes	Yes	Yes	Yes	Yes	Yes
	$L_{gC} \cong L_{rand}$	No	No	No	No	Yes	Yes
	$L_{gC} \leq L_{rand}$	Yes	No	No	Yes	No	No
Network local structure	True power law degree distribution	No	No	No	No	No	No
	Interlinked stars	No	No	Yes	No	No	No
	Core periphery	Yes	Yes	No	No	Yes	Yes

*Note:* For each macro-sector network, indicated with its relative code (in columns), we report a number of descriptive metrics. The terms  $C_{gC}$  and  $L_{gC}$  indicate the local clustering coefficient and the shortest path length of the observed network  $G$ , respectively. The terms  $C_{rand}$  and  $L_{rand}$  refer to the arithmetic mean of the average local clustering coefficient and the average shortest path of a number of simulated equivalent random networks, respectively. For a more detailed definition of metrics, refer to Sect. 3. The macro-sector relative to a code is indicated in Sect. 4

**Table 7** Areas 2 and 13 meso-sectors—model checking

Meso-sectors	02/A1	02/A2	02/B1	02/B2	02/C1	02/D1
Model behavior						
Small world ( $C_{gC} \geq C_{rand}$ , and $L_{gC} \cong L_{rand}$ )	No	No	Yes	No	Yes	No
$C_{gC} \geq C_{rand}$	Yes	Yes	Yes	Yes	Yes	Yes
$L_{gC} \cong L_{rand}$	No	No	Yes	No	Yes	No
$L_{gC} \leq L_{rand}$	Yes	No	No	No	No	No
True power law degree distribution	No	No	No	No	No	No
Interlinked stars	No	No	No	No	No	No
Core periphery	Yes	Yes	Yes	Yes	Yes	No
Meso-sector	13/A2	13/A3	13/B2	13/B3	13/B4	13/B5
Model behavior						
Small world ( $C_{gC} \geq C_{rand}$ , and $L_{gC} \cong L_{rand}$ )	Yes	No	Yes	No	Yes	Yes
$C_{gC} \geq C_{rand}$	Yes	No	Yes	Yes	Yes	Yes
$L_{gC} \cong L_{rand}$	Yes	No	Yes	No	Yes	Yes
$L_{gC} \leq L_{rand}$	No	No	No	No	No	No
True power law degree distribution	No	No	No	No	No	Yes
Interlinked stars	No	No	No	No	No	No
Core periphery	Yes	No	Yes	Yes	Yes	No
Meso-sector	13/C1	13/D2				
Model behavior						
Small world ( $C_{gC} \geq C_{rand}$ , and $L_{gC} \cong L_{rand}$ )	Yes	Yes				
$C_{gC} \geq C_{rand}$	Yes	Yes				
$L_{gC} \cong L_{rand}$	Yes	Yes				
$L_{gC} \leq L_{rand}$	No	No				
True power law degree distribution	No	No				
Interlinked stars	No	No				
Core periphery	No	No				
Network local structure						
Interlinked stars						
Core periphery						

*Note:* For each meso-sector network, indicated with its relative code (in columns), we report a number of descriptive metrics. The terms  $C_{gC}$  and  $L_{gC}$  indicate the local clustering coefficient and the shortest path length of the observed network  $G$ , respectively. The terms  $C_{rand}$  and  $L_{rand}$  refer to the arithmetic mean of the average local clustering coefficient and the average shortest path of a number of simulated equivalent random networks, respectively. For a more detailed definition of metrics, refer to Sect. 3. The meso-sector relative to a code is indicated in Sect. 4



by comparing the actual value of average path length ( $L_{gC}$ ) with that registered in ERs ( $L_{rand}$ ). In most cases, the distance among nodes which are linked by the relations formed at random in ERs, is lower than that among the nodes connected by actual relations. It follows that scientists tend to interact in small groups, being clumped into different and distant areas, even when embedded in the same largest component. However, this is not always the case. A more efficient configuration of the distance among nodes, similar to that observed in ERs, is observed in macro-sectors 13/B and 13/C (Table 6) and meso-sectors 02/B1, 02/C1, 13/A2, 13/B2, 13/B4, 13/B5, 13/C1, and 13/D2 (Table 7).

Taken together, our results point to the presence of a specific model of interaction for researchers in some sectors. When the clustering coefficient is higher than that registered in ERs, that is, some sort of specialized collaborations emerge among groups of scientists, and the distance between scientists is similar to that observed in ER, which means that the diffusion of information is relatively fast, we then find evidence for small-world behavior: a peculiar network structure with unique properties of local specialization and efficient information transfer. Such small-world behavior seems to be compatible with all networks found with a high value of average path length: that is, macro-sectors 13/B and 13/C and meso-sectors 02/B1, 02/C1, 13/A2, 13/B2, 13/B4, 13/B5, 13/C1, and 13/D2. The researchers affiliated with these sectors can rely on the fast and efficient exchange of information with their colleagues because of the network structure in which they are embedded. Overall, it seems that the researchers in Area 13 tend to interact more according to this mechanism with respect to researchers affiliated to Area 2.

Finally, we focus on the degree distribution of large components. In particular, we are interested in finding evidence in favor of or against a power law distribution with the parameter  $\alpha$  ranging between 2 and 3 (for more details, see Albert & Barabási 2002; Clauset et al. 2009). When this is the case, the authors follow a preferential attachment behavior, that is, scientists prefer to activate collaborations with those who already have many collaborations in place and who are pivotal in their sector. By looking at Tables 6 and 7, almost none of the network degree distributions fit with a true power law distribution in the macro-sectors and meso-sectors. The only exception is the network of those affiliated with meso-sector 13/B5, for which there is evidence of a preferential attachment behavior.

## 5.2 Analysis at the Local Level: Centrality Measures

We now turn to an analysis of the centrality measures (Freeman 1979) associated with the largest component of the networks. In particular, we focus on the relation between (i) degree centrality vs betweenness centrality and (ii) degree centrality vs the clustering coefficient (or local transitivity).

When the degree centrality is positively correlated with both betweenness centrality and the clustering coefficient, the network features a core-periphery structure where nodes located at the core of the subnetwork are densely connected

with one another (high degree centrality), acting as brokers (high betweenness) for the nodes situated at the periphery of the network. A core-periphery structure points to an uneven exposure to information among researchers: only those located at the core of the network can easily access to new information, while those located at the periphery tend to be excluded from the process of knowledge diffusion.

By contrast, when degree centrality is positively correlated with betweenness centrality and negatively correlated with the clustering coefficient, we say that the network has a structure similar to that of interlinked stars: few researchers play the role of a hub (high betweenness centrality) for others who are loosely connected with each other (low clustering coefficient). In this case, the diffusion of information becomes problematic. Most researchers will rely on a small number of colleagues to access knowledge produced in different areas of the networks. In other words, a small number of scientists act as information gatekeepers in these networks because the diffusion of knowledge heavily depends on the extent to which they are prone to receive new information from one part of the network and transmit it to different parts.

The analysis of the correlation between centrality measures is summarized in Tables 6 and 7. We observe that many of the largest components feature a core-periphery structure (02/A, 02/B, 13/B, 13/C) at the macro-level, and only macro-sector 02/C is characterized by relations arranged like interlinked stars. Moreover, no clear structure arises for the largest component of macro-area 02/D. As for meso-sectors, we detect the presence of a core-periphery structure in more than 60% of the giant components (meso-sectors 02/A1, 02/A2, 02/B1, 02/B2, 02/C1, 13/A2, 13/B2, 13/B3, and 13/B4).

### 5.3 *Network Attack*

Next, we test the resilience of the networks' architecture in macro-sectors by simulating different breakdown scenarios. Specifically, this is done by looking at global changes in the network topology after deleting 5% of the nodes. The results of our simulations for the macro-level networks are presented in Table 8.

The first row of the table indicates the number of components generated after deleting random nodes in the network. The third row reports the same statistics when attacking the topmost connected nodes (i.e., those with the highest degree centrality). The second and fourth rows report the ratio between the number of nodes in the giant component before and after the attack, respectively, when this is random or targeted. We see that by targeting random nodes, giant components remain substantially unaltered. By contrast, when an attack is targeted, the giant components of macro-sectors 02/A, 02/B, and 02/C lose 50% of the nodes. This is somewhat similar to what happens to macro-sector 13/B. This means that the topmost connected scientists in these networks are almost all embedded in the giant component, and they play a crucial role in sustaining the core of the collaborations in the macro-sectors. Even stronger is the effect in macro-sector 13/C, where the

**Table 8** Areas 2 and 13 macro-sectors—network attack

Macro-sector	02/A	02/B	02/C	02/D	13/B	13/C
Number of generated components—random	237	331	351	56	435	379
Giant component (after/before)—random	1.00	0.90	0.92	1.00	1.00	1.00
Number of generated components—target	266	353	368	59	463	407
Giant component (after/before)—target	0.50	0.50	0.50	1.00	0.60	0.29

*Note:* For each macro-sector network, indicated with its relative code (in columns), we report the number of components generated after deleting 5% of random nodes in the network (first row), and the ratio between the number of nodes in the giant component isolated after node deletion (second row), the number of components generated after deleting 5% of topmost connected nodes in the network (third row), and the ratio between the number of nodes in the giant component isolated after node deletion (fourth row). For a more detailed description of this procedure, refer to Sect. 5.3. The macro-sector relative to a code is indicated in Sect. 4

giant component loses about 70% of its members. Interestingly, there is no effect in macro-sector 02/D. This suggests that most collaborative scientists in this network are not embedded in the giant component; instead they work separately from the area where most researchers are involved.

The effect of random failures is less drastic than that produced by a targeted attack when evaluating the number of components generated by our simulations. The latter attacks consistently produce a higher number of components. This suggests that scientists benefiting from the diffusion of information channeled throughout network components heavily rely on the presence of the topmost connected authors.

The results remain substantially unchanged when testing the same effect at the meso-level, as reported in Table 9. Most meso-sectors rely on the topmost connected authors for the general connectivity of their networks.

Our findings hint to some policy indications: for example, replacing an eminent scientist collaborating with many laboratories (e.g., a node with high degree centrality) may compromise the chances of his/her colleagues rapidly finding in both macro- and meso-level networks new collaborators outside their research group or to access new information. In fact, by removing him/her from the network, his/her collaborators will remain isolated in small components with no direct or indirect connections to colleagues located in different zones of the network.

### 5.4 Co-authorship Networks and Scientific Performance: A Regression Analysis

In this section, we provide some insights into the relation between individual researchers’ network position and their average VQR scores as obtained by the evaluation of the papers they submitted for the evaluation exercise in 2011–2014. We carry out a linear regression analysis where the dependent variable is the

**Table 9** Areas 2 and 13 meso-sectors—network attack

Meso-sector	02/A1	02/A2	02/B1	02/B2	02/C1	02/D1	13/A2
Number of generated components—random	237	226	134	203	169	56	261
Giant component (after/before)—random	1	0.90	1	1	1	1	1
Number of generated components—target	266	236	141	220	181	59	275
Giant component (after/before)—target	0.50	0.80	0.40	0.50	0.50	1.00	0.50
Meso-sector	13/A3	13/B2	13/B3	13/B4	13/B5	13/C1	13/D2
Number of generated components—random	104	161	129	38	45	17	119
Giant component (after/before)—random	1	1	1	1	1	1	1
Number of generated components—target	109	177	140	39	48	18	124
Giant component (after/before)—target	0.67	0.33	0.50	0.50	0.67	1.00	0.80

*Note:* For each meso-sector network, indicated with its relative code (in columns), we report the number of components generated after deleting 5% of random nodes in the network (first row), and the ratio between the number of nodes in the giant component isolated after node deletion (second row), the number of components generated after deleting 5% of topmost connected nodes in the network (third row), and the ratio between the number of nodes in the giant component isolated after node deletion (fourth row). For a more detailed description of this procedure, refer to Sect. 5.3. The meso-sector relative to a code is indicated in Sect. 4

**Table 10** Distribution of researchers according to scientific area

Area	Frequency
13 (Economics and statistics)	1268
02 (Physics)	892

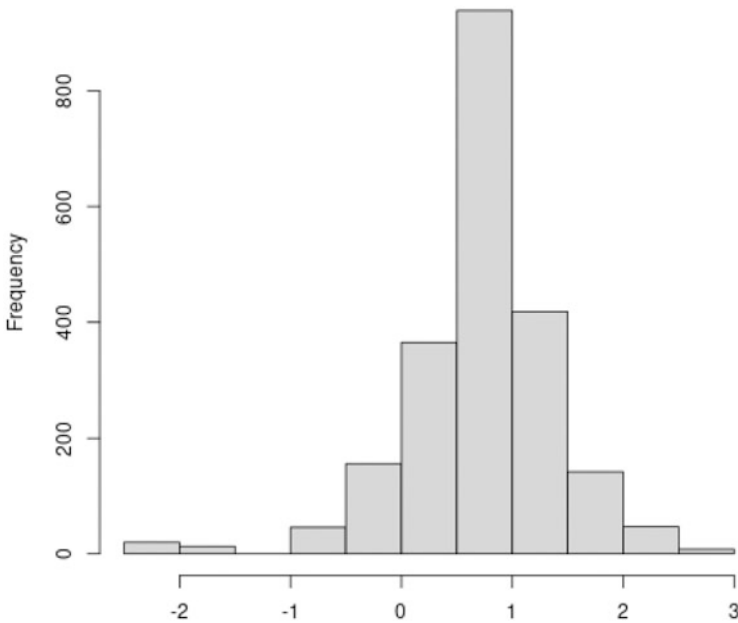
**Table 11** Distribution of researchers according to their gender

Gender	Frequency
Female (F)	513
Male (M)	1647

“excellence” measure (VQR scores) of researchers active in physics or economics and covariates represented by some of the available individual characteristics and individual network indices. In particular, as individual characteristics, we used the scientific areas (physics and economics and statistics, the latter treated as a reference category), gender (female as the reference category), and geographic location of the university to which researchers were affiliated (S, Southern Italy, I, Islands, NE, North-Eastern Italy, NO, North-Western Italy, and C, Central Italy, this latter considered as reference category). The distribution of such covariates is reported in Tables 10, 11, and 12, respectively. As far as the individual network indices are concerned, we adopted the centrality measures defined in Sect. 3, namely node degree, betweenness, and transitivity (i.e., clustering coefficient). The variable excellence of the authors is the average VQR score of the authors’ papers. As already stated, the VQR score is a 5-point scale with the scores reported in Table 2 and described in Sect. 4. In this analysis, we used the scores in Table 2 to compute the average VQR evaluation for each author. In Fig. 1, we depict the distribution of

**Table 12** Distribution of researchers according to geographic location of their hiring university

Geographic location	Frequency
Central Italy (C)	520
Islands (I)	159
North-Eastern Italy (NE)	472
North-Western Italy (NO)	568
Southern Italy (S)	441



**Fig. 1** Histogram of log(excellence). 19 researchers with an overall excellence value equal to 0 are excluded

the log transform of the excellence variable that we used as dependent variable in the regression model.

One of the distinct properties of several network characteristics measured on the level of researchers is their asymmetric distribution. A specific feature of the analyzed network is its high level of fragmentation with a large number of small components and isolates, hence preventing the calculation of network statistics for some of the units.

To meet the assumptions of the regression analysis, the network-based variables included in the model were categorized. Transitivity and betweenness were dichotomized into categories indicating zero and nonzero values; degree (number of connections) was categorized into three categories (0, 1–10, and 11–66) indicating degree centrality of the researchers.

### Model and Interpretation

Model results are reported in Table 13. It can be noted that gender differences have no significant effect on scientific performance; this means that after controlling for the geographical area and network characteristics of the researchers, the gender gap is not present in the analyzed scientific areas, differently from other studies (Aksnes et al. 2019). Moving to the geographic location, the model assumes universities located in Central Italy a reference category (baseline). The category indicating researchers working in universities in North-Eastern Italy has negative and significant effect, so their performance is significantly lower than Central Italy. No significant differences are found among Central, Islands, South, and North-Western Italy.

As can be noted from the *Physics* coefficient value, the authors in the physics area have higher performance than the authors in economics and statistics area (baseline category of *Physics*).

For the network indices of co-authorship network, the authors are more likely to achieve a higher VQR score if they have a greater degree and betweenness higher than 0. The same holds for transitivity as an indicator of working in clustered research groups. This shows that working with several co-authors (high degree) and being part of multiple clustered research groups (high betweenness) matter in terms of successful research.

The differences between disciplines become even higher when the authors have between 1 and 10 co-authors (estimated parameter of the interaction effect between *Physics* and Degree[1,10]), which means that highly central researchers are more successful in a bibliometric discipline, likewise physics macro-sectors. Above this

**Table 13** Regression model estimates and fit

	Dependent variable		<i>t</i> -value	Pr(>   <i>t</i>  )
	log (Excellence)			
	Estimate	Std. Error		
(Intercept)	0.456 * **	0.0378	12.06	0.0000
Gender (male)	-0.004	0.0298	-0.14	0.8852
Geographic area (Islands)	0.004	0.0522	0.07	0.9454
Geographic area (North-East)	-0.101 * **	0.0368	-2.74	0.0062
Geographic area (North-West)	-0.037	0.0352	-1.06	0.2901
Geographic area (South)	0.009	0.0375	0.25	0.8030
Physics	0.114 * **	0.0358	3.19	0.0015
Transitivity	0.101 * *	0.0455	2.23	0.0259
Degree[1,10]	0.292 * **	0.0349	8.35	0.0000
Degree[10,66]	0.830 * **	0.1895	4.38	0.0000
Betweenness	0.289 * **	0.0531	5.44	0.0000
Physics:degree[1,10]	0.110 * *	0.0535	2.06	0.0396
Physics:degree[10,66]	0.317	0.1946	1.63	0.1035

Note: \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$

threshold (so for very high degree values), the effect of the number of co-authors on performance is positive on average, independently on the scientific area. This latter effect is in line with previous findings (Abbasi et al. 2011; De Stefano & Zaccarin 2016; Lee & Bozeman 2005) and shows that having high number of different co-authors can lead to positive effects on scholars' scientific performance in different fields.

## 6 Concluding Remarks

The present chapter illustrated the use of SNA tools for co-authorship in the context of the Italian research evaluation exercise that ran from 2011 to 2014. In particular, we analyzed the results at the different network levels (global, subgroup, and individual actor levels), here considering their relations with the performance of researchers. The analysis is in line with the literature on scientific collaboration. In fact, research collaboration is often reported as a driver of scientific quality and productivity (Abbasi et al. 2011). For this reason, the analysis of collaboration and co-authorship networks provides essential information for the design of many academic policies. An in-depth understanding of the interactions among scientists provides useful insights into the conditions underlying creativity and genesis of scientific discovery, and it may provide information on new tools and policies that have the potential to accelerate science (Fortunato et al. 2010). This is particularly relevant when considering the interdisciplinary fields required to tackle complex problems in innovative ways and bridge disciplinary silos, such as the fight against climate change or the current COVID-19 crisis. The study of collaboration networks can also be leveraged to provide scientists with access to new and non-redundant information allowing them to engage in more innovative studies. For this reason, the researchers have been progressively stimulated by new policies to activate new forms of collaboration and improve their position in the co-authorship network. For instance, this is the case of scientific policies providing research funding conditional on the activation of a new intellectual collaboration or the case of internal department tenure policies that require candidates to have a minimum amount of publications but that do not fully discount articles by the number of authors (see Ductor 2015 for a recent discussion). Knowledge of one's collaboration network is also an essential tool to forecast one's future research output and productivity (Ductor et al. 2014); therefore, it provides crucial information for conducting good recruitment in a department and hiring talented researchers. Moreover, the structure of scientific collaboration networks is a powerful source of information on the dependence of a research team from the presence of so-called academic stars (Azoulay et al. 2010; Waldinger 2010). Therefore, this finding provides useful suggestions to design a system of incentives for "superstar" scholars to (i) remain in the university and maintain an efficient network of collaborations and (ii) increase the involvement of their collaborators in research projects, to reduce the dependency of the overall network from their own work. Finally, collaboration networks are

important predictors of the level of peer pressure suffered by an individual; this can be altered to improve a scientist's working environment and correct undesired situations, such as the presence of gender or other kinds of disparities (Lindenlaub & Prummer 2021). The results presented in the work, despite being retrieved from a small sample of publications of scholars in specific areas, suggest that even in the Italian scenario, it would be worth fostering intra and interdisciplinary collaboration to improve group and individual scientific productivity and performance. This is especially true based on the insights on the importance of a network position in producing quality research outputs. To perform this task and introducing new policies in this direction, comprehensive knowledge of the network structure in disciplines is crucial. The understanding of network patterns by means of the tools presented can guide in the detection of those researchers in certain structural position who may be the target of some network-based interventions (e.g., scale-free networks because the one observed on meso-sector 13/B5 relies on few important nodes acting as hubs). We believe that the results are promising, but we think that a future analysis would benefit from the availability of richer datasets containing a larger set of individual publication records for retrieving a more comprehensive co-authorship network.

## References

- Abbasi, A., Altmann J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5, 594–607.
- Abbasi, A., Chung, K. S. K., & Hossain, L. (2012). Egocentric analysis of co-authorship network structure, position and performance. *Information Processing and Management*, 48, 671–679.
- Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794), 378
- Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74, 47–97.
- Abel, G. J., Muttarak, R., Bordone, V., & Zaghen, E. (2019). Bowling together: Scientific collaboration networks of demographers at European population conferences. *European Journal of Population*, 35, 543–56.
- Abramo, G., D'Angelo, C. A., & Di Costa, F. (2018). The collaboration behavior of top scientists. *Scientometrics*, 118(1), 215–232.
- Abramo, G., D'Angelo, C. A., & Di Costa, F. (2019). A gender analysis of top scientists' collaboration behavior: Evidence from Italy. *Scientometrics*, 120, 405–418.
- Abramo, G., D'Angelo, C. A., & Murgia, G. (2014). Variation in research collaboration patterns across academic ranks. *Scientometrics*, 98(3), 2275–2294.
- Aksnes, D. W., Piro, F. N., & Rørstad, K. (2019). Gender gaps in international research collaboration: A bibliometric approach. *Scientometrics*, 120, 747–774.
- Azoulay, P., Zivin, J. G., & Wang, J. (2010). Superstar extinction. *Quarterly Journal of Economics* 25, 549–589.
- Beaver, D. (2001). DeB.: Reflections on scientific collaboration (and its study): Past, present, and future. *Scientometrics*, 52, 365–377.
- Bellotti, E. (2012). Getting funded. Multi-level network of physicists in Italy. *Social Networks* 34, 215–229.



- Bellotti, E., Kronegger, L., & Guadalupi, L. (2016). The evolution of research collaboration within and across disciplines in Italian Academia. *Scientometrics*, *109*, 783–811.
- Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703.
- Commission of European Community. (2006). Amended proposal for Council Decision concerning the Specific Programme “Cooperation” implementing the Seventh framework programme (2007–2013) of the European Community for research, technological development and demonstration activities.
- Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy*, *38*, 293–305.
- De Stefano, D., Fuccella, V., Vitale, M. P., & Zaccarin, S. (2013). The use of different data sources in the analysis of co-authorship networks and scientific performance. *Social Networks* *35*, 370–381.
- De Stefano, D., & Zaccarin, S. (2016). Co-authorship networks and scientific performance: An empirical analysis using the generalized extreme value distribution. *Journal of Applied Statistics*, *43*, 262–279.
- Digiampietri, L., Rego, L., Costa de Souza F., Ospina, R., & Mena-Chalco, J. (2017). Brazilian network of PhDs working with probability and statistics. *Brazilian Journal of Probability and Statistics* *32*, 755–782.
- Ferligoj, A., Kronegger, L., Mali, F., Snijders, T. A. B., & Doreian, P. (2015). Scientific collaboration dynamics in a national scientific system. *Scientometrics* *104*, 985–1012.
- Fischbach, K., Putzke, J., & Schoder, D. (2011). Co-authorship networks in electronic markets research. *Electron Markets*, *21*, 19–40.
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., & Vespignani, A. (2010). Science of science. *Science*, *359*(6379).
- Freeman, L. C. (1979). *Centrality in networks: I. Conceptual clarification*. Social networks (Vol. 1, pp. 215–239).
- Fronczak, A., Fronczak, P., & Hoyst, J. A. (2003). Mean-field theory for clustering coefficients in Barabási-Albert networks. *Physical Review E*, *68*(4), 046126.
- Fuccella, V., De Stefano, D., Vitale, M. P., & Zaccarin, S. (2016). Improving co-authorship network structures by combining multiple data sources: Evidence from Italian academic statisticians. *Scientometrics* *107*, 167–184.
- Ductor, L., Fafchamps, M., Goyal, S., & Van der Leij, M. J. (2014). Social networks and research output. *Review of Economics and Statistics*, *96*(5), 936–948.
- Ductor, L. (2015). Does co-authorship lead to higher academic productivity? *Oxford Bulletin of Economics and Statistics*, *77*(3), 385–407.
- Goyal, S., Van der Leij, M. J., & Moraga-Gonzalez, J. L. (2006). Economics: An emerging small world. *Journal of Political Economy*, *114*, 403–412.
- Kronegger, L., Mali, F., Ferligoj, A., & Doreian, P. (2012). Collaboration structures in Slovenian scientific communities. *Scientometrics*, *90*, 631–647.
- Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, *35*, 673–702.
- Leone Scialolazza, V., Vacca, R., Kennelly Okraku, T., & McCarty, C. (2017). Detecting and analyzing research communities in longitudinal scientific networks. *PLoS One*, *12*, e0182516.
- Leone Scialolazza, V., Vacca, R., & McCarty, C. (2020). Connecting the dots: Implementing and evaluating a network intervention to foster scientific collaboration and productivity. *Social Networks*, *61*, 181–195.
- Lindenlaub, I., & Prummer, A. (2021). Network structure and performance. *The Economic Journal*, *131*(634), 851–898.
- Maggioni, M. A., & Uberti, T. E. (2011). Networks and geography in the economics of knowledge flows. *Quality & Quantity*, *45*, 1031–1051.

- Mali, F., Kronegger, L., Doreian, P., & Ferligoj, A. (2012). Dynamic scientific co-authorship networks. In: A. Scharnhorst, K. Börner, & P. van den Besselaar, P. (Eds.), *Models of science dynamics* (pp. 195–232). Berlin, Heidelberg: Springer.
- Melin, G. (2000). Pragmatism and self-organization: Research collaboration on the individual level. *Research Policy*, 29, 31–40.
- Ministry of Education and Science. (2006). Spanish National Plan of Research and Development 2008–2011. <http://www.ingenio2010.es/>
- Moody, J. (2004). The structure of a social Science: Disciplinary cohesion from 1963 to 1999. *American Sociological Review*, 69, 213–238.
- National Institute of Health. (2007). NIH launches interdisciplinary research consortia. <https://www.nih.gov/news-events/news-releases/nih-launchesinterdisciplinary-research-consortia>
- Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, 101, 5200–5205.
- Ponomarev, B., & Boardman, C. (2016). What is co-authorship? *Scientometrics*, 109, 1939–1963.
- Sà, C. M. (2008). Interdisciplinary strategies' in U.S. research universities. *Higher Education*, 55, 537–552.
- Uddin, S., Hossain, L., & Rasmussen, K. (2013). Network effects on scientific collaborations. *PLoS ONE* 8(2), e57546.
- Waldinger, F. (2010). Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany. *Journal of Political Economy*, 118, 787–831.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.
- Watts, D. (1999). Networks, dynamics and the small-world phenomenon. *American Journal of Sociology*, 105, 493–527.
- Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316, 1036–1039.
- Yan, E., & Guns, R. (2014). Predicting and recommending collaborations: An author-, institution-, and country-level analysis. *Journal of Informetrics*, 8, 295–309.

**Domenico De Stefano** (Ph.D. University of Naples Federico II) is Professor of Social Statistics at the University of Trieste. His research interests focus on social network analysis methods and applications in knowledge diffusion networks, multivariate data analysis and statistical modelling.

**Luka Kronegger** (Ph.D. University of Ljubljana) is Assistant Professor at the University of Ljubljana. His main research is focused on knowledge transition in the higher education mentorship networks and investigation of European stakeholders involved in the field of childhood obesity.

**Valerio Leone Sciabolazza** (Ph.D. Sapienza University of Rome) is Assistant Professor at Sapienza University of Rome. His research interests are on network diffusion processes, economic growth and migration, the role of political patronage networks in determining legislators' activities.

**Maria Prosperina Vitale** (Ph.D. University "G. D'Annunzio" Chieti-Pescara) is Associate Professor of Social Statistics at the University of Salerno. Her current research interests are in the field of network analysis with a focus on student mobility choices and co-authorship relationships.

**Susanna Zaccarin** is Professor of Social Statistics at the University of Trieste. Her research interests focus on data collection methods, statistical modelling and network analysis.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.



# Topic-Driven Detection and Analysis of Scholarly Data



**Alfio Ferrara, Corinna Ghirelli, Stefano Montanelli, Eugenio Petrovich, Silvia Salini, and Stefano Verzillo**

**Abstract** The chapter presents a topic mining approach that can be used for a scholarly data analysis. The idea here is that research topics can emerge through an analysis of epistemological aspects of scholar publications that are extracted from conventional publication metadata, such as the title, the author-assigned keywords, and the abstract. As a first contribution, we provide a conceptual analysis of research topic profiling according to the peculiar behaviours/trends of a given topic along a considered time interval. As a further contribution, we define a disciplined approach and the related techniques for topic mining based on the use of publication metadata and natural language processing (NLP) tools. The approach can be employed within

---

The information and views set out in this work are those of the authors and do not reflect the official opinion of the European Union, the Bank of Spain, or the Eurosystem. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the any use that may be made of the information contained therein.

---

A. Ferrara · S. Montanelli (✉)

Department of Computer Science, Data Science Research Center, Università degli Studi di Milano, Milan, Italy  
e-mail: [alfio.ferrara@unimi.it](mailto:alfio.ferrara@unimi.it); [stefano.montanelli@unimi.it](mailto:stefano.montanelli@unimi.it)

C. Ghirelli

Banco de Espana, DG Economics, Statistics and Research, Madrid, Spain  
e-mail: [corinna.ghirelli@bde.es](mailto:corinna.ghirelli@bde.es)

E. Petrovich

Department of Economics and Statistics, Università degli Studi di Siena, Siena, Italy  
e-mail: [eugenio.petrovich@unisi.it](mailto:eugenio.petrovich@unisi.it)

S. Salini

Department of Economics, Management and Quantitative Methods, Data Science Research Center, Università degli Studi di Milano, Milan, Italy  
e-mail: [silvia.salini@unimi.it](mailto:silvia.salini@unimi.it)

S. Verzillo

European Commission, Joint Research Centre (JRC), Ispra, Italy  
e-mail: [stefano.verzillo@ec.europa.eu](mailto:stefano.verzillo@ec.europa.eu)

© The Author(s) 2022

D. Checchi et al. (eds.), *Teaching, Research and Academic Careers*,  
[https://doi.org/10.1007/978-3-031-07438-7\\_8](https://doi.org/10.1007/978-3-031-07438-7_8)

a variety of topic analysis issues, such as country-oriented and/or field-oriented research analysis tasks that are based on scholarly publications. In this direction, to assess the applicability of the proposed techniques for use in a real scenario, a case study analysis based on two publication datasets (one national and one worldwide) is presented.

**Keywords** Natural Language Processing · Scholarly Data Analysis · Topic Mining

## 1 Introduction

In contemporary science policy debates, one of the most heated discussions concerns the role and effects of research performance metrics in research assessment frameworks. According to the advocates of using these metrics, indicators based on citations and publications would be more objective than the traditional peer review system, hence allowing for breaking ‘old boys circles’ and hampering nepotism, cronyism, and other inappropriate academic practices (Geuna & Martin, 2003). Moreover, by setting measurable thresholds and benchmarks, performance metrics would stimulate both the quantity and quality of scientific production (Bonaccorsi, 2015; Geuna & Martin, 2003; Moed, 2017). Finally, research evaluation based on metrics would be less expensive than peer review, it would save taxpayers’ money (Geuna & Piolatto, 2016), and as recent evidence shows, it may provide comparable results, at least when the need is to assess research performance at the institutional level (Checchi et al., 2021). In addition, at the individual level, the predictive power of bibliometrics is superior to peer review in almost all disciplines (except medicine). On the other hand, critics insist on the ‘unintended consequences’ of using metrics and on the ‘constitutive effects’ that their pervasive presence has on the behaviour of researchers (Dahler-Larsen, 2014). These effects include goal displacement (scoring high on the metrics becoming a target in and of itself), promotion of the unethical use of citations (excessive self-citation, creation of citation cartels, the strategic exchange of citations, etc.), task reduction (academic activities that are not considered in the calculation of the indicators, such as teaching and public engagement, being avoided), and an artificial increase in productivity by ‘salami slicing’ (dividing one scientific work into multiple publications) (Fochler et al., 2016; de Rijcke et al., 2015). Even the recent rise in retractions and research misconduct (e.g. fabrication of results, plagiarism, ‘p-hacking’, etc.) has been linked to the increasing pressure of metrics (Biagioli et al., 2019).

One of the most interesting criticisms raised against metrics in research evaluations is that they would not only affect the behaviour of researchers, but also the epistemic content of the science they produce (i.e. ideas, research themes, methods, etc.). In particular, the excessive weight of metrics would damage the pluralism of scientific enquiry, rewarding the mainstream approaches not because of their scientific merit, but only because of their (transient) popularity or connection

with academic power. For instance, in a recent joint declaration, the Académie des Sciences, Leopoldina, and the Royal Society (Académie des Sciences et al., 2017) write that ‘undue emphasis on bibliometric indicators [...] may also hinder the appreciation of the work of excellent scientists outside the mainstream; it will also tend to promote those who follow current or fashionable research trends, rather than those whose work is highly novel and which might produce completely new directions of scientific research’ (p. 2). Metrics have been blamed for inducing risk avoidance in science: the researchers, under the pressure of scoring well on indicators, would focus on topics, research programmes, and methods that are more likely to be rewarded. By contrast, they would avoid revolutionary ideas, out-of-the-box innovation, and interdisciplinarity because these would be deemed to be too risky enterprises. Thus, orthodoxy and conformism would be promoted at the expense of critical thinking, damaging scientific progress.

Until recently, the presence and magnitude of the effects of metrics on scientists and science have been debated more often than empirically investigated. In recent years, the empirical study of the effects of research evaluations on research practices has begun. An increasing body of literature has started documenting how researchers react under competitive conditions, which may affect their likelihood of promotion, particularly how research evaluation frameworks based on ‘metrics’ have induced a change in the publishing behaviour of researchers. In Italy, for instance, recent evidence shows that the introduction of research evaluation procedures has promoted strategic behaviours among researchers via the creation of ‘citation clubs’ that are aimed at artificially inflating bibliometric outcomes (Baccini et al., 2019; Scarpa et al., 2018; Seeber et al., 2019).

However, studying the impact of metrics and the evaluation on the epistemic content of research, that is, on the theories and ideas that are produced by the scientific community under the regime of metrics-based evaluation, is still in its infancy (Muller & de Rijcke, 2017). In particular, the accuracy of *the mainstream criticism* outlined above is still to be addressed by empirical studies.

In this chapter, we propose a mining approach for the detection and analysis of ‘mainstream topics’. The proposed idea is that topics featuring mainstream research can emerge through an analysis of the epistemological aspects of scholarly publications extracted from conventional publication metadata, such as the title, the author-assigned keywords, and the abstract. As a first contribution, we provide a conceptual analysis of the notion of mainstream research that is exploited to enforce mainstream profiling based on peculiar behaviours/trends of research topics along a considered time interval. As a further contribution, we define a disciplined approach and the related techniques for topic mining based on the use of publication metadata and natural language processing (NLP) tools. Finally, a case study analysis is presented to assess i) the applicability of the proposed techniques to a real scenario based on a publication dataset of Italian scholars and ii) the scalability and reliability of some of the case study results when the proposed approach is based on a richer and comprehensive database of all international publications, as collected by Scopus Elsevier over 14 years.

The chapter is organised as follows: In Sect. 2, studies on the epistemic impacts of metrics and techniques for automatic topic extraction from scholarly publications are briefly presented. In Sect. 3, a conceptual analysis of mainstream and what it comprises is provided by highlighting the different and sometimes opposite meanings that are attached to this concept in the literature. In Sect. 4, we present our modelling considerations about mainstream profiling. The proposed approach and techniques to topic mining are illustrated in Sect. 5. In Sect. 6, the results obtained by applying the proposed techniques to both a real publication dataset of Italian scholars and to the whole Scopus publication set on the same disciplines are discussed. Finally, our concluding remarks are given in Sect. 7.

## 2 Literature Review

By the *epistemic impacts* of metrics, we are focusing on the array of changes induced by metrics-based evaluation regimes on the epistemic processes of knowledge production and their outputs (scientific ideas, theories, research programmes, etc.) (Muller & de Rijcke, 2017). Epistemic impacts should be analytically distinguished from the effects of metrics on the social structure of science and, specifically, on its reward system, even if, in concrete situations, both kinds of impacts are likely to occur together. The decline of interdisciplinary research and reduction of scientific pluralism are examples of the epistemic impacts of metrics. By contrast, the rise of self-citations and gift authorships are examples of changes in the reward system of science.

The reward system-related effects are relatively easier to capture using quantitative methods because they can be inferred from analysing publications and citations. Starting from the pivotal study of Butler (2003) on the effects of the Australian research evaluation system, most scientometric studies so far have focused on these quantitative indicators to investigate the changes in researcher behaviour under the pressure of metrics (Abramo et al., 2019; Abramo et al., 2021; Baccini et al., 2019). Epistemic impacts, on the other hand, are more difficult to track for three main reasons. First, epistemic concepts, such as interdisciplinarity, scientific pluralism, and scientific mainstream, do not have standard, uncontested definitions. Second, the quantitative operationalisations of these notions frequently run the risk of reducing complex phenomena to monodimensional measures that miss important epistemological nuances. Third, there is no consensus on what epistemic factors contribute the most to scientific progress. For instance, philosophers of science have long debated what degree and what kind of scientific pluralism is beneficial to scientific enquiry (see (Viola, 2018) for a detailed discussion of the literature on this topic). The epistemic deviations induced by metrics are difficult to point out because there is no universally accepted baseline normative epistemology that accounts for the correct functioning of science.

In light of these methodological and theoretical impasses, most of the research on the epistemic impacts of metrics so far has turned to methodologies, such as

surveys and interviews, showing how researchers themselves perceive the pressure of metrics on their epistemic practices. In one of the first studies of this kind, Muller and de Rijcke (2017) interviews 38 Dutch and Austrian post-docs and junior group leaders in the life sciences, finding that researchers pervasively ‘think with indicators’. Indicators such as the journal impact factor do not intervene only in the evaluation of research after the fact, but also inform the entire research process, from the very conception of research projects to the choice of scientific collaborators and even the animal models used. Castellani et al. (2016) reach a similar conclusion after giving out a questionnaire to 12 Italian scientists from several disciplines. Their interviewees underline the risk that metrics in research evaluation can promote uniformity in the scientific community and discourage ground-breaking approaches. Also, the interviewees argue that metrics worsen the ‘publish or perish’ culture and induce scientists to publish low-quality material just to score better on productivity indicators. Feenstra and Lopez-Cozar (2021) interview 14 Spanish researchers in philosophy and ethics about the effects of metrics in their disciplines. Even though the interviewed researchers identify some positive effects, such as more transparent policies in the academic promotion process, they deem the impact on research agendas, publication language, and mental health as negative. In particular, metrics would hamper intellectual diversity in philosophical research and even lead to research misconduct. These studies highlight how metrics and indicators have gained a prominent place in the ‘epistemic living space’ of researchers, both in the natural and social sciences (Felt, 2009).

One limitation of these studies, however, is that they do not discuss the epistemic concepts used by the researchers to frame their experience of metrics. In this sense, then, they offer a valuable but partial perspective on epistemic impacts. By contrast, the present study is the first attempt to ground an investigation of an epistemic phenomenon, that is, the scientific mainstream, in a conceptual analysis of the related epistemic concept.

Further related work focuses on the methods and techniques for the classification of scholarly publications. Usually, a combination of automated procedures and manual activities/practices has been proposed (Glenisson et al., 2005). Solutions based on the use of human-assigned metadata, such as superimposed subject categories of articles and journals, represent a popular solution (Borner, 2010). This approach is effective when the choice of subject categories is shared by the final users and the classification results provide a scholarly picture in which the actors (i.e. the publication authors) can self-recognise the categorisation of their scientific products. However, manually defined subject categories are characterised by several well-known weaknesses. For instance, predefined categories are typically inadequate for dealing with publications about emerging topics characterised by recent formation and a new epistemic body (Suominen & Toivanen, 2016). Machine learning and unsupervised classification/clustering approaches have recently been proposed for overcoming such limitations. For instance, in Boyack et al. (2011) and Talley et al. (2011), topic modelling and clustering solutions are exploited to provide a visual, graph-based representation of a publication dataset extracted from the MEDLINE repository and the National Institutes of Health (NIH), respectively.



Similar approaches have been investigated (Nichols, 2014; Yan et al., 2012) for the information retrieval field and for the National Science Foundation awards, respectively. On the other hand, the construction of a map of science merely derived from scholarly data by using automated classification algorithms is characterised by possible limitations, as well. For instance, automated solutions are generally weak in capturing the minor trends within a discipline, even if they provide a relevant contribution from the historical and epistemic point of view. A recent comparison between unsupervised learning and human-assigned approaches to classification of scholarly data has been provided (Suominen & Toivanen, 2016); in the study, a topic modelling solution based on the latent Dirichlet allocation (LDA) algorithm is exploited. The results show that it is difficult to argue the superiority of one method (human-based scholarly data classification) over the other (algorithm-based scholarly data classification) (Suominen & Toivanen, 2016). However, it is well recognised that machine-generated scholarly data classifications provide a strong contribution in terms of practicality (Castano et al., 2018). This means that the capability to rapidly generate thematic, interactive views of an underlying (large) scholarly publication dataset can be considered as a result, but it also represents a worth support/contribution for experts that aim to further refine/revise the obtained results to provide their own data views.

### 3 Conceptual Analysis of the Mainstream Notion

Etymologically, the term ‘mainstream’ refers to the main current of a river or a stream. According to the dictionary, the mainstream is the ‘prevailing current of thought, influence or activity’. As an adjective, ‘mainstream’ means ‘representing the prevalent attitudes, values, and practices of a society or a group<sup>1</sup>’. The term usually belongs to the context of artistic and cultural phenomena, where it is mainly used to denote trends in popular and media culture.<sup>2</sup> Sometimes, it takes a pejorative sense by subcultures who view the mainstream culture as artistically inferior.

When it is employed in a discussion about science, ‘mainstream’ preserves its nature as a common language term. However, a precise and widely accepted definition of what ‘mainstream’ means in reference to science is missing, as is an operational definition of how to measure it.

The term can be used as a noun (‘the mainstream in economics’), as well as an adjective (‘mainstream science’). In both cases, mainstream is said of many different aspects of the scientific enquiry, from the most abstract to the most practical. Mainstream can be the following:

---

<sup>1</sup> American Heritage Dictionary of the English Language, Fifth Edition (2011).

<sup>2</sup> For an overview, see <https://en.wikipedia.org/wiki/Mainstream>

- A general theoretical framework or research programme (e.g. the neoclassical approach in economics)
- A specific theory (e.g. the big bang theory in cosmology)
- A position in a debate (e.g. the functionalism in the philosophy of mind)
- A research object or topic (e.g. the quark bottom in high energy physics)
- A research methodology (e.g. participant observation in cultural anthropology)
- A technique, a research protocol, or a procedure (e.g. the PCR in molecular biology).

In the empirical study of what is mainstream, such variability must be considered to set an appropriate level for the analysis. Different empirical methods will capture the mainstream at different levels of ‘granularity’, depending on the scientific aspect being considered. However, the most important feature about the term and its usage is that it assumes *different and sometimes opposite meanings* in the literature. ‘Mainstream’ is used to reference not only different things, but also different and sometimes incompatible ways. By surveying the literature, we can analytically distinguish six key meanings, whose differences can be appreciated better when they are compared with their opposites.<sup>3</sup>

1. **‘Orthodox’ vs. ‘heterodox’ or ‘fringe’.** In this sense, the mainstream is the dominant school of thought within a certain discipline or field. The mainstream is characterised by adherence to certain scientific content (e.g., a theory, a research programme, a method, etc.). The nonmainstream schools, on the other hand, are characterised by their refusal of some of the mainstream’s tenets. In this sense, the term is used both positively and negatively. In the positive sense, the mainstream represents the standard view of the scientific community, whereas heterodox schools represent the margins of science, dangerously bordering pseudo-science (‘fringe science’) (Gottfredson, 1997). By contrast, when the term is used negatively, the closed-mindedness or refusal of the ‘pluralism’ of the mainstream is stressed (Colander et al., 2004). An instance of this meaning can be found in economics, where heterodox schools (e.g., Marxists, post-Keynesians, feminists, Old Institutionalists, and Austrians) are distinguished from the neoclassical mainstream. In Anglo-American philosophy, analytic philosophy may be considered the mainstream, whereas Continental philosophy can be thought of as the heterodox approach (Katzav & Vaesen, 2017).
2. **‘Normal’ vs. ‘Revolutionary’, ‘Ground-breaking’.** The second meaning of mainstream refers to the distinction between normal and revolutionary science (Kuhn, 1996). Mainstream science would be characterised by a step-by-step, cumulative nature, whereas nonmainstream science would be more revolutionary, ground-breaking, and frequently not understood by the mainstream because of this. Mavericks and misunderstood geniuses would be the typical makers of

---

<sup>3</sup> Note that the six meanings rarely appear in their pure form. Often, scholars and commentators mix two or more meanings together. The six meanings should be considered as ideal types for the analysis, not as simple descriptions of usage.

nonmainstream science (Heinze, 2013). Compared with the first meaning, the focus is on the mode of scientific progress rather than on the adherence to some specific theory.

3. **‘Popular’ vs. ‘Niche’**. The third meaning is neutral with respect to the scientific content of mainstream and nonmainstream science. Here, mainstream is purely characterised in quantitative terms as the research that is currently done by most of the researchers in a field. No judgement on the orthodoxy or progress of the mainstream is implied. Nonmainstream science is not considered as heterodox or ground-breaking, just as topics that are addressed by fewer researchers. This meaning of mainstream can be compared with the concept of ‘impact’ in a citation analysis. When the number of citations received by a publication is not considered a proxy of its scientific quality (a normative concept), it can nonetheless be considered a measure of popularity or influence (a descriptive concept).
4. **‘Trendy’, ‘Short-lived’, ‘Passing’ vs. ‘Stable’, ‘Long-Lasting’**. In this sense, the focus is on the temporal extension of mainstream science. The mainstream is equated with the currently ‘hot’ areas of research. However, the short life of these areas is also implied. From this point of view, a mainstream researcher is a researcher who follows the trends, doing what everybody does at that moment. Mainstream topics are the ones that are currently fashionable in the research community, for whatever reason. These mainstream topics have the highest chance of being published in the highest-ranked journals and produce a high impact in terms of citations. In the literature on the perverse effects of research metrics, mainstream is frequently intended in this sense (e.g. de Rijcke et al., 2015).
5. **‘Supported by (academic) power’ vs. ‘Underground’**. This meaning focuses on the socioacademic dimension of the mainstream, stressing the connection between mainstream and power. Mainstream science is what is defended by academic elites in prestigious universities and supported by economic and industrial powers. By contrast, nonmainstream science is seen as resistant. The underground approaches are not published in mainstream journals and are unlikely to receive funding through normal channels, even though they might receive funding from alternative sources. This meaning echoes the distinction between the underground or independent labels and the ‘majors’ in the music industry.<sup>4</sup> In economics, it is not unusual to describe the difference between mainstream and heterodox schools by pointing out not only the theoretical divergences, but also the different relationships that they entertain with economic powers (Cedrini & Fontana, 2018; Colander et al., 2004).
6. **‘Core’, ‘Western’ vs. ‘Periphery’, ‘Non-Western’**. This meaning of mainstream is eccentric compared with the others and is found only in the bibliometric literature on the scientific production of developing countries (e.g. Gasparyan et al., 2017). In this literature, mainstream is used as a synonym of Western, and

---

<sup>4</sup> [https://en.wikipedia.org/wiki/Music\\_industry](https://en.wikipedia.org/wiki/Music_industry)

Mainstream meaning	Opposite	Focus on	Example
1 Orthodox	Heterodox	Intellectual contents	[Gottfredson, 1997]
2 Normal	Revolutionary	Mode of scientific progress	[Heinze, 2013]
3 Popular	Niche	Quantity	-
4 Trendy, short-lived, passing	Stable, long-lasting	Temporal extension	[de Rijcke, 2015]
5 Supported by (academic) power	Underground	Power	[Colander, 2004]
6 Core	Periphery	Geopolitics	[Gasparyan, 2017]

**Fig. 1** The six different meanings of ‘mainstream’ in reference to science. The notion of mainstream has no universal meaning in discussions about science and science policy. Specifically, six different meanings can be analytically distinguished in the literature, noting that sometimes two or more meanings are intended at the same time. In the first column of the table, the key terms that capture each meaning are presented, along with their opposites (second column) that contribute to specifying their semantic content. Each meaning stresses a different dimension of the notion of mainstream, focusing on various aspects of the scientific activity. In the last column of the table, examples of studies that employ each meaning of the notion are provided

mainstream science is the scientific research either produced by highly developed countries or published in international outlets. By contrast, nonmainstream science is the science produced in developing countries and published in local journals.

Note that the six meanings, even if they are closely related, should not be considered synonyms. In fact, they are not mutually implied. For instance, a molecular biologist can deal with a niche topic (nonmainstream according to meaning 3) by applying a standard experimental method (mainstream according to 1). As a further example, an astrophysicist can investigate a ‘trendy’ celestial object (mainstream according to meaning 4) but in the context of a heterodox cosmological model (nonmainstream according to meaning 1). In Fig. 1, a summary of the six meanings, their opposites and the aspect of *mainstreamness* they highlight are provided.

## 4 Modelling Mainstreams

Previous approaches to the mainstream definition have led to different operational definitions of it.

The meanings 1 and 5 (‘orthodox’ and ‘supported by power’) require considerable expert knowledge of the scientific fields to assess whether a publication belongs

to the mainstream. To empirically investigate the mainstream that is intended in this sense implies gathering the opinion of several experts, with evident limitations in the number of publications that can be considered. Meaning 6 ('core') is easier to treat with quantitative methods because the geographical information can be retrieved automatically from the publications' metadata. However, this meaning of mainstream is less interesting from the point of view of the debate on research metrics.

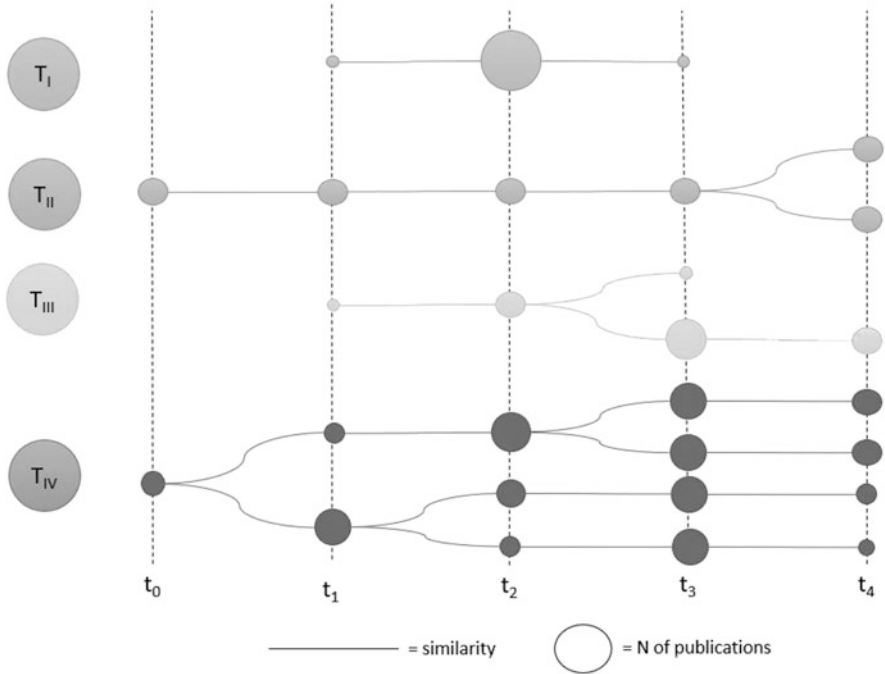
Hence, we remain with meanings 2, 3, and 4, that is, 'normal science', 'popular', and 'trendy'. With relative ease, meanings 3 and 4 can be translated into quantitative measures. Popularity can be measured by the number of publications addressing a topic, whereas the trendiness of a topic can be measured by its temporal extension. Meaning 3 is particularly interesting because it refers to the epistemological concepts of normal versus revolutionary science advanced by Kuhn. Some observations by Kuhn and Lakatos can help us translate (partially) these notions into measures. According to Kuhn, during the normal science period, a paradigm is 'articulated' by the researchers, that is, it is expanded in different directions. Lakatos calls these paradigm articulations 'progressive research programmes' (1978). A progressive research programme can be recognised by its capacity to produce new research lines, that is, by its fruitfulness (Ivani, 2019). Thus, meaning 2 can be measured as a factor of productivity or the fruitfulness of a topic.<sup>5</sup>

The proposed approach to mainstream detection integrates the following three meanings of the term: *popularity* (meaning 3), *trendiness* (meaning 4), and *fruitfulness* (meaning 2). They constitute the three dimensions of what is mainstream that will be considered in our study. Based on them, several profiles of mainstream can be outlined (Fig. 2):

1. **Spot.** This profile corresponds to a short-lived topic characterised by a short burst of attention from the research community that is focused in a limited interval of time. This profile mostly relies on meaning 4 ('trendy').
2. **Persistent.** This profile of mainstream is based on meaning 3, popularity. It describes a topic that enjoys stable attention from the research community but that has low productivity in terms of new research lines.
3. **Impasse.** This profile describes the behaviour of a research programme that progressively decreases in importance until it becomes marginally important.
4. **Boosting.** This mainstream profile corresponds to a fruitful research programme of the normal scientific phase, hence relying mostly on meaning 2. It is characterised by a long life with a high number of descendant topics.

---

<sup>5</sup> Clearly, both Kuhn's and Lakatos' theories of scientific change are far more complex and richer than the sketchy picture offered in this report. In fact, we do not aim to offer a full operationalisation of these theories. Our limited goal is to draw on some epistemological topics to better design our methodology.

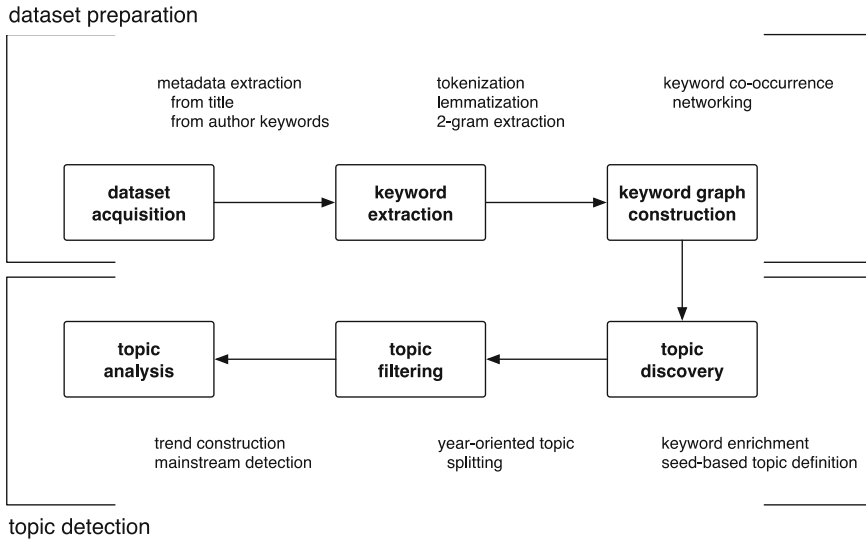


**Fig. 2** Mainstream profiles. By combining the three meanings or aspects of the notion of mainstream that can be quantified (i.e. popularity, trendiness, and fruitfulness), it is possible to delineate the various temporal profiles of a mainstream topic, that is, the various modes in which a mainstream topic may develop over time. The figure shows four of these modes. From the top to the bottom of the figure, they are as follows: spot topic (a short-lived topic that attracts a burst of attention in the research community), persistent topic (a topic that enjoys stable attention in the community but does not produce new research lines), impasse topic (a topic that branches in research lines, some of which decay), and boosting topic (a topic characterised by high fruitfulness that produces several new research lines). In the figure, the relation of filiation within a topic is represented by lines, whereas the size of the research lines (quantified in terms of publications) that form a topic is represented by circles

In different ways, each of these ideal profiles of mainstream integrates the three core aspects of meanings 2, 3, and 4. Our method aims at individuating instances of such profiles into the scientific production of our case studies.

## 5 Semiautomatic Topic Detection

Consider a dataset of scholarly publications  $P = \{p_1, p_2, \dots, p_n\}$ . For topic detection in  $P$ , we propose the approach shown in Fig. 3 based on a pipeline characterised by *dataset acquisition*, *keyword extraction*, *keyword graph construction*, *topic discovery*, *topic filtering*, and *topic analysis*. In the following, we first present



**Fig. 3** The proposed mining approach to topic detection. The proposed topic mining approach is based on a pipeline where the initial publication dataset with related metadata is first submitted to a *keyword extraction* stage aimed at extracting relevant tokens. The tokens are then organised in a graph based on keyword co-occurrences within publications (i.e. *keyword graph construction*). The subsequent steps of *topic discovery* and *topic filtering* are applied to generate the set of topics emerging from the publications. Finally, a *topic analysis* is enforced to determine trends over topics and mainstream behaviours

dataset acquisition, keyword extraction, and keyword graph construction as the preparation steps; then, we focus on the subsequent activities related to topic discovery, filtering, and analysis.

## 5.1 Dataset Preparation

Dataset preparation has the goal of extracting keywords from publications that are representative of the study's focus. Moreover, once those keywords have been extracted, preparation aims at explicitly representing the distribution of keywords over publications so that the co-occurrence of the same keywords in publications is highlighted.

An initial step of *dataset acquisition* is extracting the metadata of each publication  $p \in P$ , namely the title and the author-assigned keywords. The *keyword extraction* step is then executed on the publication metadata by applying conventional NLP techniques, such as tokenisation, lemmatisation, and 2-gram recognition based on mutual information (Manning et al., 2008). A *keyword set*  $K_p$  is associated with each publication  $p \in P$  as a result. The step of *keyword graph construction* is finally

executed to highlight when keywords co-occur in the publication descriptions, namely in the associated keyword sets. The result is a graph  $G = (N, E)$ , where  $N = \bigcup_{i=1}^n K_{p_i}$  is the set of nodes constituted by the overall set of publication keywords extracted from the metadata and  $E$  is the set of graph edges connecting pairs of keyword nodes. An edge  $e_{ij} = (n_i, n_j, w_{ij})$  denotes that the keyword represented by the node  $n_i$  co-occurs with the keyword of node  $n_j$  in the keyword sets of the publication descriptions. The weight  $w_{ij}$  denotes the strength/relevance of the  $n_i, n_j$  co-occurrence, namely the number of publications in which  $n_i, n_j$  co-occur.

*Example* As an example of data preparation, we consider publication p1 with the title ‘Bologne et le Cardinal Légat Bertrand du Pouget’ and the following author-assigned keywords: avignon, bertrand du pouget, bologne, cardinal légat. The following keyword set  $K_{p1}$  is extracted for the publication p1:

$$K_{p1} = \{ \underline{\text{avignon}}, \text{bertrand}, \text{bertrand du pouget}, \underline{\text{bologne}}, \text{cardinal}, \text{cardinal légat}, \text{légat}, \text{pouget} \}$$

Similarly, consider a further publication p2 characterised by the following keyword set  $K_{p2}$ :

$$K_{p2} = \{ \underline{\text{avignon}}, \text{bertrand du pouget}, \underline{\text{bologne}}, \text{histoire de l'Église}, \text{jean xxii} \}$$

In the keyword graph construction step, each item of the sets  $K_{p1}$  and  $K_{p2}$  becomes a node of the graph  $G = (N, E)$ . Call  $n_a$  the graph node for the keyword avignon and  $n_b$  the node for the keyword bologne. The edge  $e_{ab} = (n_a, n_b, 2)$  is defined in  $G$  to denote that the keywords avignon and bologne co-occur in two publications (i.e. p1 and p2); thus, the weight of the edge between their respective nodes is 2.

## 5.2 Topic Discovery

The keywords used for describing publications are characterised by *sparseness*, meaning that the terms appearing in the keyword set  $K_p$  of a publication  $p$  are usually highly focused and are rarely employed in the keyword set of other publications. To reduce the impact of keyword sparseness and capture possible topic overlaps among publications, we exploit the idea that the keywords of a publication can be enriched with the keywords of other publications when these keywords are frequently co-occurring and, thus, when they are used within the same terminological context. For topic discovery, each publication  $p \in P$  is associated with an *enriched keyword set*  $\overline{K}_p$  that has the goal of describing the publication  $p$  with keywords that are general enough to reveal the publication topic instead of the publication focus.

For a publication  $p \in P$ , the construction of the set  $\overline{K}_p$  is described in the following way: Consider the keyword graph  $G = (N, E)$  built during dataset preparation and consider a keyword  $k_i \in K_p$ . We call *keyword co-occurrence context*



the set  $K_i^* = \{k_j : \exists e_{ij} (n_i, n_j, w_{ij}) \in E\}$  such that there is at least one co-occurrence relation between  $k_i$  and  $k_j$  in  $G$  (i.e. the two keywords co-occur in the description of at least one publication). Given the publication  $p$ , we call the *publication co-occurrence context* the set  $K_p^* = \bigcup_{k_i \in K_p} K_i^*$ . The set  $K_p^*$  contains keywords that are not directly used to describe the publication  $p$  but that co-occur with the keywords of  $K_p$  in other publications. Each keyword  $k_j \in K_p^*$  is associated with a weight  $\omega_j$  to denote the relevance of the keyword  $k_j$  in describing the topic of the publication  $p$ . For a keyword  $k_j \in K_p^*$ , the weight  $\omega_j$  is calculated as follows:

$$\omega_j = \frac{1}{\max_{k_z \in K_p^*} \omega_z} \sum_{k_i \in K_p} \sum_{k_j \in K_i^*} \alpha + w_{ij}$$

where  $\alpha \in \mathbb{N}$  is a constant parameter and  $w_{ij}$  is the weight associated with the edge  $e_{ij}$  in the graph  $G$ , which denotes the number of co-occurrences in the publications of the keywords  $k_i \in K_p$  and the keyword  $k_j \in K_p^*$ . The  $\alpha$  parameter is introduced to support a flexible definition of the weight  $\omega_j$  associated with a keyword  $k_j \in K_p^*$ . In particular, the value of  $\alpha$  is added to the weight  $\omega_j$  each time a keyword  $k_i \in K_p$  co-occurs with a keyword  $k_j \in K_p^*$ . When low values of  $\alpha$  are considered (i.e.  $\alpha = 0$  or  $\alpha = 1$ ), the weight  $\omega_j$  mostly depends on the weight  $w_{ij}$  of the co-occurrences of the keyword  $k_j$  with the keywords of  $k_i \in K_p$ . When high values of  $\alpha$  are considered, the weight  $\omega_j$  is increased each time a co-occurrence of the keyword  $k_j$  is found with the keywords of  $k_i \in K_p$ , despite the strength of the weight  $w_{ij}$ . This means that when  $\alpha$  is high, we assign more importance to the keywords  $k_j \in K_p^*$  that have numerous co-occurrences with the keyword  $k_i \in K_p$  and give less importance to the weight  $w_{ij}$  of such co-occurrences.

Finally, the enriched keyword set of a publication  $p$  is defined as  $\overline{K}_p = \{k_j : k_j \in K_p^* \wedge \omega_j \geq th\}$ , where  $th$  is a prefixed threshold to distinguish relevant versus nonrelevant keywords to include in  $K_p$ . Finally, a new graph  $\overline{G} = (N, \overline{E})$  is generated according to the enriched keyword sets  $\overline{K}$ . In  $\overline{G}$ , the edges  $\overline{E}$  denote the keyword co-occurrence in the enriched keyword sets  $\overline{K}$  of the publications.

According to the enriched co-occurrence graph  $\overline{G} = (N, \overline{E})$ , we provide the following topic definition:

**Topic** A topic  $T_s$  is a set of featuring keywords that describes a common research argument. A topic  $T_s$  is defined around a *seed keyword*  $k_s$  that represents the label/name of the research argument. Given a seed keyword  $k_s$  associated with a corresponding keyword node  $n_s \in N$  in  $\overline{G}$ , the topic  $T_s$  corresponds to the set of keywords associated with the nodes  $N_s \subseteq N$  connected with  $k_s$  in the enriched co-occurrence graph  $\overline{G} = (N, \overline{E})$ , namely  $N_s = \{n_j : \exists \overline{e}_{sj} (n_s, n_j, w_{sj}) \in \overline{E}\}$ .

We say that a publication  $p$  is about a topic  $T_s$  when at least one common keyword exists between the enriched keyword set  $K_p$  and topic  $T_s$ , namely  $K_p \cap T_s \neq \emptyset$ .

$K_p^*$	$K_p$			$K_p^*$	$\alpha$ value		$\omega_j$ values	
	papacy	xiv century	avignon		$\alpha = 1$	$\alpha = 4$	$\alpha = 1$	$\alpha = 4$
modern era	5			modern era	5	9	0.62	0.56
papacy		1	2	papacy	3	11	0.37	0.69
xiv century	1			xiv century	1	5	0.12	0.31
avignon	2			avignon	2	6	0.25	0.37
history		3		history	3	7	0.37	0.44
church history	1	2	1	church history	4	16	0.50	1.00
middle ages	2	6		middle ages	8	16	1.00	1.00

**Fig. 4** Example of keywords and topics. An example of topic discovery. Given a publication  $p$  with keywords  $K_p$  and context  $K_p^*$ , we show the keyword co-occurrences in the publications of the dataset (left side) and the weight  $\omega_j$  of each keyword  $k_j \in K_p^*$  with two different setting of the  $\alpha$  parameter (right side)

*Example* As an example of topic discovery, in Fig. 4, we show the excerpt of a keyword set  $K_p$  and related co-occurrence context  $K_p^*$ :

$$K_p \subseteq \{\text{avignon, papacy, xiv century}\}$$

$$K_p^* \subseteq \{\text{avignon, church history, history, middle ages, modern era, papacy, xiv century}\}$$

On the left side of Fig. 4, we show the number of co-occurrences between any pair of keywords  $k_i \in K_p$  and  $k_j \in K_p^*$ . For instance, given  $k_i = \text{papacy}$  and  $k_j = \text{modern era}$ , the value  $w_{ij} = 5$  is shown in Fig. 4 as denoting that papacy and modern era co-occur in five publications, namely  $e_{ij} = (n_i, n_j, 5)$  is set in the graph  $G$ .

On the right side of Fig. 4, we show the weight  $\omega_j$  of each keyword  $k_j \in K_p^*$  when two different settings of the  $\alpha$  parameter are considered. When  $\alpha = 1$ , the keywords  $k_j \in K_p^*$  with a higher weight  $\omega_j$  are middle ages and modern era, which are the keywords with highest  $w_{ij}$  value. It is interesting to note that modern era has a high  $\omega_j$  weight, even if this keyword only co-occurs with papacy in  $K_p$ . When  $\alpha = 4$ , the keywords  $k_j \in K_p^*$  with a higher weight  $\omega_j$  are church history and middle ages, which are the keywords that co-occur with most of the publication keywords in  $K_p$ . It is interesting to note that the weight  $\omega_j$  of modern era is strongly reduced when  $\alpha = 4$ . According to this example, by considering a threshold  $th = 0.8$ , the enriched keyword set  $\overline{K_p}$  is defined as follows:

$$\overline{K_p} = \{\text{middle ages}\} \text{ (when } \alpha = 1 \text{);}$$

$$\overline{K_p} = \{\text{church history, middle ages}\} \text{ (when } \alpha = 4 \text{).}$$

### 5.3 Topic Filtering and Analysis

The ultimate goal of the proposed approach to topic detection is to analyse topics over time to highlight possible mainstream behaviours. To this end, *topic filtering* is executed to split the co-occurrence graph  $\overline{G}$  into a set of subgraphs  $\overline{G}_Y = (N_Y, \overline{E}_Y)$ , where each one is related to keyword co-occurrences in a specific year  $Y$ . A graph  $\overline{G}_Y \subseteq \overline{G}$  is constituted by *i*) the nodes  $N_Y = \bigcup_{i=1}^{i=k} \overline{K}_{p_i}$ , where the sets  $\overline{K}_{p_i}$  are the enriched keyword sets of the publications from the year  $Y$  and *ii*) the edges  $\overline{E}_Y$ , where an edge  $\overline{e}_{ij} \in \overline{E}_Y$  is defined as  $\overline{e}_{ij} = (n_i, n_j, w_{ij})$  and connects two keyword nodes  $n_i, n_j \in N_Y$ . The weight  $w_{ij}$  denotes the number of publications from the year  $Y$  in which the keywords  $n_i, n_j$  co-occur. We note that the number of publications per year can be (very) different from one year to another. As a result, for comparison of keyword co-occurrence weights across consecutive years, given an edge as  $\overline{e}_{ij} = (n_i, n_j, w_{ij})$  in the year  $Y$ , the number of co-occurrences  $w_{ij}$  is normalised by the overall number of publications from the year  $Y$ .

Given a seed keyword  $k_s$  with the associated keyword node  $n_s$ , a topic  $T_s$  in the year  $Y$  corresponds to the set of keyword nodes  $N_{Y_s} \subseteq N_Y$  in the subgraph  $\overline{G}_{Y_s} \subseteq \overline{G}_Y$  where  $N_{Y_s} = \{n_j : \exists \overline{e}_{sj} (n_s, n_j, w_{sj}) \in \overline{E}_Y\}$ .

As a result of topic filtering, a topic  $T_s$  can change over time because the set of keywords  $T_s$  that characterises the topic can vary from one year to another. Moreover, when a topic is associated with a stable pair of keyword nodes  $n_i, n_j$  in two consecutive years  $Y$  and  $Y + 1$ , it is possible that the co-occurrence weight  $w_{ij}$  is different in the two considered years. As a result, the *topic analysis* step is executed to observe the behaviour of topics along time/years. In particular, the goal of this step is to recognise the possible mainstream topics according to the following definition:

**Mainstream Topic** A mainstream topic  $M$  is a topic whose trend within a certain time interval of years  $[Y_1, Y_2]$  follows one of the mainstream profiles presented in Sect. 4, namely spot, persistent, impasse, or boosting.

*Example* As an example, we consider the following enriched keyword sets associated with six publications in the time interval from 2015 to 2017:

- $\overline{K}_0(2015)$ : church history, middle ages
- $\overline{K}_1(2015)$ : christianity, middle ages, history
- $\overline{K}_2(2016)$ : christianity, catholic church, philosophy
- $\overline{K}_3(2016)$ : middle ages, philosophy
- $\overline{K}_4(2017)$ : catholic church history, middle ages
- $\overline{K}_5(2017)$ : philosophy, middle ages

In Fig. 5, we show a tabular representation of the graph  $\overline{G}$  built according to the above keyword sets  $\overline{K}$ . Consider a seed keyword  $k_s = \text{middle ages}$ . All the keywords of Fig. 5 have at least one co-occurrence with the seed keyword; thus, they are all belonging to the considered topic  $T_s$  about ‘Middle Ages’. The strength of the

	church history	middle ages	christianity	history	catholic church	philosophy
church history	-	1	0	0	0	0
middle ages	1	-	1	2	1	2
christianity	0	1	-	1	1	1
history	0	2	1	-	1	0
catholic church	0	1	1	1	-	1
philosophy	0	2	1	0	1	-

**Fig. 5** Example of co-occurrence graph  $\overline{G}$  for a set of enriched keywords. As an example in the framework of topic discovery, the figure reports the number of co-occurrences within the publications for each pair of the considered keywords

$T_{middle\ ages}$	2015	2016	2017
church history	0.5	0	0
christianity	0.5	0	0
history	0.5	0	0.5
catholic church	0	0	0.5
philosophy	0	0.5	0.5

**Fig. 6** Example of keyword weight in the years 2015–2017 about the topic ‘Middle Ages’. As an example of topic trend, we show the weight of the keywords associated with the topic ‘Middle Ages’ in the time interval of the years 2015–2017

co-occurrences between the seed keyword and keywords of Fig. 5 in the years 2015–2017 is shown in Fig. 6 (normalised value by the number of publications in each considered year). By observing the keyword strength in the considered time interval, we can envisage possible mainstream topic profiles. In particular, for the topic  $T_{middle\ ages}$ , the keyword history denotes a *persistent topic* behaviour (despite showing little fluctuation in 2016). We also note that the keywords church history and christianity denote an *impasse topic* behaviour, while the keyword philosophy denotes a *boosting topic* behaviour for the topic  $T_{middle\ ages}$ . As a final consideration of the observed mainstreams, we could claim that in the context of the Middle Ages studies, an initial interest in the History of the Church and Christianity shifted towards more philosophical studies about Catholicism.

## 6 Case Study Analysis

In this section, we present the results obtained by applying the proposed approach and related techniques for topic mining on a real publication dataset taken from selected institutional research archives of Italian universities. The main idea is to provide a clear description of the results we obtained, here by focusing on a

few disciplines and using institutional publications data provided by four Italian universities. Regarding the case study analysis, an Online Appendix is provided with complementary figures and comments. The Appendix is available for download at the following link: [http://islab.di.unimi.it/content/maverick\\_data/appendix.pdf](http://islab.di.unimi.it/content/maverick_data/appendix.pdf).

## 6.1 Dataset Description

The proposed case study is based on a publication dataset collected from selected Italian universities. In the early 2000s, most Italian universities started to populate and maintain institutional research archives for persistently storing publications and research products. In particular, each university supported the creation of its own repository based on products published (and compulsorily uploaded) by its affiliated scholars. In selecting both universities and research areas to consider for building the dataset of the case study, we relied on the following recommendations: i) choose large, representative Italian universities, ii) choose a few selected research areas, and iii) compose a dataset that is representative of both bibliometric and nonbibliometric research areas according to the Italian regulation for research evaluation. As a result, the following four Italian universities have been selected: UNIBO—University of Bologna (with 2896 academic researchers—data consulted on April 15, 2021, from <https://cercauniversita.cineca.it/php5/docenti/cerca.php>), UNIMI—University of Milan (with 2258 academic researchers), UNIRM—University of Rome ‘La Sapienza’ (with 3350 academic researchers), and UNITO—University of Turin (with 2086 academic researchers). Moreover, among all the available disciplines, we focus on those publications authored by all scholars of the following research areas (as defined by The Italian National University Council—CUN): A01 (mathematics and informatics), A11 (history, philosophy, pedagogy, and psychology), and A13 (economics and statistics).

A summary view of the collected dataset is provided in Fig. 7. The dataset contains 123,504 publications labelled with 124,820 author-defined keywords.

	UNIBO	UNIMI	UNIRM	UNITO	Total
A01 pubs	4,831	7,227	8,889	7,705	28,652
A11 pubs	13,313	8,196	24,737	17,007	63,253
A13 pubs	11,817	3,455	12,436	6,762	34,470
Total pubs	27,223	18,805	46,002	31,474	123,504
Keywords	39,624	22,634	38,105	24,457	124,820

**Fig. 7** Summary picture of the Maverick dataset. Number of publications and keywords in the Italian case study by university (Univ. of Bologna, Univ. of Milan, Univ. of Rome, and Univ. of Turin) and discipline (scientific area of study as classified by ANVUR)

For topic mining purposes, the keywords and publication titles are exploited. It is important to note that 58,585 publications of the dataset do not provide any keywords. For these publications, only the keywords extracted from the title are then used for topic mining. As a further remark, we observe that the number of publications per year is not constant. In fact, at the beginning of the 2000s, only a few publications were inserted into the selected archives by their authors, and this practice became a regular—and often compulsory—routine only around the year 2005. For this reason, the considered publications in this empirical exercise cover the years from 2005 to 2018. It is also important to stress that the number of publications per year is continuously increasing throughout the whole observed period because of the increasing role of performance-based exercises in Italy; thus, a normalisation step is required when the analysis focuses on the consistency of topics across different years.

## 6.2 General Results in Italian Academia

The results obtained on the considered dataset are briefly discussed by separately exploiting the publications of each considered research area, namely A01, A11, and A13.

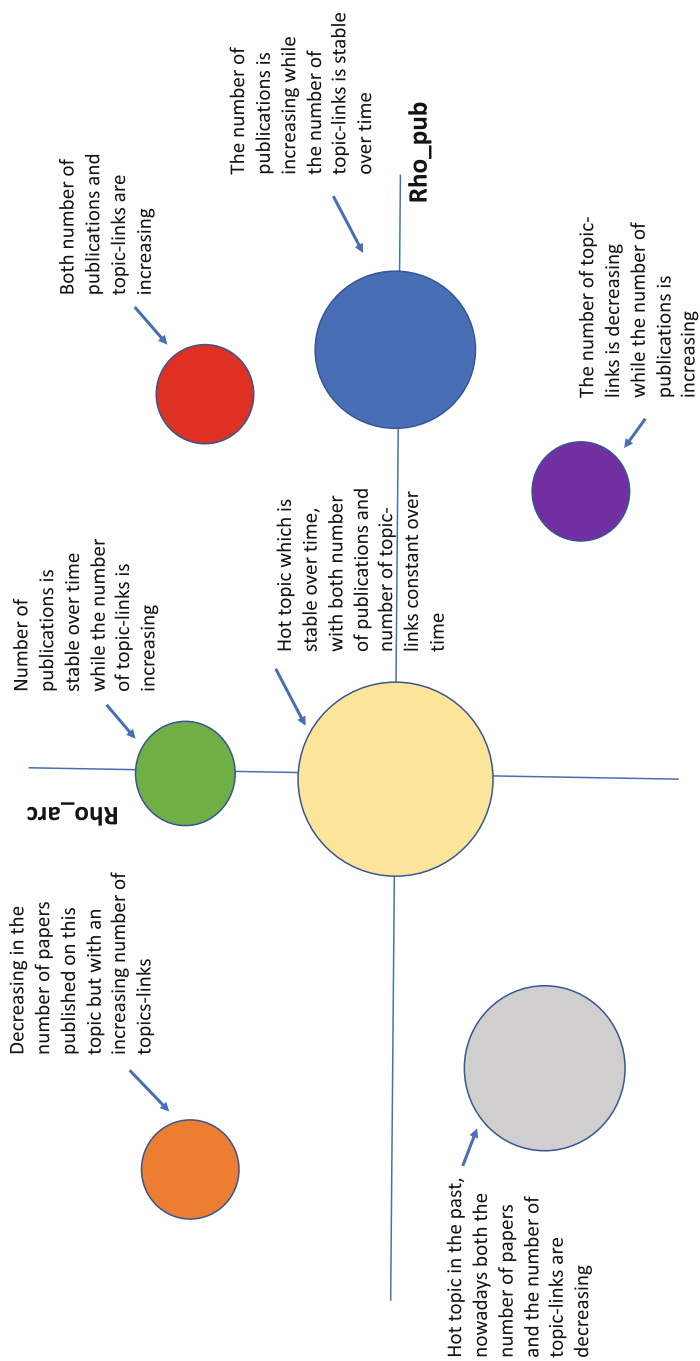
First, to identify the mainstream profiles defined in Sect. 4 using the bibliometric data collected for this case study, we started defining a couple of synthetic operators to describe a topic's behaviour over time. For any given topic  $k$  belonging to its  $G$  graph, we first generate a matrix of the number of links with all the other existing topics within the same discipline during the observed years, where in each of its cells, we have the corresponding number of papers. Then, we compute two simple correlation coefficients,  $\rho_{kt,j}$ , for any identified topic  $k$ : a pair of  $\rho$ -s (namely  $\rho_{k,t}$  and  $\rho_{k,j}$ ) that represents the correlation coefficients of the topic's number of publications published over time ( $t$ ) and of the number of topic links that the selected topic ( $k$ ) establishes with the other topics ( $j$ ) in the discipline, respectively.

Figure 8 shows how a topic may behave according to different combinations of the values defined by its pair of  $\rho$ -s coefficients. Using the computed  $\rho$  coefficients, it is possible to broadly map the mainstream topic profiles defined in Sect. 4 in the area defined by the two  $\rho$  pairs.

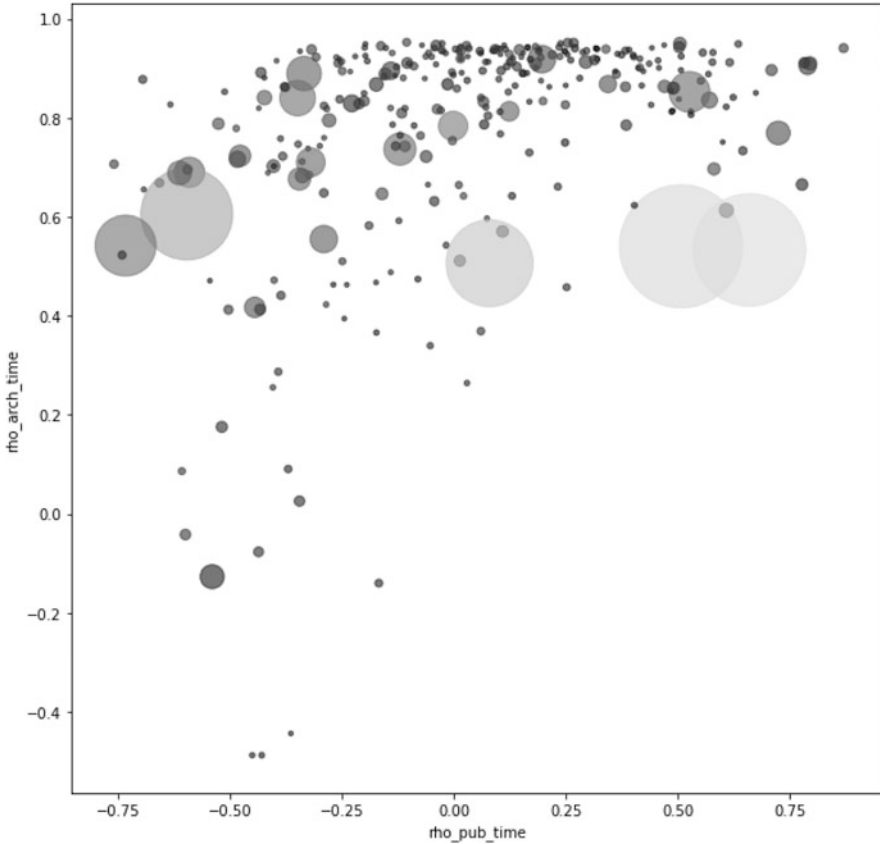
A *spot topic*, for example, corresponds to a short-lived topic that, after a burst of attention from the research community in the past, is now abandoned. A 'trendy' topic within a discipline appears in the bottom left area of Fig. 8 (e.g. grey circle).

An *impasse topic* describes the development of a research programme having some topic links that died in recent years; this is in the bottom middle area of the graph (e.g. below the big yellow circle).

A *persistent topic* identifies a mainstream topic that enjoys stable attention from the research community but with low productivity in terms of links with new research lines. This topic appears in the bottom right part of Fig. 8 (e.g. purple circle).



**Fig. 8** Topic behaviour in the Italian case study according to  $\rho$ -s coefficients. Figure 8 shows how a topic may behave according to different combinations of the values defined by its pair of  $\rho$ -s coefficients. Examples of topic characteristics are provided in different zones of the provided space to show how the mapping of mainstream topic profiles defined in Sect. 4 can be obtained in the  $\rho$ -s coefficients space



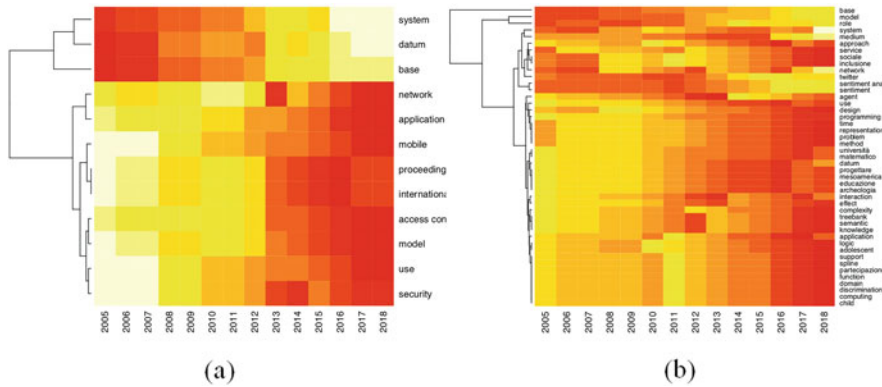
**Fig. 9** Mathematics and informatics. Each of the considered disciplines in this empirical study is visualised on a map (Mathematics and informatics only is reported above), showing a circle for each topic characterising the discipline during the period under investigation. Circle size and colour represent the number of topic links and topic size (e.g., number of publications), respectively

A *boosting topic*, which has been described in Sect. 4 as a topic characterised by a long life and a high number of connections with other topics, is the top centre or top right corner of Fig. 8 (e.g. green and/or red circles).

In addition to this, Fig. 8 may be useful to identify *niche* topics, like the orange-like circle in the top left area, which is characterised by a decreasing number of papers published in the past few years along with an increasing number of topic links.

For each one of the disciplines considered in this empirical study, a figure has been created that visualises a map similar to Fig. 8, with a circle for each topic characterising the discipline during the period under investigation. Circle size and colour represent the number of topic links and topic size (e.g. number of publications), respectively. Each map describes the corresponding discipline using the proposed topic approach (Fig. 9).





**Fig. 10** Examples of a heatmap for mathematics and informatics. In (a) the topic *privacy*, in (b) the topic *social*. Examples of a heatmap for mathematics and informatics. Each measure of topic link evolution over time is standardised by its dimension to generate a comparable heatmap that clearly visualises the temporal topic’s dynamics in the discipline

Then, for each of the relevant topics of the identified mainstream categories, we compute a matrix describing the evolution over time of all the topic links generated by the topic itself, here based on the structure described in Fig. 6. Each measure of the evolution of a topic link over time is standardised by its dimension to generate a comparable heatmap that clearly visualises the temporal topic’s dynamics. The case study on A01 is reported here, including the figures mentioned above. For areas A11 and A13, the figures are reported and described in the Online Appendix (see Figs. 1–4 in the Appendix).

**Area 1—Mathematics and Informatics** As a first example, Fig. 10a provides the heatmap of the topic ‘**privacy**’. This is an *impasse topic* with negative values for both the correlation coefficients ( $\rho$ -s). Each row of the heatmap represents the topics with which the topic ‘privacy’ reports links over time, and the colour indicates the intensity of such links. Red means few links, whereas white means many links. According to this heatmap, the topic ‘privacy’ used to be linked to topics like ‘access’, ‘security’, and ‘network’ in the past, whereas recently, it started to be associated with different topics such as ‘data’ and ‘systems’. This dynamic seems to be pretty much in line with the current increasing availability of new sources of (individual) data, for example, hospital individual data or credit bank transactions, which consequently challenges new issues related to data ‘privacy’ concerns.

Figure 10b provides the heatmap of the topic ‘**social**’. Both  $\rho$ -s are positive, which makes it a *boosting topic*. The corresponding heatmap suggests that this topic is now (in 2016–2018) very much linked with topics like ‘sentiment analysis’ and ‘social networks’ (e.g. Twitter).

This is again very much in line with the new and fast-growing literature that uses ‘big data’ to extract indicators to summarise, for example, users’ opinions. By contrast, at the beginning of the sample period, the topic ‘social’ was associated

with more traditional topics like ‘education’, ‘participation’, ‘university’, and ‘discrimination’.

**Area 11—History, Philosophy, Pedagogy, and Psychology** Turning to Area 11, the topic ‘ageing’, represented in detail in Fig. 2a of the Online Appendix, provides another interesting example of an *impasse topic* with both negative correlation coefficients ( $\rho$ -s).

In the most recent years, this topic has very much been associated with ‘experience’, ‘activity’, ‘creativity’, ‘life’, and ‘health’, while in the previous years, it used to be linked with discussions and studies more focused on the past (i.e. ‘history’ and ‘wars’).

Because the problem of the ageing of the population is increasingly and extremely relevant, the topic ‘ageing’ seems to be now more associated with discussions related to aged people’s quality of life (both in terms of health and wealth) and their occupations rather than their past historical memories.

In addition, the topic ‘female studies’ is a good example of a *boosting topic* (both rho-s are positive and high). According to graph Fig. 2b in the online supplementary material, this topic has been recently associated with topics like ‘child’ and ‘adolescent’ (which suggests an emerging focus on the relationship between mothers and children), ‘male’ (which points to gender-related studies), or ‘patient’ and ‘effect’ (which relates to the literature of causal analysis of health issues, which often may provide heterogeneous effects by gender).

By contrast, in the past, ‘female studies’ have been associated with topics related to women’s mental status (e.g. ‘mental health’, ‘stress’, ‘personality’, ‘attention’, ‘perception’, ‘memory’, ‘brain’, and ‘neuro’). In addition, it used to be associated with ‘work’ and ‘quality of life’, which may refer to work–life balance issues that appeared commonly in the literature.

**Area 13—Economics and Statistics** Figure 3 in the Online Appendix shows a pretty different scenario for Area 13 compared with Area 11 and Area 1.

In fact, Fig. 4 in the Online Appendix shows three heatmaps for the three following topics: ‘development’, ‘taxation’, and ‘network analysis’.

‘**Network analysis**’ can be defined as “*a set of integrated techniques to depict relations among actors and to analyse the social structures that emerge from the recurrence of these relations*” (see Smelser & Baltes, 2001).

From our analysis, it may be characterised as a boosting topic that exhibits positive values of rho-s. Although in the past it focused on theory (being related with abstract analysis) and empirical analysis, it has been recently applied among economists, econometricians, and statisticians to topics such as ‘sentiment analysis’, ‘Twitter’, and ‘social media’, generating a new strand of literature studying a ‘network analysis’ taking advantage of the new sources of (big) data now available.

On the contrary, a clear example of an *impasse topic* in economics and statistics is represented by the topic ‘**taxation**’. The economics of taxation mainly collects studies regarding both the effects and consequences of taxes on economic decisions, as

well as on how to efficiently design tax systems (e.g. income, capital, environmental taxes).

For this topic, both *rho*-s are negative, meaning that there has been a decreasing interest in this topic over the past decade. However, looking carefully at its development over the past few years (see Fig. 4c in the Online Appendix), it seems quite reasonable to identify dead links with topics such as ‘literature review’, ‘inequality measures’, ‘country taxation’, and ‘equity’ in favour of new emerging trends with topics like ‘income distribution’, ‘evidence’, and ‘effects’, which are very much in line with recent works on the global evolution of inequality, taxation top income dynamics, progressive wealth taxation, and so forth.

Finally, an example of a persistent topic is also identified in the economics and statistics area when looking at the topic ‘**development**’, for which both *rho*-s are almost close to zero. Development economics is a branch of economics that focuses on studies of economic, health, education, and social conditions in developing countries (especially low-income ones) compared with developed ones.

The heatmap for this topic, as represented by Fig. 4b in the Online Appendix, makes evident how this topic turns from being historically related with topics like ‘global’, ‘sustainability and growth’, or ‘industry’ in the past to new research frontiers aiming to explore how to estimate the ‘effects of policies’ and field interventions in emerging countries, often following—also in Italian academia—the studies winning Nobel Prizes in 2019 on the use of randomised clinical trials (RCTs) in this field to measure their ‘performances’, as well as on ‘innovation’ and ‘new perspectives’ in general.

### ***6.3 Robustness of the Proposed Approach Using an International Dataset over 14 Years***

To show the ability of the proposed approach to identify the existing publication topics, their evolution over time and their topic links in a broader (not only restricted to the national context as in the Italian case described before) and international context, we rely on different data sources: Scopus Elsevier. Through the Elsevier API service, we downloaded all the Scopus research products published between 2005 and 2018 that are classified as instances of at least one of the following subject areas (each journal may belong to more than one subject area): business, economics and econometrics, decision sciences, statistics and probability, and demography.

The dataset contains 1,700,286 unique papers published as articles (articles in press, editorial, erratum, and business articles), chapters, books, conference papers, notes, reviews, letters, and short surveys between 2005 and 2018 written by 1,433,297 different authors and labelled with 1,168,680 author-defined keywords. The obtained dataset is 12 times larger than the database analysed in the Italian case and covers almost all papers published in the selected disciplines by all the authors who are active in these research fields around the world. This database, even if not

perfect in terms of its coverage for all the existing disciplines (it is well known how social sciences and humanities or medicine are not perfectly represented by Scopus, see for example Archambault et al., 2006), provides a good set of information to explore in depth the ability of the approach to identify and group the topics in the literature.

From these data, we have selected two topics (**‘development’** and **‘taxation’**)—out of the several topics analysed before for the Italian case—to demonstrate the scalability of the proposed approach to a broader data source and the ability of the method to go deeper into the identification of the relevant topic links. As a matter of fact, the main contribution of this chapter is to propose a methodological approach to topic mining, showing its applicability to different disciplines and its reliability—in terms of the obtained results—when datasets of different richness and size are considered.

Obviously, the topic description may be slightly different depending on the different sets of authors and publications considered. For example, consider that in a given discipline, the Italian authors could have a different publishing behaviour in the 14 years analysed when compared with the authors who are active in the international literature. In this case, the two analyses may not perfectly overlap.

As for the **‘development’** topic, the approach applied to the Scopus dataset can identify several different topics within the ‘development’ field. From the scholarly publications in ‘development’ journals, the proposed approach identifies eight (sub)topics. A first topic, which is the most relevant in terms of publications, is broadly named ‘development’ and is classified as a boosting topic (both rho-s are positive and high) in the Scopus dataset. That is, it has around 3000 papers published (increasing over time) with an increasing number of links with other topics. This topic is a persistent one in the Italian case study. In addition to this, seven additional subfields in the ‘development’ area have been identified, describing a relevant heterogeneity within the field. Topics like ‘development finance’ and ‘development funding’ are both classified as impasse topics (with negative rho-s), exhibiting ended links with other topics and reduced attention from the researchers publishing in the discipline over the considered years. At the same time, this approach provides evidence of some new boosting topics (with both rho-s positive and large) in subfields like ‘development economics’ and ‘development strategies’. Moreover, a ‘development’ heatmap (Fig. 6 in the Online Appendix) shows how the emerging topics over the past few years of the analysed sample have focused on studying cultural, educational, agricultural, trade, and migration issues, along with managerial, institutional, and governance strategies, with special attention given to sustainability, climate change, and the evaluation of the effects of policies in the context of African and Asian countries (like China and India). This more detailed description of the field is in line with the one offered in the Italian case study but with an improved degree of available details on both the thematic issues and specific countries.

As for the **‘taxation’** topic, we have now a richer set of subtopics identified by our approach. Although the overall ‘taxation’ topic has received decreasing interest over the past decade in the Italian case study (as shown in the previous section),

the Scopus dataset shows how this field is highly heterogenous. Some topics show a decreasing interest from scholars (like ‘tax competition’), while other topics are clearly emerging (‘boosting topics’) in terms of both the number of papers and the number of topic links. Examples of these boosting topics are ‘tax incentives’, ‘tax havens’, ‘tax compliance’, and ‘tax morale’, which are identified as new emerging topics over the past 15 years.

A first heatmap on ‘taxation’ as a whole shows how the evolution of the international literature spans from links with topics like ‘regulation’, ‘redistribution’, and ‘welfare’ analysis in the early years to more recent developments in the field focused first on ‘income inequality’ (from 2011 to 2015) and then on ‘income distribution’, ‘tax reforms’, and ‘policy evaluation’ of interventions in countries like USA, Australia, and United Kingdom (Fig. 7 in the Online Appendix). Taking advantage of the richness of the considered Scopus dataset, we can go deeper in the analysis of these topics by looking at the heatmaps representing the links that even smaller subtopics in ‘taxation’ have established over time. For example, if we focus on the smaller topics identified by the proposed approach within the taxation field, we can identify ‘tax havens’ as a new emerging topic (which exhibits positive and large rho-s) with a rising number of publications from 2008 onwards. In addition, this subtopic has been interlinked since the very beginning (2008/2010) with topics such as ‘tax competition’, while from 2014 to 2018, it has been associated with topics like ‘tax avoidance’ and ‘tax evasion’, probably reflecting very recent contributions in the literature following the international debate on tax havens (e.g., the ‘Panama papers’ debate). Note that a similar degree of precision in describing the emerging or declining fields of research within a more general discipline like taxation studies could not be found when dealing with national subsamples of publications like the ones described in the Italian case study.

To conclude, the representation and discussion of some selected topics of the three disciplines analysed here shows how the proposed approach may be useful in describing both the geography of topics and the evolution and interlinkages of topics within a discipline by means of two datasets: i) institutional publications provided by four Italian universities and ii) international publications over 14 years. Moreover, the examples show how having a more comprehensive database of the worldwide production of papers is essential to prove the scalability of the proposed approach and reliability of the obtained results. All in all, richer and sizable datasets allow clearer and in-depth analyses to be provided. In particular, for a given number of papers available from the literature, the larger the number of the analysed topics, the sparser the topic links matrix. The cell size of this matrix is crucial to obtain reliable information on the temporal evolution of detailed topics and their interlinkage with new emerging or declining ones. Therefore, large bibliometric databases with millions of records can provide enough information to make this possible.

## 7 Concluding Remarks

The results from the case study show that the proposed approach to topic mining is capable of revealing trends of publication keywords and changes of these trends over time both when country-level data are available and when—even better—the larger international literature in a given field is considered. In combination with the contribution about mainstream modelling, this result represents a promising achievement, allowing us to recognise topic behaviours that can be associated with one of the profiles of the mainstream research defined in the project (i.e. spot, persistent, impasse, and boosting). In the following, we provide some considerations about possible extensions and applications based on our results.

**Possible Research Extensions** Future research activities could focus on i) the extension of the case study dataset in the Italian context to get complete coverage of the topic evolution of a given discipline in Italian academia, ii) the use of a discipline-specific keyword dictionary for a more refined topic cleaning, and iii) the comparison of case study results against third-party datasets similar to the case described in Sect. 6.2. The extension of the Italian case study dataset requires including the institutional research archive of additional Italian universities and may be a powerful tool to comprehensively analyse the topic's evolution and interlinkages across the different disciplines of Italian academia. On this point, we note that institutional research archives started to be populated at the beginning of the 2000s for almost all Italian universities. This means that i) the dataset adopted for the case study cannot be improved in terms of the size of the considered time interval and ii) the initial years of the considered time interval (i.e. the years from 2000 to 2004) are marginally useful for topic mining because few publications are present in the archives. As a result, through the extension of the available data, we aim to improve the richness of the publication corpus and increase the relevance of the case study in providing meaningful insights into the Italian picture in the period 2005–2018. A progressive inclusion of very recent publications from the considered universities is also required to keep the case study up to date. This may allow researchers from the various fields in the social sciences to study the evolution of their disciplines and the temporal changes that occurred in relation to a number of features, for example, new generations of researchers being more open towards international academia, the introduction of the research assessment exercises on the studied topics, and so forth. Moreover, the proposed approach applied to a complete country-specific bibliometric database may enable policy makers, such as ANVUR (Italian Agency for Evaluation of the University and Research System) or MIUR (Italian Ministry for Education, University, and Research) to design new policy interventions (which is their institutional mission) based on a solid analysis of 'what worked' (or not) in the past (as described in more detail in the possible applications to research evaluation provided below). A further issue for future research activities is the specification of a keyword dictionary for topic cleaning, possibly with a more detailed approach for each specific discipline. Sometimes, very general and poorly relevant keywords are included in the results of topic mining

activities. A manually defined dictionary of keywords can be set up to refine the results of keyword extraction and improve the quality of the discovered topics. Finally, a comparison of the obtained results against a third-party dataset can be also envisaged, here following the lines described in Sect. 6.2, to compare the topics found in the case study with the Italian community against a dataset that is representative of the international academic community. The goal is to observe possible similarities and/or peculiar behaviours of the Italian community compared with a larger, international group of scholars.

**Possible Applications to Research Evaluation** The topic trends that have emerged by applying the proposed techniques can be exploited to analyse changes in the publication practices of researchers along the temporal dimension. It is possible to apply these techniques to a publication dataset that is representative of the overall Italian Academy, meaning that almost all the institutional research archives of the Italian universities will be considered. A possible application scenario is to consider a ‘median scholar’ and the corresponding set of authored publications. By extracting the featured keywords from the ‘median scholar’ publications, it is possible to compare and correlate their research production against the topic trends associated with the scholarly keywords. In this way, shifts in the ‘median scholar’ interests can be tracked, as well as possible changes in terms of publication practices over time so that it is possible to observe whether the scholar’s behaviour endorses a topic whose trend can be recognised as mainstream according to specific time intervals. Similarly, one can focus on identifying heterogeneous publication patterns along the ability distribution, for example, studying if ‘top scholars’ behave differently than median or bottom ones. As a further application scenario, a similar approach can be enforced to analyse the changes that occur over time regarding a reference publication source (e.g. top journals) within a specific research area. In this way, it is possible to observe the evolution of ‘hot research topics’ in certain publication sources in correlation with the topic trends emerging from the already available results in that research area.

In addition to this, having access to the relevant data for the worldwide production of papers belonging to a specific discipline (as collected by standard bibliometric sources such as Scopus or Web of Science) may also enable a comparison of the national evolution of a discipline in a specific country (e.g. in Italy in our case) with respect to its own international benchmark. Moreover, a similar approach may also be adopted to study the effects of introducing a performance-based assessment exercises—as has happened in several countries around the world over the past decades—on the topics’ evolution in different disciplines at the local level.

Does the system of incentives provided by a performance-based assessment exercise have an impact on the evolution and choice of topics studied by academic scholars in their disciplines? Is there any evidence of a temporal shift towards international mainstream research (e.g. leaving niche topics aside) following the introduction of this type of assessment exercise? If so, is it socially optimal? All

these research questions (and probably many others) will be part of the future research agenda in this strand of the literature.

**Author Contributions** All the chapter authors equally contributed to the conceptualization, data curation/analysis, and interpretation of the results of this study. Sections 1 and 2 were primarily written by Eugenio Petrovich, Stefano Verzillo, and Stefano Montanelli; Sects. 3 and 4 by Eugenio Petrovich; Sect. 5 by Stefano Montanelli and Alfio Ferrara; Sect. 6 by Silvia Salini, Stefano Verzillo, and Corinna Ghirelli; Sect. 7 by Stefano Montanelli and Stefano Verzillo.

## References

- Abramo, G., D'Angelo, C. A., & Costa, F. (2019). When research assessment exercises leave room for opportunistic behavior by the subjects under evaluation. *Journal of Informetrics*, *13*, 830–840.
- Abramo, G., D'Angelo, C. A., & Grilli, L. (2021). The effects of citation-based research evaluation schemes on self-citation behavior. *Journal of Informetrics*, *15*(4), 101204. [abs/2102.05358](https://doi.org/10.1002/inf.1204).
- Académie des Sciences, Leopoldina, & Royal Society. (2017). Statement by three national academies on good practice in the evaluation of researchers and research programmes. *Vestnik Rossijskoj Akademii Nauk*, *88*(11), 979–981.
- Archambault, É., et al. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, *68*(3), 329–342.
- Baccini, A., De Nicolao, G., & Petrovich, E. (2019). Citation gaming induced by bibliometric evaluation: A country-level comparative analysis. *PLoS One*, *14*(9), e0221212.
- Biagioli, M., Kenney, M., Martin, B. R., & Walsh, J. P. (2019). Academic misconduct, misrepresentation and gaming: A reassessment. *Research Policy*, *48*(2), 401–413.
- Bonaccorsi, A. (2015). *La Valutazione Possibile: Teoria e Pratica della Valutazione della Ricerca*. Il Mulino.
- Borner, K. (2010). *Atlas of science: Visualizing what we know*. MIT Press.
- Boyack, K. W., Newman, D., Duhon, R. J., Klavans, R., Patek, M., Biberstine, J. R., Schijvenaars, B., Skupin, A., Ma, N., & Borner, K. (2011). Clustering more than two million biomedical publications: Comparing the accuracies of nine text-based similarity approaches. *PLoS One*, *6*(3), 1–11.
- Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation*, *12*, 39–46.
- Castano, S., Ferrara, A., & Montanelli, S. (2018). Topic summary views for exploration of large scholarly datasets. *Journal on Data Semantics*, *7*, 155–170.
- Castellani, T., Pontecorvo, E., & Valente, A. (2016). Epistemic consequences of Bibliometrics-based evaluation: Insights from the scientific community. *Social Epistemology*, *30*(4), 398–419.
- Cedrini, M., & Fontana, M. (2018). Just another niche in the wall? How specialization is changing the face of mainstream economics. *Cambridge Journal of Economics*, *42*(2), 427–451.
- Checchi, D., Ciolfi, A., De Fraja, G., Mazzotta, I., & Verzillo, S. (2021). Have you read this? An empirical comparison of the British REF peer review and the Italian VQR bibliometric algorithm. *Economica*, *88*(352), 1107–1129. <https://doi.org/10.1111/ecca.12373>
- Colander, D., Holt, R., & Rosser, B. (2004). The changing face of mainstream economics. *Review of Political Economy*, *16*(4), 485–499.
- Dahler-Larsen, P. (2014). Constitutive effects of performance indicators: Getting beyond unintended consequences. *Public Management Review*, *16*(7), 969–986.



- de Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2015). Evaluation practices and effects of indicator use — A literature review. *Research Evaluation*, 25(2), 161–169.
- Feenstra, R. A., & Lopez-Cozar, E. D. (2021). The footprint of a metrics-based research evaluation system on Spanish philosophical scholarship: an analysis of researchers' perceptions. *ArXiv*. abs/2103.11987. <https://doi.org/10.48550/arXiv.2103.11987>
- Felt, U. (Ed.). (2009). *Knowing and living in academic research: Convergences and heterogeneity in research cultures in the European context*. Institute of Sociology of the Academy of Sciences of the Czech Republic.
- Fochler, M., Felt, U., & Muller, R. (2016). Unsustainable growth, hyper-competition, and worth in life science research: Narrowing evaluative repertoires in doctoral and postdoctoral scientists' work and lives. *Minerva*, 54(2), 175–200.
- Gasparyan, A. Y., Nurmashev, B., Udovik, E. E., Koroleva, A. M., & Kitas, G. D. (2017). Predatory publishing is a threat to non-mainstream science. *Journal of Korean Medical Science*, 32(5), 713.
- Geuna, A., & Martin, B. R. (2003). University research evaluation and funding: An international comparison. *Minerva*, 41(4), 277–304.
- Geuna, A., & Piolatto, M. (2016). Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while). *Research Policy*, 45(1), 260–271.
- Glenisson, P., Glanzel, W., & Persson, O. (2005). Combining Full-text Analysis and Bibliometric Indicators. a Pilot Study. *Scientometrics*, 63(1), 163–180.
- Gottfredson, L. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1), 13–23.
- Heinze, T. (2013). Creative accomplishments in science: Definition, theoretical considerations, examples from science history, and bibliometric findings. *Scientometrics*, 95(3), 927–940.
- Ivani, S. (2019). What we (should) talk about when we talk about fruitfulness. *European Journal for Philosophy of Science*, 9(1), 4.
- Katzav, J., & Vaesen, K. (2017). Pluralism and peer review in philosophy. *Philosophers' Imprint*, 17(19), 1–20.
- Kuhn, T. S. (1996). *The Structure of Scientific Revolutions* (3rd ed.). University of Chicago Press.
- Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge University Press.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- Moed, H. F. (2017). *Applied evaluative Informetrics*. Springer International Publishing, Cham.
- Muller, R., & de Rijcke, S. (2017). Thinking with indicators. Exploring the Epistemic Impacts of Academic Performance Indicators in the Life Sciences. *Research Evaluation*, 26(3), 157–168.
- Nichols, L. (2014). A topic model approach to measuring Interdisciplinarity at the National Science Foundation. *Scientometrics*, 100(3), 741–754.
- Scarpa, F., Bianco, V., & Tagliafico, L. A. (2018). The impact of the National Assessment Exercises on self-citation rate and publication venue: An empirical investigation on the engineering academic sector in Italy. *Scientometrics*, 117(2), 997–1022.
- Seeber, M., Cattaneo, M., Meoli, M., & Malighetti, P. (2019). Self-citations as strategic response to the use of metrics for career decisions. *Research Policy*, 48(2), 478–491.
- Smelser, N. J., & Baltes, P. B. (Eds.). (2001). *International Encyclopedia of the Social & Behavioral Sciences* (Vol. 11). Elsevier.
- Suominen, A., & Toivanen, H. (2016). Map of science with topic modeling: Comparison of unsupervised learning and human-assigned subject classification. *Journal of the Association for Information Science and Technology*, 67(10), 2464–2476.
- Talley, E. M., et al. (2011). Database of NIH Grants using machine-learned categories and graphical clustering. *Nature Methods*, 8(6), 443–444.
- Viola, M. (2018). Evaluation of research(ers) and its threat to epistemic pluralisms. *European Journal of Analytic Philosophy*, 13(2), 55–78.

Yan, E., Ding, Y., Milojevic, S., & Sugimoto, C. R. (2012). Topics in dynamic research communities: An exploratory study for the field of information retrieval. *Journal of Informetrics*, 6(1), 140–153.

**Alfio Ferrara Alfio Ferrara** (Ph.D., University of Milan) is a Professor of Computer Science at the University of Milan. His research interests are focused on data science methods for natural language processing, information retrieval, and text mining.

**Corinna Ghirelli Corinna Ghirelli** (Ph.D., University of Ghent) is a research economist at the Bank of Spain. Her research interests are applied econometrics, labour economics, policy evaluation, and textual analysis.

**Stefano Montanelli Stefano Montanelli** (Ph.D., University of Milan) is an Associate Professor at the University of Milan. His main research interests include semantic web, data matching, web data classification and summarisation, and crowd-collaborative data management.

**Eugenio Petrovich Eugenio Petrovich** (Ph.D., University of Milan) is a post-doctoral researcher at the University of Siena. He works on scientometrics and quantitative science studies.

**Silvia Salini Silvia Salini** (Ph.D., University of Milano-Bicocca) is an Associate Professor of Statistics at the University of Milan. Her main research interests focus on statistical models for social science, multivariate statistics, statistical learning methods, robust statistics, and scientometrics.

**Stefano Verzillo Stefano Verzillo** (Ph.D., University of Milan) is a Senior Research Scientist at the Joint Research Centre of the European Commission. His research interests are education, labour and health economics, and evaluation of public policies.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.



**Part V**  
**Research Quality and Impact on Teaching**

# The Relationship Between Teaching and Research in the Italian University System



Maria Rosaria Carillo, Alessandro Sapio, and Tiziana Venittelli

**Abstract** We study the relationship between the quality of research and teaching in the Italian university system, at the study program level. We run a cross-sectional econometric analysis by using a very rich dataset collected by the Italian National Agency for the Evaluation of Universities and Research Institutes on the BA and MA-level degrees of all universities in Italy in the academic year 2016/2017. We find that a positive relationship between teaching quality and research performance emerges if we take account of yardstick competition among study programs belonging to the same department. Indeed, previous theoretical results suggest that, despite the individual trade-off between teaching and research faced by individual academics, in multi-unit universities adopting a budget sharing rule based on both research performance and number of students, the negative relation between teaching and research is reduced or even completely counterbalanced. We find a confirmation of this hypothesis by proxying yardstick competition with the number of study programs activated per department. However, the teaching–research relationship is positive and stronger where study programs are relatively few and immediately comparable by the department managers. Such results emerge more strongly in MA-level degrees, where teaching is more aligned with individual research interests.

---

We are grateful to the National Agency for the Evaluation of the University and Research Systems (ANVUR) for providing us with the data (in anonymized form) and assistance. We also thank seminar participants at the 2019 ANVUR Workshop (Rome, 12 November 2019) for useful comments and suggestions. We acknowledge funding from ANVUR (III Concorso Pubblico di Idee di Ricerca). The usual disclaimer applies.

---

M. R. Carillo (✉) · A. Sapio  
University of Naples Parthenope, Napoli, Italy  
e-mail: [carillo@uniparthenope.it](mailto:carillo@uniparthenope.it); [alessandro.sapio@uniparthenope.it](mailto:alessandro.sapio@uniparthenope.it)

T. Venittelli  
University of Naples Federico II, Napoli, Italy  
e-mail: [tiziana.venittelli@unina.it](mailto:tiziana.venittelli@unina.it)

**Keywords** Teaching quality · Degree program · Research quality · University

**JEL Classification** I20

## 1 Introduction

In recent years, a substantial number of universities around the world have become increasingly research-oriented. Most universities adopt reward systems that favour academics with high-ranking publications and guarantee career prospects to those with a high productivity in research, by reserving only marginal attention to the teaching effectiveness (ter Bogt & Scapens 2012; Parker 2012; Douglas 2013; Cadez et al. 2017). Nevertheless, contrary to what seems to emerge from such orientation, universities are interested also in high-quality teaching, since both research and teaching are leading missions for them. Hence, it is important to understand what consequences a reward system so skewed towards research may have on the quality of teaching. In fact, if the two activities are substitutes, a reward system based mainly on research might reduce the quality of teaching. The contrary happens if the two activities are complements: in this case, rewarding research allows also the teaching quality to rise.

Although this is a crucial issue for the university system, the literature has not reached a wide consensus on the nature and sign of the relationship between research and teaching. On the one hand, there are those who claim that the relationship is positive because the abilities in running the two activities are complementary, since excellent researchers may also provide high-quality teaching, being people with deeper insights on scientific topics that they transfer through teaching (Braxton 1996; Sullivan 1996; Rodriguez & Rubio 2016). On the other hand, there are those who emphasize substitutability, arguing that the abilities in teaching and research are independent and both activities need time and effort which are limited resources for researchers. As a consequence, an incentive scheme more skewed towards research might drastically lower the time and effort that individual researchers dedicate to teaching activity as well as its effectiveness (Barnett 1992; Marsh 1987; Ramsden & Moses 1992; Parker 2012). Empirical analysis run on the question has not solved the puzzle. Several papers find a positive relationship between research and teaching quality, others a negative or even a null one. Moreover, results show a high variability since they change according to the level of degree programs, the proxies used to measure quality of teaching and research, and the variables capturing the context within which the two activities are performed.<sup>1</sup>

The large variability in the empirical results could also derive from the fact that, although the relationship between teaching and research depends both on the behaviour of academic professors and on the organization of the universities and

---

<sup>1</sup> See Marsh and Hattie (2002) and Qamar uz Zaman (2004), for two very comprehensive surveys.

departments within which the two activities are carried out, the aims of the professors and universities do not completely coincide, and, under some circumstances, there might be even a conflict between them. While universities are multitasking institutions, for which teaching and research are complementary activities, this is not necessarily true for a single researcher, for whom the two activities are more likely substitutes (Barnett 1992; Hattie & Marsh 1996; Cadez et al. 2017). In fact, universities derive funds from tuition fees and research funds, while researchers derive their wages, tenure, and scientific reputation mainly by research productivity. Since both teaching and research require effort and time which are limited resources, they might be perceived as substitutes by the individual researcher. This framework can be further complicated by the fact that between universities and academic professors there is a principal–agent relation (Gautier & Wauthy 2007; Bak & Kim 2015; De Philippis 2020). While universities can observe research productivity, they cannot perfectly observe teaching effectiveness. This implies that if universities adopt an incentive scheme based on research performance, in order to solve the agency problem, this strategy may have an unintended detrimental effect on the teaching quality, since professors would choose to put more effort on research activity by free riding on teaching activity, which is perceived as a sort of public good (Gautier & Wauthy 2007; Payne & Roberts 2010).

In this chapter, we study the relationship between research and teaching in the Italian university system at department and study program level. We will explore the role of the department organization in reducing the detrimental effects on teaching quality, which derive from an incentive scheme based on research performance, adopted in order to solve the agency relation between professors and departments. In particular, we will consider the multi-unit nature of departments, which in Italy typically supply different study programs both at bachelor and at master level, and the fact that university departments are financed by funds received from the government both for their research productivity and for the number of students enrolled in their programs. These characteristics of the institutional context are conducive to a yardstick competition between study programs, which reduces the incentive for individual professors to free ride on teaching activities and also the trade-off between teaching and research.

In our empirical analysis, we exploit a very rich dataset collected by the Italian National Agency for the Evaluation of Universities and Research Institutes (ANVUR) and providing information on almost five thousand degree programs belonging to all Italian universities, public and private, telematic or traditional. To measure the quality of teaching for each study program, we consider objective as well as subjective measures. Objective measures include the initial efficacy of the study programs and the regularity of study paths. As subjective indicators, we consider the graduates' satisfaction with the degree program they graduated from. All measures are based on data collected for the 2016–2017 academic year. We measure the academic research performance by using an indicator of quality rather than quantity: the R indicator provided by ANVUR, calculated at level of study program and department. More precisely, the R indicator is the average score that researchers, who teach in a given study program in the 2016–2017 academic

year, received during the 2011–2014 Italian Research Assessment, normalized by scientific macro-field.

From our analysis, a positive teaching–research quality relationship emerges rather clearly when we carefully represent the nexus between research and teaching in the framework of yardstick competition among study programs, proxied by the number of study programs activated within the same department. In particular, we interact research performance with our proxy for yardstick competition and find its coefficient to be positive for degree programs facing relatively few competitors in the same department and negative if the number of competing degree programs is larger, such that free-riding behaviours are harder to detect. Less clear is instead the relationship between research quality and teaching in BA-level programs, where topics are typically far from research interests of the faculty. Other interesting results regard the student–instructor ratio: in study programs with a below-median student–instructor ratio, research quality and teaching quality are more strongly associated, likely because classes with lower students allow to relax the time and energy constraints faced by professors. However, we find that when the student–instructor ratio is below median, additional students per instructor tend to weaken the positive relationship between research and teaching, as shown by the coefficient estimates of an interaction term. This can be explained by the fact that, when departments’ budget is based on both research productivity and the number of students, smaller classes imply also a lower amount of funds available to departments, which limits the scope of winner-picking in research activity and rises incentives for teaching. Results about control variables, accounting for heterogeneity in terms of average instructors’ age, qualification, gender composition of the faculty, research funding, and degree program internationalization, follow expectations.

The empirical literature analysing the relationship between teaching and research at the university (or department) level for the Italian case is quite scant. In this regard, we mention the contributions of Sylos Labini and Zinovyeva (2011) and Braga et al. (2014), which focus, respectively, on the teaching performance of the departments of all the Italian universities and on the teaching effectiveness of academic staff of some degree programs of Bocconi University. Results of both seem to suggest the existence of a weak positive correlation between the two phenomena. The existence of a weak positive correlation between teaching and research is also confirmed in De Philippis (2020), who analyses the case of Bocconi University. By comparing the results before and after the application of an incentive scheme more biased towards research, she finds evidence of a negative effect of research productivity on the teaching effectiveness at individual level, but a positive effect at the university level due to a composition effect. Although these results are particularly interesting, they are not completely transferable to the whole Italian university system.<sup>2</sup> Hence, a wider empirical analysis is needed in order to obtain more general results. This chapter may contribute to fill this gap, which is important

---

<sup>2</sup> Bocconi University is a private university that has different incentive scheme and slightly different recruitment rules.

given also the great interest showed by policy makers in setting policies aimed at enhancing the effectiveness of teaching and the productivity of research in Italian university system.

The remainder of the chapter is organized as follows. Section 2 draws on the existing literature to overview the main insights on the relationship between teaching and research at the department level. Section 3 focuses on the Italian institutional context. Section 4 describes the data. Section 5 presents the econometric model. Section 6 includes the main results. Finally, Sect. 7 concludes.

## 2 The Relationship Between Teaching and Research Within Higher Education Institutions

The relationship between teaching and research has long been debated in the literature on scientific productivity. The topic, however, has been analysed mainly at the single researcher level, by paying less attention to the organizations within which both activities are carried out, such as departments and faculties (Marsh & Hattie 2002). Moreover, the theoretical justifications provided for the existence of a negative or positive relationship between the two activities are mainly at the individual level, while factors that explain the existence of a relationship between the two activities at the department level have been much less analysed.

We believe that the most appropriate unit of analysis to study such relationship is the department and/or the university, for several reasons. First, the research activity is often conducted by teams which are composed by members of the same department.<sup>3</sup> Second, there is no reason to expect that what holds for individual academics holds in the aggregate, since individual- and organization-level goals do not fully overlap. Moreover, the relationship between professors and universities can be understood in the principal–agent framework, where universities are the principal. As an implication, the incentive scheme adopted by universities is of crucial importance for determining whether the two activities are complements or substitutes, for which the relation is, respectively, positive or negative. Third, universities and departments may affect the relationship between teaching and research through their organization of human resources by favouring the specialization among faculty members, or the emergence of positive externalities through different forms of collaborations among members of departments (Båkera & Goodallb 2020; Bradford et al. 2014; Carillo et al. 2013), or again by adopting a type of organization which reduces (or support) the administrative tasks carried out by professors. All these aspects may reduce the trade-off between the two activities.

---

<sup>3</sup> In some research fields, which make larger use of laboratories and expensive equipments, the proportion of within-department research collaborations is very high.



## ***2.1 The Multitasking Nature of Universities and the Principal–Agent Relation Between University and Professors***

As already stated, universities are multitasking institutions for which teaching and research are complements since the universities' budget is composed by both research funds and students' fees. This is not necessarily true for researchers individually, who have to allocate time and effort between the two tasks, which makes the two activities substitutes rather than complements. The agency problem further complicates the framework, since the non-observability of teaching effectiveness induces universities to adopt an incentive scheme biased towards research, by incentivizing professors to free ride on teaching activity. In fact, as predicted by Holmstrom and Milgrom (1991), if a multitasking principal adopts a performance-based scheme, this drives the agents to reallocate time and resources towards the more rewarding task, at the expense of the less rewarding one. Hence, the final results may radically change according to the incentive scheme adopted by universities and whether universities are able to counterbalance the free riding of professors through other aspects of their organization.

Recently, several papers have adopted this framework in analysing the relation between research and teaching, by asking how multitasking universities may solve the principal–agent problem. Gautier and Wauthy (2007) assume that departments (or universities) are multi-unit organizations, and budget allocation among units depends both on the number of students and on research productivity. The authors show that such allocation rule induces yardstick competition among units, which reduces the substitutability between research and teaching effectiveness. In particular, yardstick competition among the different units reduces the incentive to free ride and rises the complementarity between teaching and research if the number of units is not too high. In the same line is the paper by De Philippis (2020), who also studies the allocation of professors' efforts between research and teaching when there is an agency problem and universities adopt an incentive scheme biased towards research. In particular, she focuses on the relationship between research and teaching abilities in order to assess the effect of such an incentive scheme on the relation between the two activities. She shows that in such a framework, the degree of substitutability between the two tasks arises because when the reward is biased towards research, the cost of effort in teaching is higher for academics who are more involved in research. However, the negative effect can be counterbalanced by a composition effect which occurs if the ability for teaching is complement with the ability for research. In this case, incentives highly skewed towards research attract a supply of academics with high ability, thus counterbalancing the negative effect at individual level. Also, Bak and Kim (2015) adopt the multitasking theory for analysing the research and teaching relationship in the case of the Korean university system. The authors find that in a context where the incentive scheme is more skewed towards research, there is a reduction in teaching effectiveness. However, the negative effect is higher for

undergraduate programs, for which the substitutability between the two tasks for individual researchers is higher.

An incentive structure biased towards research may reduce teaching effectiveness also by modifying the type of research. If it incentivizes the quantity rather than quality of research, the possibility of transferring new scientific knowledge to students is reduced, and hence the sign of the correlation is more likely negative (Shin 2011).

Finally, several authors suggest that not only explicit but also implicit rewards are important in shaping the relation between teaching and research but implicit rewards as well (Marsh & Hattie 2002; Carillo & Papagni 2014). A departmental ethos that gives more emphasis on research (or on teaching) could lead academics to place greater importance to research (or to teaching). If colleagues are particularly committed to research or teaching, then it is more likely that there are intrinsic rewards and higher reputation for excellence in that activity. Ramsden and Moses (1992) suggested that “high departments are populated by staff who are on average less effective teachers and vice-versa” (p. 287).

## 2.2 *Specialization*

At the department level, the way in which tasks and duties are allocated among the department members affects the time required for the implementation of teaching activities. For example, the involvement of PhD students and research assistants in teaching activities can improve the quality of teaching and at the same time relax the time constraints faced by senior scholars. A division of labour between senior and junior academic members, which gives more administrative duties related to teaching activities to senior academics, can also achieve the same results. Bäkera and Goodallb (2020) find that in departments where junior members have a low administrative burden, their research activity improves and there is less substitutability between the two tasks at the individual level. Also, Garcia-Gallego et al. (2015), by exploring the case of Castellona University in Spain, ask whether the specialization arising within the university for which some professors specialize more in administrative and teaching duties may reduce the substitutability between the two activities. They find that all phenomena arising within departments which increase specialization and collaborations among their members give rise to a positive correlation between research and teaching at the department level.

## 2.3 *Positive Scientific Externalities*

Another important factor is the existence of positive externalities generated by the scientific activity of the members of the same department. Positive scientific externalities within the department can spread through scientific collaborations

between members of the same department, the organization and participation in seminars, participation in funded research projects, or even just through the exchange of ideas and information sharing (Carillo et al. 2013; Carillo & Papagni 2005). This implies that in an environment with high scientific externalities, it is possible to obtain a certain level of scientific production while investing less time and resources, which in turn improves the time and resource constraints on individuals and the trade-off between teaching and research activities.

## ***2.4 The Level of Education***

Another important feature of universities and departments that affect the relationship between teaching and research is the level of education they offer. Several authors (Brew & Boud 1995; Griffiths 2004; Brew 1999; Healey 2005; Palali et al. 2018) argue that undergraduate university programs offer less space for transferring the new frontier knowledge into teaching, while in more advanced education levels, such as masters or doctoral programs, this transfer is wider if not a necessary part of the teaching activity. Brew and Boud 1995 and Griffiths 2004 have focused on how departments define teaching activities: when they define it as a “student learning process,” research is closely related to teaching. Obviously, this definition is more suitable for higher level education. This result is confirmed by Palali et al. (2018). The authors run an empirical analysis on professors in the Netherlands, to find a positive relationship in case of master students and for students in the last year of their bachelor degree, while a negative one for lower degrees. De Philippis (2020) finds a similar result for Bocconi University. Hence, when professors can bring their research into class and disseminate it to students, the substitutability between teaching and research does not apply.

## ***2.5 Research Fields***

Finally, the nexus between the two activities varies according to the disciplines that characterize a department or a faculty, because of differences in epistemology, research methods, and types of academic cultures existing among them. Shin (2011) and Shin and Kim (2017) in empirical papers on the Korean university system find that in hard science departments the relationship between teaching and research is null or even negative in low-level education, while it becomes weakly positive in high-level ones. The contrary happens in social and humanities sciences. The authors argue that this can derive from the fact that research in hard sciences produces more articles in international journals, while humanities and social sciences produce more books and articles in domestic journals. These characteristics make easier for the humanities and social sciences to transfer the new knowledge in undergraduate programs. The contrary happens in higher levels

of education, where students are more accustomed with formal reasoning and have good knowledge of foreign languages: in this case, hard sciences can more easily transfer the new knowledge to students. Walstad and Allgood (2005) for example find that US Economic fields are too much aimed and too rewarding towards research activity, if compared to fields in Business, Engineering, Mathematics, and Statistics.

### 3 The Italian University System

The Italian university system has been profoundly transformed after the Gelmini reform of the university system implemented in the 2010 and the introduction of the National Scientific Qualification (*Abilitazione Scientifica Nazionale—ASN*), which jointly characterize it as a system wherein public funding is allocated to universities mostly based on teaching indicators, but individual careers depend on research performance.

After the reform, universities are organized in departments, which have responsibilities on research, teaching, and the related recruitment, within the budget allocated by the university. Each department can manage one or more degree programs, including both BA-level and MA-level programs. Each program is managed by a council, including a number of professors affiliated to the department or to other departments. Such professors are termed reference professors and can take this role only in one degree program.<sup>4</sup> The department is responsible for proposing to the university the structure of the degree programs, namely, the list of subjects, their weights in terms of ECTS, and the allocation of instructors among subjects.

The enrolment fees are collected by the university and contribute to its budget, along with transfers from the Ministry of University and Research (MUR). Such transfers are based on the number of enrolled students, as well as on teaching quality indicators and, for a limited share, on research assessment outcomes performed by ANVUR. Part of the budget is used by universities for recruitment of academic staff. This can amount to new recruits or to upgrading the position of the existing academic staff.

Academic staff members can apply for career upgrades within their university, provided they have obtained the National Scientific Qualification to that position in the relevant academic field. Qualification is awarded by national committees, by considering above all the scientific quality of the publications submitted by candidates (originality, impact, editorial collocation, coherence with the field), provided that candidates satisfy certain threshold values in terms of a number of publications. Some teaching-related aspects are also taken into account, such

---

<sup>4</sup> Reference professors are a subset of the instructors who teach in the degree program. One of the reference professors is elected as coordinator by her/his peers.

as teaching fellowships in foreign universities or PhD board membership, but their weight in the evaluation is very minor. Significantly, no indicator about undergraduate teaching is considered.

To sum up, for the purposes of studying the teaching–research relationship, one can summarize the Italian university system as follows. Universities collect enrolment fees from students and transfers from the ministry and use part of them to finance new recruits or career upgrades. Though, candidates for academic positions compete in terms of research performance. Hence, in the aggregate, opportunities for academic careers depend on the ability of universities to attract students by carefully balancing tuition fees and teaching quality; however, individual opportunities do not depend on teaching efforts and could in fact be hampered by allocating too much effort away from research.

It is worth noting that the multi-unit structure of (Italian) universities adds a further layer of incentives that may affect the teaching–research trade-off. Universities can allocate their funding for recruitment among departments and degree programs based on their relative performances in attracting students. Degree programs with more students and/or with students who report better satisfaction or job market placement may be allocated larger shares of the recruitment budget. Competition among degree programs, based on better teaching indicators, is what provides the best researchers with larger opportunities for their career concerns. But there may not be enough incentives for the individual academic to improve his/her teaching performance since positions are awarded based on research quality.

### ***3.1 The Italian Evaluation of Research Quality***

The Italian assessment of research quality (VQR) has been carried out by ANVUR, on behalf of MUR, since 2011, to evaluate the scientific production of Italian universities and departments. Researchers have to submit a limited number of research papers, presumably their best papers,<sup>5</sup> which are evaluated by a panel of experts, selected by ANVUR for each macroarea of scientific research. The evaluation process is based on two evaluation methods: bibliometric analysis, based on bibliometric indicators (i.e. citations of the paper and the impact factor of the journal in which the paper is published) and informed peer-review evaluation by external experts, named by the panel. Each product receives a score ranging from 1 (excellent) to 0.7 (good), 0.4 (fair), 0.1 (acceptable), and 0 (limited or inadmissible). Hence, the research productivity is valued in terms of quality rather than in terms of quantity.

The contribution of each researcher to the scientific performance of the university is significant, given that the results of the research evaluation contribute to determining the share of the fund that MUR allocates to each University. However,

---

<sup>5</sup> Eligible products are: journal articles, books, book chapters, conference proceedings, etc.

only a small part of this fund depends on the results of VQR. In 2017, after the publication of the VQR results that referred to the evaluation of scientific production in the period 2011–2014, this share represented 80% of the “reward fund” (*quota premiale*), which in turn consisted of 23% of the ordinary fund.

## 4 Data and Variables

For the purposes of this research, we have obtained by ANVUR data on 4858 degree programs activated by all public and private Italian universities in the 2016–2017 academic year. We consider also programs provided by online universities, as they are exposed to the same hiring rules and incentives as all other universities in Italy.

The dataset includes a number of variables that proxy for quality of research and teaching in Italian universities. In particular, in order to measure the quality of research performed by members of a study program, we rely on a variable that represents the key indicator within the 2011–2014 Italian Research Assessment, the so-called  $R$  indicator.<sup>6</sup> More specifically, the  $R$  indicator is calculated as the ratio between the average grade of the expected products by a given university in a certain scientific area and the average grade received by all the products of the area; the aggregate measure for the degree program is computed as the weighted sum of the area-wise  $R$  indicators, using the number of expected products of each area as weights.

Indicating with  $v_{i,j,k}$  the sum of the evaluations of the  $k$ -th degree program of the  $i$ -th university in the  $j$ -th area and with  $n_{i,j,k}$  the number of products expected for the VQR of the  $k$ -th degree program of university  $i$  in the  $j$ -th area and defining as  $q_{i,j,k}$  the share of professors belonging to area  $j$  who teach in the  $k$ -th degree, we have

$$R_{ik} = \sum_{j=1}^{N_j} q_{i,j,k} \frac{\frac{v_{i,j,k}}{n_{i,j,k}}}{\frac{\sum_{i=1}^{N_i} v_{i,j}}{N_j}} = \sum_{j=1}^{N_j} q_{i,j,k} \frac{\frac{v_{i,j,k}}{n_{i,j,k}}}{\frac{V_j}{N_j}}, \quad (1)$$

where  $N_i$  and  $N_j$  are the cardinalities of, respectively, universities and areas.

This indicator captures the relative research performance of researchers teaching in a given degree program, with respect to research performances in the scientific areas involved in the degree program. Values below (above) 1 indicate a below-average (above-average) research performance. We recall that individual grades, which make up the sum  $v_{i,j,k}$  for each degree/university/field combination, range

---

<sup>6</sup> On our request, ANVUR has computed the  $R$  indicator by study program, i.e. with reference to the researchers who teach in a given study program in the 2016–2017 academic year, and normalized by scientific macro-field (CUN areas).

from 1 (excellent) to 0.7 (good), 0.4 (*discreto*), 0.1 (acceptable), 0 (limited or inadmissible).

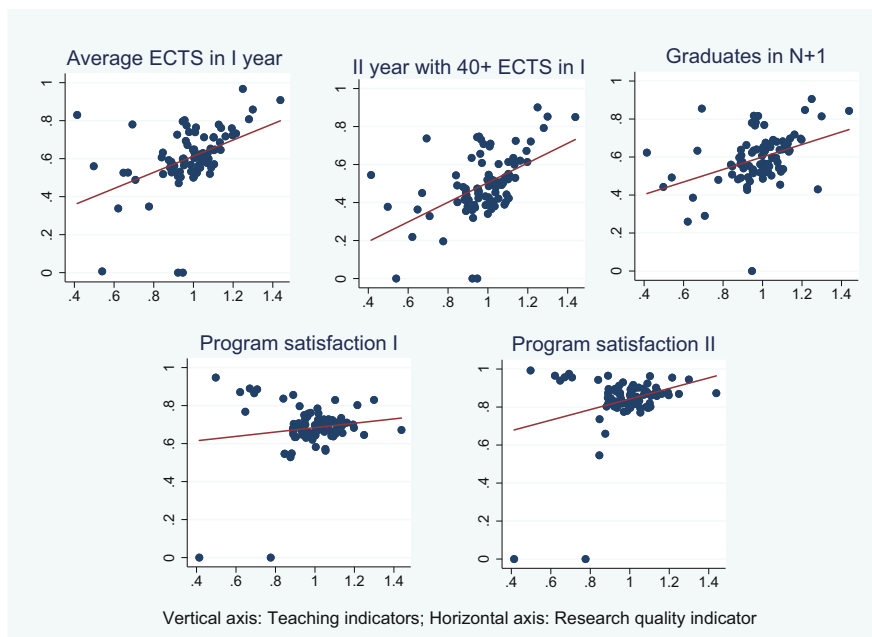
Teaching quality indicators that we consider measure both the academic performance of students and the satisfaction of graduates. Among indicators of students' academic performance, we use the percentage of credits obtained in the first year with respect to the total number of credits to be obtained in the first year, the percentage of students who have obtained at least 40 credits in the first year and then enrol to the second year, and finally, the percentage of freshmen who graduate not later than one year after the ordinary duration of the study program. The first two indicators would capture the initial efficacy of study programs, i.e. the ability of university teaching staff to allow students a fairly swift transition from the first year courses, when students apprehend the basics, to second and third year courses that are more specifically aimed at preparing students for the job market. If students struggle to pass first year exams, it may as well be due to poor selection of freshmen, to ineffective organization of first year courses,<sup>7</sup> or to a teaching staff who set very high standards. For these reasons, students may decide to transfer to another university where they expect to find a better match, or to give up university at all. In both cases, one may argue that the university has failed in its teaching mission. The third indicator of students' academic performance would capture the regularity of study paths since it is achieved when students complete their curriculum in due time. Such indicator refers to cohorts of students who have managed to pass first year exams. However, some students may still find difficulties in passing second and third year exams, which may require the application of basic notions learned in the first year, as well as learning more advanced concepts and analytical tools. Policy makers tend to have a negative assessment of universities in which students struggle to graduate in time, as this may prevent an effective school-to-work transition. On the other hand, students may as well take longer to graduate because they engage in activities that improve their chances of a successful school-to-work transition, such as internships or advanced dissertation topics.

A final category of teaching quality indicators concerns the satisfaction of graduates. We consider the percentage of graduates who would enrol again in the same degree program and the percentage of graduates who are overall satisfied about their degree program.<sup>8</sup> Students may be satisfied about their university choice for several reasons. Perhaps the straightest reason concerns the job market outcomes. Students who quickly find jobs that correspond to their labour market expectations or ambitions are supposedly more satisfied than average. Yet, satisfaction may originate from having attended classes given by highly skilled professors, from spending time in a well-organized university environment, or from the sheer interest

---

<sup>7</sup> Such as lack of clarity in prerequisites and evaluation criteria, inadequate balance between teaching materials and teaching hours, obsolete teaching methods, mismatch between topics and skills of the instructors, and insufficient availability of tutors.

<sup>8</sup> These indicators are provided by the AlmaLaurea Interuniversity Consortium. For non-consortium universities, information on program satisfaction is requested directly by ANVUR to each university.



**Fig. 1** Correlation between teaching indicators and research quality

of the discipline—regardless of labour market outcomes. All teaching indicators are provided by ANVUR in respect of AVA (Autovalutazione—Valutazione periodica—Accreditamento) obligations on universities and refer to academic year 2016–2017.

In Fig. 1, we present the relationship between teaching indicators and the  $R$  measure of program-level research quality. The three scatter plots in the upper panel of Fig. 1 refer to students' academic performance and show a positive association between the research performance of teaching staff and all indicators of teaching quality, i.e. the average number of credits obtained in the first year, the percentage of students enrolled in the second year with 40 credits in the first year and the percentage of students who graduate within one year by the legal duration of the study program. Instead, we observe no association or even a weakly positive association between program research quality and the satisfaction of graduates according to the scatter plots at the bottom of Fig. 1, which refer to the percentage of graduates who would enrol again in the same degree program (*Program satisfaction I*) and the percentage of graduates who are overall satisfied about their degree program (*Program satisfaction II*).

As suggested in Sect. 2 when summarizing the theoretical insights on the research–teaching relationship, it is essential to take into account the organization of departments. In particular, we have to consider in our case the multi-unit nature of departments in Italy, since generally with few exceptions, departments may house



more than one degree program, often at least a BA-level and a MA-level degree. The degree programs organize the teaching activity and establish the actions to be taken in order to improve the teaching quality indicators (e.g. tutoring of students who struggle to pass exams, recommendations in order to have syllabi that match the students expectations, organization of internships). However, degree programs are designed by university departments, which decide their goals, modules, as well as the allocation of teaching personnel among them.

To measure the inner organization of departments, we use different variables. First, the number of professors allocated to each degree program. Second, the ratio between the number of students over the number of instructors, which indicates how much relevant is students' fees in the budget of department but also the effort required by teaching activity. Third, the number of degree programs per department, which captures the yardstick competition arising within department given that degree programs compete each other for obtaining funds and resources from department. Moreover, we also include two further ANVUR indicators that are study program-specific. These are the percentage of professors who teach in basic subjects and are at the same time reference professors for the study program (i.e. directly engaged in the management of the study program) and the percentage of credits obtained abroad by students. High values of the former may signal that the management strategy defined by the professors who coordinate the program directly affects the process of basic knowledge acquisition by the students and the selectiveness of the program. The latter (credits abroad) can be seen as a proxy of the intrinsic motivation of students and of their income. Indeed, although students in international mobility receive a small scholarship, students coming from lower income families may not afford to pay for the full cost of a foreign stay. Typically, students who are less motivated will not apply to Erasmus programs.

Other important aspects are the shares of full and associate professors, the share of post-docs, and research funding per capita. The average teaching experience of the department professors (as proxied by their role) and the availability of younger colleagues who may help them carry out research and teaching tasks (post-docs) sound like useful control variables. The department staff composition tells something about the division of labour within a department, which may be a key driver of teaching quality, as well as about the pattern of intra-department externalities (see Sect. 2). Also, higher research funding per capita may alter the trade-off between teaching and research efforts, as it may be reflective of an incentive structure biased in favour of research, possibly to the detriment of teaching.

Finally, we control for some characteristics of individual professors such as the average age of professors and the share of women. Younger professors may master the most advanced methodological tools, yet they may lack experience. Women may face a tighter work–life constraint and therefore may have to choose between excelling in teaching and in research.

We consider also fixed effects. We include university dummies, to control for unobserved university-specific features that may affect performances.<sup>9</sup> We control for the level of education: BA-level degree (*laurea*), MA-level degree (*laurea magistrale*), and *laurea magistrale a ciclo unico* (a 5-year degree). Indeed, the knowledge base and motivation of students in different degree types change considerably (see Sect. 2): MA-level students are “better selected” and are interested in more applied topics. We also control for the geographical area (North, Centre, South), as the socio-economic differentials that characterize Italy may have an impact on students’ performances. In the South, with less infrastructures and lower per capita income, students may have less resources for their education and lower expectations about job opportunities and therefore may underperform even if their universities are well-organized and house highly skilled professors.

Moreover, our estimates will take account of the irreducible specificities of scientific areas, as discussed in Sect. 2.5, by including the degree type dummies (i.e. Economics, Humanities, Mathematics, Medicine, etc.) and performing estimates on area-specific subsamples (bibliometric vs. non-bibliometric areas). All control variables refer to the academic year 2016–2017.

Descriptive statistics for the variables considered in this study are displayed in Table 1. Means are computed for the whole sample (column 1) and by type of degree program (BA- and MA-level degrees, respectively, in columns 2 and 3). In column 4, we compute t-tests to verify if there are statistically significant differences between BA- and MA-level programs (column 4). With regard to ANVUR indicators on students’ academic performance, we see that, on average, students obtain about 60% of the required ECTS credits within the first academic year, while the percentages of those who progress to the second year with at least 40 credits are on average 49%; finally, about 61% of students graduate within one year beyond the legal duration of the study. According to column 4, there is a substantial difference between BA- and MA-level programs according to all the indicators of students’ performance, with MA-level degree students outperforming BA-level degree students by, respectively, 9.5%, 7.5%, and 23.5%, which confirms the higher ability of students who self-select in master programs vis-a-vis those who enrol in bachelor programs.

As for the ANVUR indicators on the satisfaction of graduates, Table 1 shows that the percentage of graduates declaring they would enrol again in the same programs (*Program satisfaction I*) or to be completely satisfied about the program they attended (*Program satisfaction II*) is relatively large, that is, 67% and 84%, respectively. The t-tests in column 4 highlight that the percentage of satisfied graduates is higher for MA-level degree students with respect to the former indicator only, while there is no statistically significant difference between BA- and MA-

---

<sup>9</sup> It is worth noting that the data supplied by ANVUR do not allow to identify the universities. We do not know the name and location of the universities in our sample. Hence, we are unable to include variables describing the socio-economic context (e.g. labour market conditions) neither the university reputation.

Table 1 Descriptive statistics

	(1)	(2)	(3)	(4)
	All programs Mean	BA-level Mean	MA-level—II years Mean	Diff. <i>t</i> -test
Average ECTS in I year	0.613	0.565	0.659	0.095***
II year with 40+ ECTS in I year	0.489	0.450	0.525	0.075***
Graduates in <i>N</i> + 1	0.608	0.505	0.741	0.235***
Program satisfaction I	0.669	0.647	0.692	0.045***
Program satisfaction II	0.839	0.839	0.834	-0.005
Research quality	1.017	0.990	1.046	0.056***
Students/instructors (degree program)	13.202	17.557	7.597	-9.960***
% of reference instructors in the degree program	0.902	0.928	0.860	-0.068***
ECTS abroad/total ECTS (degree program)	0.021	0.009	0.034	0.025***
Average age of instructors (degree program)	51.377	51.089	51.577	0.488***
% Female instructors (degree program)	0.380	0.398	0.358	-0.040***
# Instructors (degree program)	43.260	57.345	22.050	-35.296***
Full professor in department(%)	0.234	0.229	0.239	0.010***
Associate professor in department(%)	0.364	0.359	0.373	0.014***
Post-docs in department (%)	0.291	0.288	0.305	0.017*
Per capita research funds (department)	49958.843	49890.57	51772.816	1882.159
# of degree programs in department	8.00	8.64	7.56	-1.08***

level degree students in relation to the percentage of graduates, who are completely satisfied about the program.

On average, the R score at the degree program level, i.e. the research quality indicator, is about 1 and it is slightly larger for MA-level programs (1.046) than for BA-level programs (0.99). Looking at the other covariates at degree program level, we show that the student–instructor ratio is higher for BA-level than for MA-level programs (17.6 vs. 7.6), while the percentage of instructors teaching basic topics who are “reference professors” is almost 90% for both BA- and MA-level programs (93% and 86%, respectively). The percentage of female instructors is also slightly higher for the BA-level program case. Finally, the percentage of ECTS obtained abroad is very low and equal to 2.1% and less than 1% in the case of students enrolled in BA-level programs. This significant differences between the BA- and MA-level programs in most of the variables we consider in our analysis may signal that the teaching–research quality relationship could work differently depending on the degree type.

## 5 Econometric Model

The relationship between research quality and teaching quality in study programs is estimated through the following model:

$$\text{Teaching quality}_{i dk} = \alpha + \beta R_{i dk} + \gamma X_{i dk} + \theta Z_{i d} + \varepsilon_{i dk}, \quad (2)$$

where  $i$  denotes the generic university,  $d$  the department, and  $k$  the generic study program;  $\text{Teaching quality}_{i dk}$  is a teaching quality indicator for study program  $k$  in university  $i$  and in the department  $d$ ;  $R_{i dk}$  represents the research quality indicator based on the Research Assessment grades; and  $X_{i dk}$  is a matrix of control variables for study program  $k$  in university  $i$ , including also dummies accounting for fixed effects, and  $Z_{i d}$  is a matrix of control variables for department  $d$  in university  $i$ .  $\beta$  is our coefficient of interest. If positive, it testifies to a positive correlation between the research quality of the professors teaching in a study program and the performance of the study program according to teaching quality indicator (such as student academic performance and graduate satisfaction).

In some model specifications, we use a slightly different measure of research quality, i.e. the  $R$  indicator normalized by instructor-specific academic discipline rather than by academic field, which gives a more precise estimator of research quality because it captures relevant differences in research performance, which are field-specific.<sup>10</sup>

Some of the control variables at study program level and at department level are of particular interest since they can modify the relationship between teaching

---

<sup>10</sup> However, ANVUR provides such indicator only for MA-level degree programs; for this reason, we use the first indicator to explore the relationship for all programs.

and research; thus, in the subsequent section, we will use some of them in order to explore whether they are moderating factors for this relation. Indeed, the theoretical insights summarized in Sect. 2 suggest that, despite the existence of a trade-off between research and teaching efforts from the viewpoint of the individual academic, a positive correlation between teaching and research quality may arise at the study program level due to multi-unit nature of universities and departments. Thus, to account for the moderating role that some department characteristics may play, we estimate a model including interaction terms. In one model, the research performance indicator is interacted with the number of programs per department. In another model, the research performance indicator is interacted with the number of students per instructor at the study program level. Such interaction terms are meant to capture the effect on teaching quality of yardstick competition among programs within the same department and the effects of funds related to students' fees.<sup>11</sup>

In commenting these results, a special emphasis will also be put on the competition among teaching staff for career concerns, which is captured by the coefficient associated with the number of instructors. Indeed, the number of instructors can be seen also as a proxy for competition faced by academics, within their degree program, for potential upgrades. In degree programs with more instructors, we expect academics to focus more on research and less on teaching, on average, in order to win the competition for upgrades.<sup>12</sup> Thus, we expect a negative coefficient for the number of instructors. Such a negative effect due to career concerns may be weaker in MA-level programs, as in MA-level programs, academics typically teach topics that are closer to their research interests and they may rather prefer to reduce teaching efforts on more basic BA-level programs.

We finally compare the results obtained on subsamples of bibliometric and non-bibliometric fields.

## 6 Results

A first estimation exercise considers, as the dependent variable, proxies for teaching quality and students' progress, namely: average ECTS obtained in the first year, number of students enrolled in the second year after obtaining 40 or more ECTS in the first year, and number of graduates one year after the ordinary duration of the program. Table 2 collects these results. In detail, the first three columns of Table 2 focus on the average ECTS obtained by students in their first year. Column (1) includes estimates for a sample including all programs, whereas the following

---

<sup>11</sup> Actually, the student–instructor ratio may also be related to greater opportunities for peer interactions, both cooperative (e.g. exchange of information) and competitive (peer pressure) arising among students. Positive externalities from information and knowledge diffusion also spread more broadly in larger classes.

<sup>12</sup> In Italy, a number of the positions for associate and full professor are reserved for the internal staff of the university (art. 24, paragraph n. 6 of the law 240/2010).

**Table 2** The relationship between teaching effectiveness and research. The effect on student academic performance

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Average ECTS in 1 year		II year with 40+ ECTS in 1 year		Graduates in N + 1			
	All programs	BA-level	MA-level	All programs	BA-level	MA-level	BA-level	MA-level
Research quality	0.0301** [0.0134]	0.0264 [0.0217]	0.0364** [0.0175]	0.0294+ [0.0185]	0.0283 [0.0300]	0.0420* [0.0246]	-0.0137 [0.0151]	-0.0570** [0.0227]
# Instructors (degree program)	-0.0003*** [0.0000]	-0.0003*** [0.0000]	0.0004 [0.0003]	-0.0004*** [0.0001]	-0.0003*** [0.0001]	0.0004 [0.0004]	-0.0003*** [0.0001]	0.0003 [0.0003]
Students/instructors (degree program)	0.0001 [0.0005]	0.0002 [0.0007]	0.0004 [0.0007]	0.0002 [0.0006]	0.0000 [0.0008]	0.0015+ [0.0010]	0.0011+ [0.0007]	0.0024*** [0.0008]
# of degree programs in department	-0.0000 [0.0005]	-0.0005 [0.0007]	0.0017* [0.0009]	-0.0006 [0.0008]	-0.0010 [0.0012]	0.0021* [0.0013]	-0.0013** [0.0006]	-0.0015 [0.0011]
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.509	0.549	0.402	0.443	0.485	0.388	0.647	0.301
Observations	4173	2016	1870	4192	2022	1883	4075	1832

Notes: All regressions include control variables at degree program level (average age of instructors, % female instructors, ECTS abroad/total ECTS, ) and department level (full professor (%), associate professor (%), post-docs (%), and per capita research funds). Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by academic macro-field. Robust standard error in brackets \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ , + $p < 0.15$

two focus on, respectively, only BA-level and MA-level programs.<sup>13</sup> The results concerning the other two indicators of teaching quality are similarly organized: columns (4), (5), and (6) for the number of students with 40 or more ECTS in the first year and columns (7), (8), and (9) for graduates one year after the ordinary duration of the program.

From our results, it seems that teaching quality is not robustly associated with research quality. We find positive and (weakly) significant coefficients when focusing on the sample including all programs and only MA-level programs, but not for graduates at time  $N + 1$ . In fact, teaching quality in BA-level programs is not significantly correlated with research performance and the coefficient is even significantly negative when considering graduates at  $N + 1$ . As concerns the number of degree programs per department, its coefficient is not significant and negative for BA-level programs, but significant and positive for MA-level degrees (except for the case of graduates in  $N + 1$ ). While the student–instructor ratio is positive for all specifications, even if strongly significant only for the case of MA-level Graduates in  $N + 1$ . Hence, while yardstick competition has no effect on the teaching effectiveness for BA-level programs, it has a positive effect on MA-level programs. We interpret the positive correlation between the student–instructor ratio and the quality of teaching as a consequence of the fund allocation scheme where having more students raises the amount of funds devoted to each program.

The coefficient estimates for the number of instructors, a proxy for competition in career concerns, are instead consistent across teaching quality proxies. Apparently, degree programs with more instructors perform less in terms of teaching quality. This is corroborated by statistical significance only for BA-level programs, presumably because BA-level programs do not often allow instructors to teach subjects that are close to their research interests.

As regards control variables (reported in Table A.1 in the Appendix), the average age of instructors shows predominantly negative and significant coefficients. Indeed, younger instructors may possess more frontier knowledge on teaching methods and/or on research concepts and tools, which may be valuable especially for MA-level students. We find positive coefficients for the shares of full and associate professors and for the per capita research funds. Full and associate professors, indeed, are supposedly more talented on average than assistant professors, given their academic age, or more experienced. The availability of more research funds per capita allows to acquire equipment which may be helpful for teaching and may improve the trade-off between teaching and research efforts by relaxing the time constraint. Because this trade-off is more stringent when teaching is perceived as subtracting precious time for research, it is no surprise to find that the coefficients of per capita research funds lack significance in MA-level programs, where research

---

<sup>13</sup> We consider only BA and two-year MA programs for the analysis in columns (2) and (3), respectively, while we exclude five-year MA programs. This is due to the mixed nature of the subjects provided by these last types of degrees, characterized both by basic and more general contents and by more specific and in-depth knowledge.

and teaching are more complementary. Another variable that displays positive and significant coefficients, both for BA-level and for MA-level programs, is the number of ECTS abroad over total ECTS. This may be due to a selection effect: only students who are more motivated and have a higher than average income may afford visiting a foreign university.

Students' satisfaction, too, could depend on the research performance of instructors. Table 3 presents estimates of our econometric model including, as dependent variables, two alternative satisfaction proxies: the percentage of graduates who declared they would enrol again in the same program (*Program satisfaction I*) and the percentage of graduates who are overall satisfied about their degree program (*Program satisfaction II*).<sup>14</sup>

The most striking result in Table 3 is that, whatever the degree level and the satisfaction variable, there is no significant correlation between satisfaction and research performance. In fact, the signs are negative, except in one column. One may argue that top researchers may not possess the teaching skills, the business contacts, or the incentives to make their classes fit for the expectations of students who, after graduation, will look for jobs outside of academia. Satisfaction may rather be improved by instructors who can offer opportunities for business sector stages and job interviews. This lacking correlation would also be in line with the most recent criticisms on the use of the students' opinions to measure teaching quality (Weinberg et al. 2009; Babcock 2010; Carrell & West 2010; Braga et al. 2014). According to this literature, the most skilled researchers, if they are also more demanding as teachers, would be penalized when evaluated through students' opinions, as students may seek to minimize efforts.

Despite the lack of significant correlations between satisfaction and research performance, the estimates on graduates satisfaction bring a few interesting take-home messages (see control variable results in Table A.2 in the Appendix). One is that, for a given research performance, it pays off for degree programs to allocate basic subjects to the professors who are responsible for managing the degree. Coefficients to the corresponding variable are positive and statistically significant, both in the BA- and in the MA-level programs. Another insight is that more ECTS abroad do not help in terms of satisfaction in the MA programs case. Perhaps, students unfavourably compare their degree of origin with the foreign one, if the latter is better organized; or, they may attribute to their degree of origin the responsibility of a weak performance abroad. In these estimates, too, the average age of instructors shows negative coefficients. The coefficient is statistically significant for MA-level degrees, where younger professors—supposing they are on the scientific knowledge frontier—may be most intellectually stimulating.

Table 4 performs the same exercise as in Tables 2 and 3, but using a slightly different research quality indicator. It is represented by the average of the R values

---

<sup>14</sup> Arguably, the latter is a more reliable measure of satisfaction for our purposes, as a negative response to the former may as well mean that graduates would have rather chosen another program in the same university or even in the same department.



**Table 3** The relationship between teaching effectiveness and research. The effect on student satisfaction

	(1)	(2)	(3)	(4)	(5)	(6)
	Program satisfaction I			Program satisfaction II		
Research quality	All programs	BA-level	MA-level	All programs	BA-level	MA-level
	-0.0150	-0.0073	-0.0259	-0.0079	0.0078	-0.0233
	[0.0179]	[0.0293]	[0.0233]	[0.0169]	[0.0287]	[0.0217]
# Instructors (degree program)	0.0001 <sup>+</sup>	0.0001*	0.0003	0.0001 <sup>+</sup>	0.0001**	0.0003
	[0.0001]	[0.0001]	[0.0004]	[0.0000]	[0.0000]	[0.0004]
Students/instructors (degree program)	0.0022***	0.0017***	0.0049***	0.0027***	0.0019***	0.0055***
	[0.0004]	[0.0005]	[0.0010]	[0.0004]	[0.0005]	[0.0010]
# of degree programs in department	-0.0015**	-0.0022**	-0.0005	-0.0011	-0.0016	-0.0008
	[0.0008]	[0.0011]	[0.0012]	[0.0008]	[0.0012]	[0.0013]
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.394	0.411	0.398	0.462	0.485	0.456
Observations	4095	1989	1833	4095	1989	1833

Notes: All regressions include control variables at degree program level (average age of instructors, % female instructors, ECTS abroad/total ECTS, ) and department level (full professor (%), associate professor (%), post-docs (%), and per capita research funds). Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by academic macro-field. Robust standard error in brackets \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ , <sup>+</sup> $p < 0.15$

(VQR 2011–2014) of all university teachers belonging to each academic discipline, weighted by the ECTS of the related programs. Thus, compared to the research quality measure used so far, the new indicator would represent a more precise measure of research quality because it would capture relevant differences in research performance, which are specific to academic discipline, rather than academic macro-field, to which each instructor belongs. However, such indicator is calculated only for MA-level degree programs.

According to the results in Table 4, the positive correlation is confirmed for all teaching indicators and it emerges even more clearly, as it can be grasped from the larger point estimates, especially in columns (3), (4), and (5). The sign of the number of degree programs is still positive and significant in the case of average ECTS obtained in the first year and of the percentage of students enrolled in the second year after obtaining 40 or more ECTS in the first year. Also, the positive influence of the student–instructor ratio is confirmed. Conversely, our proxy for career concerns (the number of instructors) does not affect teaching quality negatively as we would expect and only shows a statistically significant (positive) coefficient in respect to one of the satisfaction variables (column 5). As for the other control variables (see Table A.3 in the Appendix), negative coefficients on age are confirmed, as well as the negative correlation of graduates satisfaction with per capita research funds. ECTS gained abroad keep their positive correlation with teaching quality proxies, except in the case of graduates' satisfaction.

### ***6.1 Estimates Including Interaction Terms***

As emerging from the above tables, the evidence on the effects of yardstick competition and career concerns in multi-task multi-unit universities is, at best, mixed. Though, our estimation strategy so far has probably overlooked the essentially non-linear relationship between yardstick competition, research performance, and teaching quality. If resources are distributed among degree programs in the same department through some form of yardstick competition, we expect the relationship between teaching quality and research performance to change across departments characterized by different competitive conditions. Let us set aside the extreme case of departments with a single degree program: in such a case, no competition arises, and therefore instructors will put their efforts on research, in order to achieve their career upgrades, to the detriment of teaching quality. Consider, on the opposite, a department with several degree programs. It would be difficult to avoid free-riding behaviours in that case, as strategic interaction among degree programs would be weaker, and each degree program may rather behave in a sort of "price-taking" fashion. The argument by Gautier and Wauthy (2007) may imply a positive correlation between teaching and research quality in departments with relatively few degree programs, where performance comparisons among degree programs are easier. Therefore, in further estimations—focusing on the research performance indicator studied in Table 4—we include an interaction term

**Table 4** The relationship between teaching effectiveness and research. Using an alternative measure for research quality

	(1)	(2)	(3)	(4)	(5)
	Average ECTS in I year	II year with 40+ ECTS in I year	Graduates in N + 1	Program satisfaction I	Program satisfaction II
Research quality	0.0325* [0.0175]	0.0806*** [0.0273]	0.0575** [0.0278]	0.2620*** [0.0256]	0.3285*** [0.0271]
# Instructors (degree program)	0.0004 [0.0003]	0.0004 [0.0004]	0.0004 [0.0003]	0.0005 [0.0004]	0.0005+ [0.0004]
Students/instructors (degree program)	0.0002 [0.0008]	0.0011 [0.0010]	0.0021*** [0.0008]	0.0038*** [0.0010]	0.0042*** [0.0009]
# of degree programs in department	0.0016* [0.0009]	0.0021+ [0.0013]	-0.0016 [0.0011]	-0.0002 [0.0012]	-0.0005 [0.0012]
Control variables	Yes	Yes	Yes	Yes	Yes
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.401	0.391	0.304	0.439	0.513
Observations	1870	1883	1832	1833	1833

Notes: All regressions include control variables at degree program level (average age of instructors, % female instructors, ECTS abroad/total ECTS, ) and department level (full professor (%), associate professor (%), post-docs (%), and per capita research funds). Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by specific academic discipline. Only MA degree programs (II years) are included. Robust standard error in brackets \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ , + $p < 0.15$

between the research performance indicator and the number of degree programs per department. Furthermore, we separately analyse two subsamples: one including departments with a below-median number of degree programs and those above the median. The median number of degree programs per department is 7 (considering only MA-level degrees). We expect the coefficient of the interaction term to be positive in degree programs which compete with few other programs in the same department and a negative coefficient if the number of competing degree programs is larger.

Columns (1), (2), and (3) of Panel A in Table 5 report estimates of the model, using the average ECTS in the first year as the dependent variable. The sign and significance of the coefficients associated with the interaction term in the two subsamples confirm our expectations: the coefficient is 0.0246 and significant in the below-median subsample and  $-0.0118$  and significant in the above-median subsample. Hence, the role of yardstick competition in yielding a positive relationship between teaching and research quality is confirmed, while it vanishes when the number of competing degree programs is relatively large. The direct effect of the research performance indicator keeps its positive sign in the whole sample (0.0720) and is even stronger in the above-median subsample (0.1373). Similar results hold for the direct effect of the number of degree programs (0.0064 whole sample, 0.0084 above-median subsample, while negative in the below-median subsample). Finally, we replicate the above analysis on the heterogeneity effects, using the graduates' program satisfaction as measure of teaching quality and find similar results, as reported in columns (1), (2), and (3) of Panel B in Table 5. In particular, the interaction term is positively associated with the program satisfaction where the number of degree programs in department is low (column 2), confirming the role of yardstick competition in modulating the research–teaching relationship.

Another possible source of heterogeneity in the teaching–research quality relationship may arise in reference to the student–instructor ratio. For high-performance researchers, teaching time may have a rather high opportunity cost, but this can be mitigated if students are less. A smaller number of students allow to customize teaching methods and to involve students in research-intensive activities (e.g. data collection, experiments, discussion of scholarly articles). To capture such form of non-linearity, here too, we split the sample based on the median of student–instructor ratio, which is equal to 7.325. According to the above insights, we expect a positive correlation between teaching and research quality especially in degrees with a below-median student–instructor ratio and a positive correlation between teaching indicator and the student–instructor ratio; while the interaction term research quality times student–instructor ratio should feature a negative coefficient in degrees with below-median student–instructor ratio. This is because, as showed in Sect. 2, when the budget sharing rule adopted by departments to allocate resources among study programs is based on both research performance and the number of students, as occurs in Italy, smaller classes imply also lower amount of funds available to departments for career advancements and research. This limits, on the one hand, the incentive for academic professors to free ride on teaching activity, by rising the appropriability of teaching effort by academics, and, on the other hand, the scope

**Table 5** The relationship between teaching effectiveness and research. Controlling for heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: the dependent variable is: Average ECTS in 1 year						
	Heterogeneity by # of degree programs in department			Heterogeneity by students/instructors		
		Below median	Above median		Below median	Above median
Research quality	0.0720** [0.0280]	-0.0723 [0.0662]	0.1373** [0.0580]	0.0359+ [0.0225]	0.0506+ [0.0321]	0.1025 [0.0821]
# of degree programs in department	0.0064** [0.0027]	-0.0231* [0.0124]	0.0084** [0.0041]			
Research × # of degree programs in department	-0.0050* [0.0028]	0.0246** [0.0123]	-0.0118*** [0.0043]			
Students/instructors (degree program)				0.0009 [0.0028]	0.0152* [0.0089]	0.0036 [0.0057]
Research × Students/instructors				-0.0007 [0.0027]	-0.0102 [0.0094]	-0.0036 [0.0054]
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.402	0.444	0.405	0.401	0.332	0.494
Observations	1870	1117	753	1870	1269	601

(continued)

**Table 5** (continued)

Research quality	0.3034*** [0.0419]	0.1006 [0.1125]	0.2635*** [0.0721]	0.4797*** [0.0344]	0.6419*** [0.0509]	0.0487 [0.0654]
# of degree programs in department	-0.0035 [0.0036]	-0.0434** [0.0188]	-0.0054 [0.0061]			
Research × # of degree programs in department	0.0032 [0.0034]	0.0446** [0.0180]	0.0045 [0.0051]			
Students/instructors (degree program)				0.0324*** [0.0043]	0.1378*** [0.0113]	0.0028 [0.0040]
Research × Students/instructors				-0.0293*** [0.0040]	-0.1211*** [0.0119]	-0.0020 [0.0039]
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.513	0.439	0.609	0.536	0.587	0.390
Observations	1833	1087	746	1833	1235	598

*Notes:* All regressions in Panel A and Panel B include control variables at degree program level (average age of instructors, % female instructors, # instructors, ECTS abroad/total ECTS, ) and department level (full professor (%), associate professor (%), post-docs (%), and per capita research funds). Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by specific academic discipline. Only MA degree programs (II years) are included. Robust standard error in brackets \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ , +  $p < 0.15$

of winner-picking in research activity. Results in columns (4), (5), and (6) of both Panels A and B in Table 5 reveal that our expectations are confirmed. In fact, in Panel A, we find that the positive influence of research performance, which results in the whole sample analysis, seems to be driven by the below-median analysis where the coefficient increases in size and is still significant (even if weakly). The analysis on program satisfaction in Panel B of Table 5 bolsters our expectations. The coefficient associated with research performance in the whole sample is 0.4797 and strongly significant, rising to 0.6419 in degrees with below-median student–instructor ratio, and losing significance in the above-median subsample. Finally, the interaction between research performance and the student–instructor ratio is characterized by a negative coefficient, significant and greater in the below-median subsample, but lower and not significant in the above-median subsample.

## 6.2 *Bibliometric vs. Non-bibliometric Fields*

We now explore the heterogeneity in the relationship teaching–research by bibliometric field.<sup>15</sup> Results in Table 6 show the link between the average ECTS acquired by students during the first academic years and the research indicator normalized by academic macro-field. As regards bibliometric fields, we find a positive correlation between our teaching indicator and the instructors’ performance in research, which arises both at BA- and at MA-level degree program. The positive result also for the BA-level degree programs is not surprising, given the less generalist nature of hard science programs, which makes the transfer of knowledge suitable even to the younger and unselected students. As for the non-bibliometric fields, by contrast, we find a negative role of the research in enhancing the students’ academic performance of BA level. This is probably a consequence of the most generalist nature of social science programs. However, these results could also depend on the measurement error of the indicator we use for the research performance of non-bibliometric instructors; the R indicator would at most capture the scientific products of international relevance, penalizing domestic publications, which are more present in social science (Shin 2011).

Interestingly, results are more homogeneous when we control for the moderating effect of the number of degree programs in the department (Table 7), from which we can infer the main role of the yardstick competition in influencing the teaching–research relationship for both bibliometric and non-bibliometric sectors.

---

<sup>15</sup> We include a degree program in the analysis for the bibliometric sector if the majority of the instructors involved in the degree program belong to a scientific academic discipline classified as bibliometrics. Consequently, a degree program falls within the non-bibliometric analysis in the reverse case in which the majority of instructors involved in the degree program belong to a scientific academic discipline classified as non-bibliometric.

**Table 6** The relationship between teaching effectiveness and research. Bibliometric vs. non-bibliometric sector

	(1)	(2)	(3)	(4)	(5)	(6)
The dependent variable is: average ECTS in 1 year						
	Bibliometric sector			Non-bibliometric sector		
	All programs	BA-level	MA-level	All programs	BA-level	MA-level
Research quality	0.0543*** [0.0197]	0.0448+ [0.0306]	0.0537** [0.0273]	0.0003 [0.0194]	-0.0454+ [0.0314]	0.0301 [0.0241]
# Instructors (degree program)	-0.0003*** [0.0000]	-0.0003*** [0.0000]	0.0005 [0.0004]	-0.0002 [0.0002]	-0.0001 [0.0002]	0.0006 [0.0005]
Students/instructors (degree program)	0.0006 [0.0006]	-0.0000 [0.0006]	0.0023* [0.0014]	-0.0002 [0.0007]	0.0004 [0.0008]	-0.0008 [0.0009]
# of degree programs in department	-0.0012* [0.0007]	-0.0009 [0.0009]	-0.0007 [0.0013]	0.0039*** [0.0011]	0.0036** [0.0017]	0.0051*** [0.0015]
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.508	0.589	0.367	0.524	0.509	0.432
Observations	2497	1316	994	1676	700	876

Notes: All regressions include control variables at degree program level (average age of instructors, % female instructors, ECTS abroad/total ECTS, ) and department level (full professor (%), associate professor (%), post-docs (%), and per capita research funds). Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by academic macro-field. Robust standard error in brackets \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ , + $p < 0.15$



**Table 7** Bibliometric vs. non-bibliometric sector. Heterogeneity by # of degree programs in department

(1)	(2)	(3)	(4)	(5)	(6)
The dependent variable is: Average ECTS in 1 year					
Bibliometric sector					
	Below median	Above median		Below median	Above median
Research quality	0.0709 [0.0518]	0.0395 [0.1077]	0.1488+ [0.0974]	0.0756** [0.0313]	-0.1344 [0.1044]
# of degree programs in department	0.0041 [0.0058]	-0.0023 [0.0198]	0.0035 [0.0077]	0.0102*** [0.0026]	-0.0373* [0.0196]
Research × # of degree programs in department	-0.0050 [0.0058]	0.0015 [0.0191]	-0.0129+ [0.0079]	-0.0054** [0.0027]	0.0367* [0.0199]
Control variables	Yes	Yes	Yes	Yes	Yes
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.365	0.387	0.404	0.432	0.484
Observations	994	603	391	876	514

Notes: All regressions include control variables at degree program level (average age of instructors, % female instructors, # instructors, ECTS abroad/total ECTS, ) and department level (full professor (%), associate professor (%), post-docs (%), and per capita research funds). Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by specific academic discipline. Only MA-level degree programs (II years) are included. Robust standard error in brackets \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ , + $p < 0.15$

### 6.3 *Assessment*

Overall, our estimates allow to draw some food for thought. A first take-home message is that scientific performance and teaching quality move together in line with yardstick competition among degree programs activated in the same department. The positive conditional correlation between research and teaching indicators, measured at the study program level, is stronger in departments where degree programs are relatively few and can be immediately compared and declines whenever the degree programs competing for the resources allocated by their department are many. This is the message from coefficient estimates on the interaction between research performance and the number of study programs per department. Such results suggest that the multi-unit and multi-task nature of Italian university departments—along with the lack of alignment between university goals and individual goals—subverts the trade-off that otherwise would characterize individual decisions about teaching and research efforts.

Second, the teaching–research relationship is best understood by analysing relatively homogeneous subsamples of degree programs. BA-level and MA-level students differ in terms of knowledge base, learning potential, and goals. At the same time, professors may have different expectations from students at different education levels and may tune their teaching style accordingly. Complementarity between teaching and research is less likely in BA-level programs, where the basics are taught and topics are far from the scientific interests of professors, and indeed we find the teaching–research relationship to be weaker in the subsample focused on BA-level degrees.

Our estimates can only be interpreted as correlations since we do not have information on the identity of universities and therefore cannot rely on a causality identification strategy. Yet, some insights on how to unleash an effective knowledge transmission can be drawn from further econometric exercises. In particular, the specifications including interaction terms confirm that yardstick competition and a budget sharing rule, which includes the number of students, are essential in order to allow a more effective transmission of advanced knowledge to students.

## 7 **Concluding Remarks**

The growing research orientation of universities in recent years has fostered an intense debate among academics on the consequences on teaching activities. There are, in fact, several reasons in both support of and against the complementarity between the two main university missions: teaching and research. Empirical evidence from previous studies is mixed. Therefore, the question whether being a good researcher implies being also a good teacher is still an open question.

This study contributes to the ongoing debate in that it examines the relationship between teaching and research in the Italian university system. To do so, we use a

rich dataset provided by ANVUR on the study programs of all Italian universities and measure the quality of teaching using both students' academic performance and their degree of satisfaction with the programs attended. Our analysis suggests that the involvement of good quality researchers in the program supports mostly the academic career of MA-level degree program students and increases their program satisfaction, once they graduate. On the contrary, with regard to the BA-level degree program students, we find some negative correlation between teaching and research quality.

An interesting result that emerges from our study is the heterogeneous effect in the teaching–research relationship, which stems from the multi-unit organization of departments. In particular, we find that the positive correlation between research and teaching indicators is stronger in departments where degree programs are relatively few and shrinks when the number of the degree programs competing for the resources allocated by their department increases, suggesting the major role played by yardstick competition in shaping the teaching–research relationship in the Italian universities.

## **Appendix**

**Table A.1** The relationship between teaching effectiveness and research. The effect on student academic performance—complete specification of Table 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Average ECTS in I year		MA-level	II year with 40+ ECTS in I year		MA-level	Graduates in N + 1		
	All programs	BA-level		All programs	BA-level		All programs	BA-level	MA-level
Research quality	0.0301** [0.0134]	0.0264 [0.0217]	0.0364** [0.0175]	0.0294+ [0.0185]	0.0283 [0.0300]	0.0420* [0.0246]	-0.0137 [0.0151]	-0.0570** [0.0227]	0.0172 [0.0196]
Average age of instructors (degree program)	-0.0014* [0.0008]	0.0008 [0.0013]	-0.0034*** [0.0011]	-0.0027** [0.0012]	-0.0002 [0.0019]	-0.0053*** [0.0017]	-0.0008 [0.0010]	-0.0026* [0.0015]	-0.0005 [0.0014]
% Female instructors (degree program)	0.0427** [0.0183]	0.0925*** [0.0277]	0.0052 [0.0249]	0.0434* [0.0263]	0.1470*** [0.0400]	-0.0325 [0.0354]	0.0027 [0.0213]	-0.0292 [0.0306]	0.0119 [0.0304]
# Instructors (degree program)	-0.0003*** [0.0000]	-0.0003*** [0.0000]	0.0004 [0.0003]	-0.0004*** [0.0001]	-0.0003*** [0.0001]	0.0004 [0.0004]	-0.0003*** [0.0001]	-0.0003*** [0.0001]	0.0003 [0.0003]
Students/instructors (degree program)	0.0001 [0.0005]	0.0002 [0.0007]	0.0004 [0.0007]	0.0002 [0.0006]	0.0000 [0.0008]	0.0015+ [0.0010]	0.0007 [0.0005]	0.0011+ [0.0007]	0.0024*** [0.0008]
ECTS abroad/total ECTS (degree program)	0.2807*** [0.1041]	0.5098** [0.2188]	0.2915** [0.1224]	0.3404** [0.1424]	0.4880* [0.2624]	0.3221* [0.1685]	0.3066*** [0.0780]	0.9404*** [0.1765]	0.2500*** [0.0830]
% of reference instructors in the degree program	0.0012 [0.0125]	0.0036 [0.0189]	0.0048 [0.0171]	0.0088 [0.0178]	-0.0240 [0.0251]	0.0301 [0.0250]	-0.0044 [0.0206]	-0.0437* [0.0244]	0.0164 [0.0295]
Full professor in department (%)	0.1146*** [0.0395]	0.1199** [0.0550]	0.1063* [0.0606]	0.0761 [0.0558]	0.1303* [0.0764]	0.0401 [0.0859]	0.0523 [0.0448]	0.0114 [0.0615]	0.1513** [0.0655]

(continued)

**Table A.1** (continued)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Average ECTS in I year		II year with 40+ ECTS in I year		Graduates in $N + 1$			
	All programs	BA-level	MA-level	All programs	BA-level	MA-level	All programs	BA-level
Associate professor in department(%)	0.0696** [0.0292]	0.1014** [0.0428]	0.0763* [0.0443]	0.1158*** [0.0434]	0.1562** [0.0615]	0.1160* [0.0696]	0.0656* [0.0348]	0.0141 [0.0477]
Post-docs in department (%)	-0.0157 [0.0141]	-0.0090 [0.0192]	-0.0246 [0.0215]	0.0013 [0.0200]	0.0033 [0.0256]	0.0018 [0.0312]	0.0073 [0.0181]	0.0019 [0.0219]
Per capita research funds (department)	0.0061** [0.0027]	0.0125*** [0.0038]	-0.0028 [0.0041]	0.0046 [0.0042]	0.0124** [0.0058]	-0.0052 [0.0063]	0.0044 [0.0033]	0.0067+ [0.0044]
# of degree programs in department	-0.0000 [0.0005]	-0.0005 [0.0007]	0.0017* [0.0009]	-0.0006 [0.0008]	-0.0010 [0.0012]	0.0021* [0.0013]	-0.0013** [0.0006]	-0.0007 [0.0008]
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.509	0.549	0.402	0.443	0.485	0.388	0.647	0.659
Observations	4173	2016	1870	4192	2022	1883	4075	1976

*Notes:* Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by academic macro-field. Robust standard error in brackets \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ , + $p < 0.15$

**Table A.2** The relationship between teaching effectiveness and research. The effect on student satisfaction—complete specification of Table 3

	(1)	(2)		(3)	(4)	(5)	(6)
		Program satisfaction I			Program satisfaction II		
	All programs	BA-level	MA-level	MA-level	All programs	BA-level	MA-level
Research quality	-0.0150 [0.0179]	-0.0073 [0.0293]	-0.0259 [0.0233]	-0.0259 [0.0233]	-0.0079 [0.0169]	0.0078 [0.0287]	-0.0233 [0.0217]
Average age of instructors (degree program)	-0.0031*** [0.0012]	-0.0010 [0.0017]	-0.0050*** [0.0018]	-0.0050*** [0.0018]	-0.0022* [0.0011]	0.0003 [0.0016]	-0.0046*** [0.0017]
% Female instructors (degree program)	0.0131 [0.0265]	0.0540 [0.0389]	-0.0259 [0.0380]	-0.0259 [0.0380]	0.0089 [0.0257]	0.0523 [0.0375]	-0.0211 [0.0376]
# Instructors (degree program)	0.0001+ [0.0001]	0.0001* [0.0001]	0.0003 [0.0004]	0.0003 [0.0004]	0.0001+ [0.0000]	0.0001** [0.0000]	0.0003 [0.0004]
Students/instructors (degree program)	0.0022*** [0.0004]	0.0017*** [0.0005]	0.0049*** [0.0010]	0.0049*** [0.0010]	0.0027*** [0.0004]	0.0019*** [0.0005]	0.0055*** [0.0010]
ECTS abroad/total ECTS (degree program)	-0.0220 [0.0992]	0.5882*** [0.1827]	-0.0869 [0.1066]	-0.0869 [0.1066]	0.0068 [0.1141]	0.3470* [0.1982]	-0.0319 [0.1320]
% of reference instructors in the degree program	0.5531*** [0.0171]	0.5796*** [0.0224]	0.5345*** [0.0248]	0.5345*** [0.0248]	0.6801*** [0.0193]	0.7284*** [0.0256]	0.6388*** [0.0282]

(continue)

**Table A.2** (continue)

	(1)	(2)	(3)	(4)	(5)	(6)
	Program satisfaction I		Program satisfaction II			
	All programs	BA-level	MA-level	All programs	BA-level	MA-level
Full professor in department(%)	-0.0937* [0.0535]	-0.0685 [0.0699]	-0.1716* [0.0892]	-0.1094** [0.0499]	-0.1099* [0.0609]	-0.1637* [0.0908]
Associate professor in department(%)	0.0206 [0.0417]	0.0265 [0.0562]	0.0101 [0.0682]	0.0364 [0.0396]	0.0512 [0.0527]	0.0249 [0.0677]
Post-docs in department (%)	0.0407** [0.0198]	0.0306 [0.0247]	0.0382 [0.0329]	0.0275+ [0.0190]	0.0021 [0.0238]	0.0440 [0.0319]
Per capita research funds (department)	0.0032 [0.0039]	0.0156*** [0.0053]	-0.0156** [0.0064]	0.0046 [0.0038]	0.0121** [0.0054]	-0.0070 [0.0059]
# of degree programs in department	-0.0015** [0.0008]	-0.0022** [0.0011]	-0.0005 [0.0012]	-0.0011 [0.0008]	-0.0016 [0.0012]	-0.0008 [0.0013]
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.394	0.411	0.398	0.462	0.485	0.456
Observations	4095	1989	1833	4095	1989	1833

*Notes:* Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by academic macro-field. Robust standard error in brackets \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ , +  $p < 0.15$

**Table A.3** The relationship between teaching effectiveness and research. Using an alternative measure for research quality—complete specification of Table 4

	(1)	(2)	(3)	(4)	(5)
	Average ECTS in I year	II year with 40+ ECTS in I year	Graduates in N + 1	Program satisfaction I	Program satisfaction II
Research quality	0.0325* [0.0175]	0.0806*** [0.0273]	0.0575** [0.0278]	0.2620*** [0.0256]	0.3285*** [0.0271]
Average age of instructors (degree program)	-0.0037*** [0.0011]	-0.0051*** [0.0016]	-0.0001 [0.0014]	-0.0017 [0.0018]	-0.0005 [0.0018]
% Female instructors (degree program)	0.0029 [0.0247]	-0.0351 [0.0353]	0.0113 [0.0302]	-0.0177 [0.0365]	-0.0117 [0.0353]
# Instructors (degree program)	0.0004 [0.0003]	0.0004 [0.0004]	0.0004 [0.0003]	0.0005 [0.0004]	0.0005+ [0.0004]
Students/instructors (degree program)	0.0002 [0.0008]	0.0011 [0.0010]	0.0021*** [0.0008]	0.0038*** [0.0010]	0.0042*** [0.0009]
ECTS abroad/total ECTS (degree program)	0.2963** [0.1236]	0.3254* [0.1697]	0.2468*** [0.0846]	-0.1199 [0.1174]	-0.0715 [0.1486]
% of reference instructors in the degree program	-0.0197 [0.0204]	-0.0304 [0.0306]	-0.0258 [0.0297]	0.3448*** [0.0302]	0.4005*** [0.0323]
Full professor in department(%)	0.1137* [0.0593]	0.0292 [0.0846]	0.1342** [0.0658]	-0.3100*** [0.0861]	-0.3316*** [0.0859]

(continued)



Table A.3 (continued)

	(1)	(2)	(3)	(4)	(5)
	Average ECTS in I year	II year with 40+ ECTS in I year	Graduates in $N + 1$	Program satisfaction I	Program satisfaction II
Associate professor in department(%)	0.0767* [0.0442]	0.0975 [0.0689]	0.1092** [0.0550]	-0.1155* [0.0662]	-0.1291** [0.0656]
Post-docs in department (%)	-0.0256 [0.0214]	0.0017 [0.0308]	0.0162 [0.0286]	0.0414 [0.0309]	0.0475* [0.0287]
Per capita research funds (department)	-0.0023 [0.0041]	-0.0054 [0.0062]	0.0019 [0.0049]	-0.0205*** [0.0061]	-0.0129** [0.0055]
# of degree programs in department	0.0016* [0.0009]	0.0021+ [0.0013]	-0.0016 [0.0011]	-0.0002 [0.0012]	-0.0005 [0.0012]
Geographic macroarea FE	Yes	Yes	Yes	Yes	Yes
Degree field FE	Yes	Yes	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.401	0.391	0.304	0.439	0.513
Observations	1870	1883	1832	1833	1833

Notes: Research quality is the average score assigned to the instructors teaching in a degree program, during the second exercise of VQR 2011–2014. The research measure is normalized by specific academic discipline. Only MA degree programs (II years) are included. Robust standard error in brackets  
\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ , +  $p < 0.15$

## References

- Babcock, P. (2010). Real costs of nominal grade inflation? new evidence from student course evaluations. *Economic Inquiry*, 48, 983–996.
- Bak, H., & Kim, D. (2015). Too much emphasis on research? An empirical examination of the relationship between research and teaching in multitasking environments. *Research in Higher Education*, 56, 843–860.
- Barnett, R. (1992). Linking teaching and research: A critical inquiry. *Journal of Higher Education*, 63(6), 619–636.
- Bäkera, A., & Goodallb, A. (2020). Feline followers and “umbrella carriers”: Department chairs’ influence on faculty job satisfaction and quit intentions. *Research Policy*, 49(4), 1–14.
- Bradford, L., Foltz, J., & Prager, D. (2014). Making time for science. *Research Policy*, 43(1), 21–31.
- Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students’ evaluations of professors. *Economics of Education Review*, 41, 71–88.
- Braxton, J. (1996). Contrasting perspectives on the relationship between teaching and research. *New Directions for Institutional Research*, 90, 5–14.
- Brew, A. (1999). Research and teaching: Changing relationships in a changing context. *Studies in Higher Education*, 24(3), 291–301.
- Brew, A., & Boud, D. (1995). Teaching and research: Establishing the vital link with learning. *Higher Education*, 29(3), 261–273.
- Cadez, S., Dimovski, V., & Groff, M. (2017). Research, teaching and performance evaluation in academia: The salience of quality. *Studies in Higher Education*, 42(8), 1455–1473.
- Carillo, M., & Papagni, E. (2005). Scientific research, externalities and economic growth.
- Carillo, M., & Papagni, E. (2014). “little science” and “big science”: The institution of “open science” as a cause of scientific and economic inequalities among countries. *Economic Modelling*, 43, 42–56.
- Carillo, M. R., Papagni, E., & Sapio, A. (2013). Do collaborations enhance the high-quality output of scientific institutions? Evidence from the Italian research assessment exercise. *Journal of Socio-Economics*, 47C, 25–36.
- Carrell, S. E., & West, J. E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy*, 118, 409–432.
- De Philippis, M. (2020). Multi-task agents and incentives: The case of teaching and research for university professors. *The Economic Journal*, 131(636), 1643–1681.
- Douglas, A. S. (2013). Advice from the professors in a university social sciences department on the teaching-research nexus. *Teaching in Higher Education*, 18(4), 377–388.
- Garcia-Gallego, A., Georgantzis, N., Martin-Montaner, J., & Perez-Amaral, T. (2015). How do research and administrative duties affect university professors’ teaching? *Applied Economics*, 47(45), 4868–4883.
- Gautier, A., & Wauthy, X. (2007). Teaching versus research: A multi-tasking approach to multi-department universities. *European Economic Review*, 51, 273–295.
- Griffiths, R. (2004). Knowledge production and the research-teaching nexus: The case of the built environment disciplines. *Studies in Higher Education*, 29, 709–726.
- Hattie, J., & Marsh, H. (1996). The relationship between research and teaching: A meta-analysis. *Review of Educational Research*, 66(4), 507–542.
- Healey, M. (2005). *Reshaping the university: New relationships between research, scholarship and teaching*. Chapter Linking Research and Teaching: Disciplinary Spaces (pp. 30–42). Maidenhead, UK: Open University Press.
- Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization*, 7, 24–52.
- Marsh, H. W. (1987). Students’ evaluations of university teaching: Research findings, methodological issues, and directions for further research. *International Journal Educational Research*, 11, 253–388.

- Marsh, H. W., & Hattie, J. (2002). The relation between research productivity and teaching effectiveness: Complementary, antagonistic, or independent constructs? *The Journal of Higher Education*, 73(5), 603–641.
- Palali, A., van Elka, R., Bolhaara, J., & Rudb, I. (2018). Are good researchers also good teachers? The relationship between research quality and teaching quality. *Economics of Education Review*, 64, 40–49.
- Parker, L. D. (2012). Beyond the ticket and the brand: Imaging an accounting research future. *Accounting and Finance*, 52(4), 1153–1182.
- Payne, A., & Roberts, J. (2010). Government oversight of public universities: Are centralized performance schemes related to increased quantity or quality? *The Review of Economics and Statistics*, 92(1), 207–212.
- Qamar uz Zaman, M. (2004). *Review of the academic evidence on the relationship between teaching and research in higher education*. Annesley: DfES Publications.
- Ramsden, P., & Moses, I. (1992). Associations between research and teaching in Australian higher education. *Higher Education*, 23, 273–295.
- Rodriguez, R., & Rubio, G. (2016). Teaching quality and academic research. *International Review of Economics Education*, 23, 10–27.
- Shin, J., & Kim, Y. (2017). *The changing academic profession in Hong Kong* (Vol. 19), chapter The Teaching and Research Nexus Under Research University Initiatives: A Comparative View for East Asia (Vol. 19). Cham: Springer.
- Shin, J. C. (2011). Teaching and research nexuses across faculty career stage, ability and affiliated discipline in a South Korean Research University. *Studies in Higher Education*, 36(4), 485–503.
- Sullivan, A. (1996). Teaching norms and publication productivity. *New Directions for Institutional Research*, 90, 15–21.
- Sylos Labini, M., & Zinovyeva, N. (2011). The relationship between academic research and teaching quality.
- ter Bogt, H. J., & Scapens, R. (2012). Performance management in universities: Effects of the transition to more quantitative measurement systems. *European Accounting Review*, 21(3), 451–497.
- Walstad, W., & Allgood, S. (2005). Views of teaching and research in economics and other disciplines. *American Economic Review Papers and Proceedings*, 95(2), 177–183.
- Weinberg, B. A., Hashimoto, M., & Fleisher, B. M. (2009). Evaluating teaching in higher education. *The Journal of Economic Education*, 40, 227–262.

**Maria Rosaria Carillo** (Ph.D. University of Naples, Federico II) is Professor of Economics at the University of Naples Parthenope. Her research interests include economics of science, education economics, economic development and migration.

**Alessandro Sapiro** (Ph.D. Scuola Sant’Anna, Pisa) is Professor of Economics at the University of Naples Parthenope. His research interests include economics of science, energy economics and industrial dynamics.

**Tiziana Venittelli** (Ph.D. University of Naples, Parthenope) is Postdoctoral Research Fellow at the University of Naples Federico II. Her research interests are in economic development, education, migration and labor economics.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter'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.



# Degree-Level Determinants of University Student Performance



Massimiliano Bratti, Giovanni Barbato, Daniele Biancardi, Chiara Conti,  
and Matteo Turri

**Abstract** Although features of the higher education degree programmes in which students are enrolled are likely to have an impact on their academic careers, primarily because of data limitations, research has mainly focused on individual, household and higher education institution drivers of student performance. To fill this knowledge gap, this chapter presents a study using administrative data on the complete supply of higher education degrees in Italy during 2013–2018 to carry out an analysis of the degree-programme determinants of university student performance, as measured by the National Agency for the Evaluation of the University System and Research (ANVUR) ‘quality’ indicators. After controlling for detailed degree subject–geographic macro-area fixed effects, our analysis uncovers several significant degree-programme predictors of university student performance, including the degree’s type of access (i.e. selectivity), language of instruction, composition of the teaching body, percentage of teachers in ‘core’ subjects, teachers’ research performance (for master degrees) and university spatial competition.

**Keywords** Teaching quality · University · Degree programme · Quality assurance · Italy

**JEL Codes** I23, I28

---

M. Bratti (✉) · G. Barbato · M. Turri  
University of Milan, Milan, Italy  
e-mail: [massimiliano.bratti@unimi.it](mailto:massimiliano.bratti@unimi.it); [giovanni.barbato@unimi.it](mailto:giovanni.barbato@unimi.it); [matteo.turri@unimi.it](mailto:matteo.turri@unimi.it)

D. Biancardi  
University of Turin, Turin, Italy  
e-mail: [Daniele.BIANCARDI@ec.europa.eu](mailto:Daniele.BIANCARDI@ec.europa.eu)

C. Conti  
University of Rome “La Sapienza”, Rome, Italy  
e-mail: [chiara.conti@uniroma1.it](mailto:chiara.conti@uniroma1.it)

© The Author(s) 2022

D. Checchi et al. (eds.), *Teaching, Research and Academic Careers*,  
[https://doi.org/10.1007/978-3-031-07438-7\\_10](https://doi.org/10.1007/978-3-031-07438-7_10)

## 1 Introduction

Economists have long recognized the importance of investments in human capital and education as fundamental engines of a country's economic growth (e.g. Becker 1994; Barro 2001). Together with quantity, scholars have also more recently stressed the importance of the quality of education in explaining countries' economic performance (Hanushek and Woessmann 2012). At the micro-level as well, there is evidence of positive labour market returns to university quality (McGuinness 2003; Black & Smith 2004; Di Pietro & Cuttillo 2006; Ciani & Mariani 2014; Andrews et al. 2016; Deming et al. 2016; Anelli 2018), with wage premia associated with better university reputations (MacLeod et al. 2017).

Being aware of the importance of having a highly educated workforce, several countries have made attempts to increase the number of university graduates by expanding their higher education supply to attract more students into higher education. Policies such as increasing the geographical diffusion of university branches (Oppedisano 2011) or a complete restructuring of university education—an example being the 'Bologna process' (Bondonio & Berton 2018; Di Pietro & Cuttillo 2008)—have been implemented to reach this goal. Yet, an important hurdle to increasing the number of university graduates remains the high share of students dropping out from higher education. In OECD countries, for instance, 'on average, 12% of students who enter a bachelor's programme full time leave the tertiary system before the beginning of their second year of study. This share increases to 20% by the end of the programme's theoretical duration and to 24% three years later' (OECD 2019, p. 208). Chapters "Do Financial Conditions Play a Role in University Dropout? New Evidence from Administrative Data" and "Drop-Out Decisions in A Cohort of Italian Universities" in this book provide an extensive discussion of the determinants of student dropout and new evidence based on Italian universities. It is clear that an increase in the number of graduates could be achieved by reducing important inefficiencies in higher education systems.

The extant literature has extensively investigated the individual-level determinants of university student progression and academic performance (e.g. school entry qualifications and family background). However, often owing to a lack of data studies accounting for supply-side (i.e. university) characteristics are very rare. Those investigating degree-programme characteristics are even rarer. In the current chapter, we seek to fill this important gap in the academic literature. Leveraging a new and very rich database built by merging information on university performance indicators (PIs) provided by the National Agency for the Evaluation of the University System and Research (ANVUR) with degree-programme-level information gathered within the quality assurance system for higher education (HE hereafter), this study features, to the best of our knowledge, the first comprehensive analysis of the degree-programme determinants of student performance. Our study spans the

complete HE supply in Italy (bachelor's, master's and combined bachelor's/master's degrees) for the 2013–2018 period.<sup>1</sup>

In addition to researchers in the field of higher education, other types of stakeholders are likely to be interested in our analysis as well. First are students and their families, who when making their (or their children's) enrolment choices often focus on a given degree programme rather than on higher education institutions or college majors more broadly. This study provides findings on student dropout and progression that may inform their choices. Second, our study is of interest to all stakeholders that are engaged at different levels in the governance of higher education institutions, such as the heads of degree programmes and quality assurance (QA) groups. Indeed, several countries have introduced complex QA systems to improve the quality of their educational systems (see the next section). In the Italian QA system, for instance, heads of degree programmes and QA groups represent the frontline for interventions to improve the quality and effectiveness of tertiary education. A strong stimulus to improve quality in higher education comes from the diffusion of a quasi-market, which implies that universities are increasingly competing for students and have to devote more attention to the quality of the services they provide and overall student satisfaction compared to the past. Students' enrolment choices are indeed affected by university characteristics, including teaching and research quality (Biancardi & Bratti 2019), and some students are willing to travel long distances in search of better educational opportunities (Baryla & Dotterweich 2001; De Angelis et al. 2017; Bratti & Verzillo 2019). Moreover, the analysis developed in chapter "Drop-Out Decisions in a Cohort of Italian Universities" shows that the abilities of students from outside the town/region of their university are higher than the overall population in terms of high school grades and that these students drop out significantly less than those who study in their hometowns. Although heads of degree programmes and QA groups are equipped with an extensive set of indicators to monitor degrees, only rarely are these systematically analysed as we do in this chapter. Thus, our analysis can be of interest to policymakers needing to take actions to improve the quality of higher education.

This chapter unfolds as follows. The next section discusses the evaluation of teaching activities and the introduction of quality assurance for higher education systems, while Sect. 3 describes the Italian system of quality assurance. Section 4 briefly reviews some key findings from the literature on student progression and dropout. The data and the econometric model used in our empirical analysis are presented in Sect. 5. The main results of our analysis are commented on in Sect. 6, while some robustness checks are presented in Sect. 7. Section 8 summarizes the main findings of this chapter and draws conclusions.

---

<sup>1</sup> The list of abbreviations used in this chapter is presented in the Appendix A. In the Italian context, bachelor's, master's and combined bachelor's/master's degree are *lauree di primo livello*, *lauree di secondo livello* and *lauree a ciclo unico*, respectively.

## 2 The Evaluation of Teaching Activities and Higher Education Quality Assurance (QA) Systems

Teaching quality is an important element for university student performance. Although teaching and research—and more recently, the so-called ‘third mission’—have been traditionally recognized as equally important missions of universities, the evaluation of their activities has developed quite differently over time in terms of rationales and intensity. Several scholars have recognized that the assessment of university activities has been heavily influenced by a striving for research excellence (Dill & Soo 2005). Performance-based funding mechanisms, international rankings, and even the structure of academic careers have consequently been based almost exclusively on the assessment of research performance at both the individual and the institutional level (Horta et al. 2012). Nowadays, a majority of Western European countries (but also countries in other parts of the world) have indeed adopted evaluation exercises or comparable mechanisms to assess research quality (Hicks 2012).

In contrast, the evaluation of teaching activities is younger and almost entirely expressed in the form of accreditation or QA systems, the main function of which is to verify the existence of qualitative standards and requirements through an evaluation procedure that does not affect—at least directly—the amount of public funding that universities receive from national governments. The introduction and diffusion of QA is the result of three main interrelated policy rationales and processes that have occurred, especially in Western Europe, since the late 1980s (Cheng 2015).

First, a ‘steering at a distance’ conception of HE system governance has developed, according to which national governments grant some form of institutional and organizational autonomy in exchange for external control through various mechanisms such as funding and evaluation systems. QA proved to be an instrument of such policies and clearly emerged in countries such as the UK and the Netherlands (Neave & van Vught 1991).

Second, QA has often been introduced as part of new public management (NPM)-based reforms or, more generally, of ‘market-based’ policies (Agasisti et al. 2019). At the system level, the NPM reforms aimed to steer HE systems vertically through agencies, evaluation exercises and budgetary constraints, increasing the universities’ accountability as well as supporting the overall level of competition for resources (Bleiklie & Michelsen 2013). QA systems can thus be seen as a mechanism through which to make the relationship between public funding and the quality of university’s activities more transparent. With the decrease in public funding and increasing competition, QA can also be viewed as a way of demonstrating value for money to those who bear the cost of educational services—in other words, it serves as a consumer protection device (Stensaker 2011). At the institutional level, the NPM reforms supported the introduction of a new management style that strengthened the power of the leadership and executive bodies and, at the same time, decreased the power of collegial bodies. QA was seen as a ‘top-down



managerial device' (Vidovich 2002, p. 397) to make either universities or academics more accountable and to some extent limit and control their historical autonomy and self-governance. QA mechanisms indeed help a university's top management develop clearer lines of responsibility through the definition of minimum-quality standards and their continuous monitoring, as well as the consequent centralization of information (Morley 2003; Stensaker 2008). In this way, QA can support the direction of a university both in terms of resource allocation and in terms of its organizational effort (Jarvis 2014).

Third, it is equally claimed that the spread of QA systems in Europe is also part and parcel of the consequences generated by the Bologna process (Huisman & Westerheijden 2010). In the process of developing a European Higher Education Area (EHEA), the need for a common framework was also translated into the requirement that each country would establish a national system of QA. To this end, the European Network of Quality Assurance Agencies (ENQA) was created, and the European Standards and Guidelines (ESG) were then established to provide general conditions and standards that each national QA system must adopt in relation to both the internal QA of HE providers and the external QA of national agencies (Sin et al. 2017).

Despite the increasing diffusion of QA systems, their effects on teaching and learning performance have not been fully investigated and discussed yet, whereas the literature has stressed some potential unintended consequences. QA practices have been found to be heavily bureaucratic and compliance-oriented processes (Harvey & Newton 2007). Huisman and Westerheijden (2010) claimed that internal QA systems could be considered as a good example of power's idea of 'decoupling' (Power 1997), that is, a buffer complying with the requirements and standards of external evaluation actors by providing verifiable measures that are unrelated to organizational processes. Consequently, it cannot be a coincidence that several recent studies have questioned the actual impact of QA practices on teaching and learning activities. However, these mainly address this issue through the perceptions of academics (Stensaker 2011; Cardoso et al. 2016; Tavares et al. 2017), without going into depth into the actual teaching and learning performance of either students or degree programmes (an exception is, for instance, Andreani et al. 2020).

Finally, a potential unintended consequence of QA mechanisms, and more generally of the evaluation of teaching, is the quantification of quality, as denominated by Kallio et al. (2017, p. 299). In their empirical study on Finnish higher education, they illustrated that 'the easiest way of meeting targets is by lowering standards, for instance, by letting students pass exams more easily and granting degrees with looser criteria.' These 'gaming' phenomena have indeed already been observed in other practices diffused in the public sector, as shown by Christopher and Hood (2006).

### 3 The Italian Higher Education QA System

Although the Italian HE system is one of the largest in Europe, with over 1.6 million enrolled students, more than 300,000 graduates per year and 90 universities, the first extensive QA system was only introduced in 2013 (Ministerial decree 47/2013). Occasional QA practices could be found among Italian universities before 2013, however (Rebora & Turri 2011). These were the result of either the Conference of Engineering Deans, which promoted QA and accreditation practices for engineering degree programmes, or the Conference of the Rectors of Italian Universities (CRUI), which launched accreditation procedures for degree programmes on a voluntary basis.

With the full establishment of the National Agency for the Evaluation of the University System and Research (ANVUR), the QA policy became much more comprehensive and structured. The NPM-based reform of 2010 (Law 240/2010) clearly identified ANVUR as the body in charge of monitoring the effective operation of internal QA procedures by defining quality standards and verifying that these are applied by universities (Agasisti et al. 2019). The QA model denominated AVA (*Autovalutazione, Valutazione Periodica e Accreditamento*, i.e. Self-Evaluation, Periodic Evaluation and Accreditation) is clearly inspired by the European Standards and Guidelines and consists of three interrelated stages. The first is a set of internal QA practices and procedures carried out by universities at the level of both the entire organization and individual degree programmes. Each university is indeed required by law to define its objectives and procedures for quality assurance and improvement and to perform an annual review for each degree programme. Since the internal QA procedure is mainly carried out at the level of the individual degree programme, a major part of the QA process is performed by the head of the degree programme. The internal QA process has to comply with the quality standards established by ANVUR, which provides specific requirements such as, for instance, the involvement of student representatives in the internal QA process.

Second, the external process consists of on-site visits from a group of QA experts and students forming a CEV (Evaluation Expert Committee) and appointed by ANVUR every 5 years. The main output of these CEV visits is an assessment of the compliance of degree programmes (10% of the total number of degree programmes) using the quality standards defined by ANVUR, in order to assess the effectiveness of the internal QA system. A CEV might also decide whether a degree programme needs to undertake corrective actions (within a time limit) if the final rating is not satisfactory. Third, based on the evaluations obtained by on-site visits, ANVUR recommends whether the Ministry of Education, Universities and Research (MIUR) should accredit a university. With the launch of the new ESG in 2015, ANVUR started a review and updating process of the AVA system, which resulted in new guidelines issued in December 2016 (Ministerial decree 987/2016). The review of AVA had multiple goals. First of all, it aimed to reduce the number of quality standards (from 57 to 30) that universities have to be compliant with. Second,

ANVUR also aimed to strengthen the internal self-evaluation of universities before the visits by reducing the number of degree programmes under evaluation. These two goals were particularly important in terms of reducing the potential bureaucratic burden associated with QA procedures.

Finally, ANVUR developed and introduced a list of 37 indicators to support a stronger connection between the outcomes of the internal QA system and the performance of either the entire university or particular degree programmes. The introduction of these indicators was also a way to put students, the learning process and outcomes at the centre of the QA process, instead of QA procedures merely complying with the national legislation (Andreani et al. 2020). These indicators are used by the different internal university actors who participate in the QA process and by ANVUR in the assessment phase that precedes the on-site CEV visit. The 37 indicators are structured into 6 areas and can refer to the level of either the entire university or the degree programme, as well as to both levels (Ministerial decree 6/2019):<sup>2</sup>

- I. Teaching: 9 indicators at the level of the entire university and at the degree-programme level
- II. Internationalization: 3 indicators at the level of the entire university and at the degree-programme level
- III. Environment and quality of research: 5 indicators at the level of the entire university
- IV. Economic and financial sustainability: 3 indicators at the level of the entire university
- V. Further indicators for the evaluation of teaching: 8 indicators at the degree-programme level
- VI. Further pilot indicators: 9 indicators at the degree-programme level

The value of each indicator is computed by ANVUR for three consecutive academic years to facilitate the identification of time trends. Moreover, all indicators also present the average values for other degree programmes that belong to the same scientific area within the same geographical macro-area, as well as at the national level, in order to enable benchmarking exercises. Among these 37 indicators, 29 are clearly connected to the area of learning and teaching (L&T) and belong to the above-mentioned areas, no. I, II, V and VI.

These 29 indicators can be classified according to the four domains of the L&T process and quality, as recognized by the literature (see Leiber 2019 for a framework and a literature review of this topic), namely (i) L&T environment, that is, a framework of conditions and inputs to L&T in terms of organization, staff and students, (ii) teaching processes and competences of teachers, (iii) learning processes and competences of students and (iv) learning outcomes and gains. As claimed by Leiber (2019, p. 79) and in line with ESG (2015), ‘these four constitutive domains should be considered to generate a comprehensive view on

---

<sup>2</sup> A detailed list of ANVUR indicators can be found in the Appendix (Tables D1–D4).

L&T quality issues, because L&T quality of (higher) education is multi-causally determined by the quality of inputs (L&T environment; teaching, learning and assessment competences) as well as the quality of teaching and learning processes and characterized by the quality of outcomes (learning outcomes and learning gain).’

Therefore, through the framework proposed by Leiber (2019), it turns out that the large majority of the AVA teaching indicators are concentrated in the ‘L&T environment’ (10) and ‘learning outcomes’ (18) domains, whereas there is almost an absence of indicators concerning both the ‘teaching competences’ (only 1) and the ‘learning competences’ domains. Indicators such as the ‘proportion of teaching staff who participated in pedagogical training’ and the ‘number of and duration of students’ interactions with course activities’ might indeed become more and more important in supporting the shift in paradigm from teaching to learning represented by the student-centred approach of the ESG (2015). Moreover, the number of indicators related to the ‘learning outcomes’ domain is heavily skewed in favour of metrics regarding the student success rate and the regularity of students’ careers, without covering any aspect of the learning gain process, in other words, the proper achievement and assessment of learning outcomes. In the domain of ‘L&T environment’, there are instead no indicators of the quality of incoming students and the amount of financial investment in L&T.

In the following sections, some of these indicators will be used for our empirical analysis.

## 4 Literature Review on the Determinants of University Student Performance and Dropout

There is an extensive literature on the determinants of university student progression and dropout. The number of studies is too large to summarize all of the findings here.<sup>3</sup> For the sake of space, in this section, we only report some of the key results emerging from the literature.<sup>4</sup>

### *Individual-Level Determinants of Student Dropout and Progression*

Scholars have especially worked on the individual-level determinants of university student progression and the probability of dropout. Among the demographic characteristics significantly associated with student dropout are age, i.e. older students are more likely to drop out (Montmarquette et al. 2001; Smith & Naylor 2001; Stratton et al. 2008)—although in the Italian context this may simply be due to the lower ability of older students, i.e. those who experienced grade retention—and gender,

---

<sup>3</sup> See also chapters “Do Financial Conditions Play a Role in University Dropout? New Evidence from Administrative Data” and “Drop-Out Decisions in A Cohort of Italian Universities” in this book, which analyse student dropout.

<sup>4</sup> For a comprehensive literature review on university student dropout and time-to-degree, see, for instance, Aina et al. (2021).

i.e. female students are less likely to drop out of university (McNabb et al. 2002; Arulampalam et al., 2004a, 2004b; Gury; Cappellari & Lucifora) due to their greater study effort (Stinebrickner & Stinebrickner 2012) and higher returns from education (Goldin et al. 2006). Gender differences also exist in terms of the probability of on-time graduation, although in this case the advantage of women is not ubiquitous in the literature (Häkkinen & Uusitalo 2003; Aina et al. 2011; Lassibille & Navarro Gomez 2011). The results are less clear-cut with regard to ethnicity, with some scholars finding higher dropout rates for students from minority groups (Harvey & Anderson 2005) and others finding a lower dropout rate, especially for those enrolled in selective institutions (Alon & Tienda 2005).

In most studies, dropout is negatively associated with the level of student entry qualifications and ability (Smith & Naylor 2001; Arulampalam et al. 2004a,b; Stratton et al. 2008). Good student achievement in secondary school is also associated with a shorter time to graduation (Aina et al. 2011; Lassibille & Navarro Gomez 2011), although this does not necessarily reflect a causal relation (Bound et al. 2012). Yet, in some studies, students with better entry qualifications are found to be more likely to drop out (DesJardins et al. 1999; Belloc et al. 2010). This reflects the complex nature of student dropout, which is sometimes motivated not by unsatisfactory student performance but by the availability of good opportunities in the labour market or the higher expectations of better students, which may not be met by the study programme they originally choose. Yet, early academic performance—that is, performance in the first years of enrolment—is a powerful determinant of student dropout (Montmarquette et al. 2001; Bennet 2009; Belloc et al. 2010), as students learn about their abilities during the courses and while taking exams (Stinebrickner & Stinebrickner 2014).

The network of relations cultivated during their studies also affects university student dropout behaviour. Indeed, dropout is lower when students have more interactions with professors (Tinto 1975; Pascarella & Terenzini 1978) and peers (Stinebrickner & Stinebrickner 2006), for instance, through study and learning groups (Tinto 1997). Thus, students attending more selective programmes have the additional advantage of benefiting from more able peers (Sacerdote 2011).

A student's family background is an important determinant of his/her probability of dropping out of higher education. Dropout is generally higher for students with a lower socio-economic status (Di Pietro 2004; Johnes & McNabb 2004; Cappellari & Lucifora 2009; Trivellato & Triventi 2009; Aina 2013). In these types of studies, it is generally not possible to disentangle the effect of family income from that of other family characteristics, which would be very relevant policy-wise, however. An exception is Stinebrickner and Stinebrickner (2008), who report that differences between students with different socio-economic statuses persist even in the absence of credit constraints.

According to the human capital model of Gary Becker (Becker 1994), student dropout depends on the opportunity costs of studying, which in turn depend on labour market conditions. Thus, student dropout should decrease in poor labour market conditions, such as during recessions. Results consistent with this prediction are found, for instance, by Di Pietro (2006) and Adamopoulou and Tanzi (2017).

*Institutional Determinants of Student Dropout and Progression*

Interestingly, much less work exists on the institutional determinants of student dropout. Scholars have often focused on system-wide higher education characteristics. Student aid generally contributes to increasing the participation of low-income students in higher education (Dynarski & Scott-Clayton 2013). By relaxing cash constraints, increases in student aid contribute to reducing student dropout (Singell 2004; Arendt 2013) and boost the probability of graduation for disadvantaged students (Alon 2007). However, although the introduction of a strong merit component in student aid (i.e. cut-offs in grade point average (GPA) or university credits to be achieved to maintain aid eligibility) speeds up graduation, on average (Glocker 2011; Scott-Clayton 2011; Gunnes et al. 2013; Denning 2019), it also raises educational inequalities, increasing the probability of dropout for low socio-economic background students and creating an equity–efficiency trade-off (Schudde & Scott-Clayton 2016; Scott-Clayton & Schudde 2020). Financial incentives for good performance are more effective if they are combined with support services for students (Page et al. 2019; Andrews et al. 2020), especially for women (Angrist et al. 2009).

Another important institutional feature of higher education is student fees, i.e. the amount of private vs public funding devoted to higher education. Studies credibly identifying the causal effect of fees on student dropout are in short supply, and scholars have reported mixed results. Bradley and Migali (2019) investigate the effect of the 2006 fee reform that increased university fees in England and, using a difference-in-differences (DIDs) strategy, report opposite effects for high-income and low-income students, whose dropout probabilities fell and increased after the reform, respectively. Conversely, Montalvo (2018) exploits the discontinuity in student fees by student income in a regression discontinuity design (RDD) and finds no adverse effect on student dropout, irrespective of student socio-economic status. A similar strategy has been used by Garibaldi et al. (2012), who show that an increase in tuition fees reduces the probability of late graduation, increasing the efficiency of the educational system.

Other non-monetary institutional features are likely to impact student dropout and performance, such as the quality of facilities and services (tutoring, support etc.) provided to students, the structuring of teaching activities, the type of admission criteria and the characteristics of the teaching body. Ryan (2004) shows that dropout is lower in large universities thanks to the greater availability of services and support that they can provide by exploiting economies of scale. As for admissions criteria, although some scholars have reported lower student dropout rates and a shorter time-to-graduation in systems characterized by stricter admission criteria (Bowen et al. 2009; Bound et al. 2010), this result does not seem to apply to all contexts. Francesconi et al. (2011) leverage a reform introducing selective admissions in a large private university in northern Italy and do not find any improvement in student performance, a result that they relate to the existence of several enrolment alternatives available to students. This finding seems to be confirmed by Carrieri et al. (2015), who instead report positive effects of a similar reform implemented in a public university in southern Italy, in an area where students had very few

alternatives for pursuing university studies. In addition, the way teaching activities, exams and graduation sessions are organized during the academic year affects student performance. On the one hand, Di Pietro and Cutillo (2008) find that the greater flexibility introduced thanks to the Bologna reform (2001) reduced dropout in Italy. On the other hand, a study from Sweden shows that in universities that reduced the number of thesis defence sessions, i.e. reducing flexibility, the time for degree completion fell (Löfgren & Ohlsson 1999).

Following the literature on school class size, researchers have also focused on the impact of student–teacher ratios or other measures of resource intensity at the university level, generally finding positive associations with student performance (Bound & Turner 2007; Bound et al. 2010; Aina et al. 2011; Gitto et al. 2016). Chapter “Teaching Efficiency of Italian Universities: A Conditional Frontier Analysis” in this book provides a test of university teaching efficiency using a similar indicator.

Specifically concerning the working conditions and quality of the teaching staff, Herzog (2006) reports that a higher share of tenure-track (vs temporary) professors is associated with lower student dropout. An important supply-side factor to be taken into account is the quality of teachers (Hanushek & Rivkin 2006; Hanushek et al. 2019). The impact of teaching in higher education institutions (HEIs) on student and graduate performance has been the subject of a recent strand of literature (Laureti et al. 2014; Braga et al. 2016; Brownback & Sadoff 2020) showing positive returns of teaching quality. An interesting finding is that the quality of teachers measured in terms of value added is not always reflected correctly in student teaching evaluations (Braga et al. 2014).

In the current chapter, we seek to contribute to the extant literature by moving the focus to the *degree-level determinants* of student progression, student dropout and levels of student satisfaction. Expanding our knowledge of these issues is key as the first actors called on to implement policies to reduce dropout and improve student progression are degree directors and the QA groups that support them in degree governance. These can never operate on features of the higher education system, which are determined at higher hierarchical levels—for instance, at the level of the higher education institution or even the more aggregated regional or national level (e.g. the amount and forms of student aid, the amount of student fees etc.)—and often require very long periods of time to be changed. Heads of study programmes and the QA group have much more limited policy levers and can often change only small organizational features of the degrees they manage. Thus, assessing the latter’s impact on student progression and satisfaction becomes key for effective policymaking, especially in the short run.

## 5 Data and Empirical Model

In what follows, we describe our empirical model and the main variables used in our empirical analysis, along with the data sources.

## 5.1 Model

We estimate linear regression models specified as follows:

$$y_{it} = \beta_0 + \beta_c \mathbf{C}_{it} + \beta_D \mathbf{D}_{it} + \beta_T \mathbf{T}_{it} + e_{it}, \quad (1)$$

where  $y_{it}$  are the measures of university outputs provided by ANVUR indicators. The explanatory variables are collected in three distinct vectors. The first is a vector  $\mathbf{C}_{it}$  for contextual factors or factors that are beyond the control of each degree programme. The key regressors in this vector are represented by geographic macro-area and detailed degree subject fields (i.e. class of degree) fixed effects.<sup>5</sup> An example of a class of degree is ‘LM-56’, i.e. ‘Master’s Degree in Economic Sciences’. For each degree class, the Ministry of University specifies how the syllabus must be articulated in terms of subject groups covered (SSD, ‘scientific and disciplinary sectors’) and the corresponding number of university ECTS credits.<sup>6</sup> Degree programmes in the same degree class and geographic macro-area are indeed the benchmark against which heads of degree programmes and the QA group are called on to compare the performance of their degrees. These fixed effects capture the average differences in PIs that are geographic- and subject-group-specific; a second vector ( $\mathbf{D}_{it}$ ) collects degree-programme features that do not pertain to the teaching body. They include the type of student admissions, the teaching language, the multidisciplinary character of the degree, the size of the QA group and the intensity of spatial competition; a final vector ( $\mathbf{T}_{it}$ ) of regressors collects characteristics of the teaching body such as the percentage of teachers by academic position, the number of students per teacher-tutor, the percentage of teachers in ‘core’ subjects (SSD) and the research evaluation of the teaching body measured by the most recent Research Assessment Exercise (*Valutazione Qualità della Ricerca*, VQR). Finally,  $e_{it}$  is a degree-specific error term.

To allow for degree-level specificities, the models are estimated separately for bachelor’s degrees, master’s degrees and combined bachelor’s/master’s degrees.

In the following sections, we explain the main dependent and explanatory variables included in Eq. (1) and the rationale for their inclusion.

---

<sup>5</sup> However, some of these factors are still under the control of universities, which can decide in what subject group to open degree programmes.

<sup>6</sup> European Credit Transfer and Accumulation System (ECTS) credits are a standard means of comparing the ‘volume of learning based on the defined learning outcomes and their associated workload’ (ECTS user guide) for higher education across the European Union and other collaborating European countries. One academic year corresponds to 60 ECTS credits, which is normally equivalent to a total workload of 1500–1800 h.



## 5.2 *Dependent Variables*

In the regression models, the following ANVUR indicators are used as dependent variables, and the types of degrees for which they are available are indicated in parentheses (BA = bachelor's degree; MA = master's degree; BA+MA = combined bachelor's/master's degree). For each indicator, we report the original ANVUR name and the name with which we will refer it to in the analysis (e.g. in tables):

### *Student Progression*

*iC01* (ECTS40): percentage of regularly attending students who have earned at least 40 ECTS credits during the academic year (BA, MA, BA+MA)

*iC13* (*ECTS*<sub>1</sub>): percentage of ECTS credits achieved in the first year over the total ECTS credits to be achieved in the first year (BA, MA, BA+MA)

*iC15* (*ECTS*<sub>20</sub><sub>1</sub>): percentage of students who continue on to the second year in the same degree programme having earned at least 20 ECTS credits in the first year (BA, MA, BA+MA)

*iC16* (*ECTS*<sub>40</sub><sub>1</sub>): percentage of students who continue on to the second year in the same degree programme having earned at least 40 ECTS credits in the first year (BA, MA, BA+MA)

*iC02* (GRAD): percentage of graduates within the legal programme duration (BA, MA, BA+MA)

### *Student Satisfaction*

*iC18* (ENR): percentage of graduates who would enrol in the same degree programme again (BA, MA, BA+MA)

*iC25* (SATI): percentage of students who are generally satisfied with their degree programme (BA, MA, BA+MA)

In our opinion, these are the ANVUR indicators that can be more strictly considered as degree-programme outputs related to student dropout and progression and overall levels of student satisfaction. Other ANVUR teaching indicators are related to features of the teaching body engaged in each degree programme, such as the percentage of teaching hours taught by personnel with open-ended contracts, and are included as explanatory variables in the econometric models (see the next section). The source of these data is ANVUR, who provided us with indicators for the 2013–2018 period for the purpose of the current research.

In Fig. 1, we plot the raw geographical differences for the first indicator of progression—the percentage of regularly attending students who have earned at least 40 ECTS credits during the academic year. A clear North–South divide emerges for progression. Students enrolled in higher education institutions located in the North are much more likely to have completed at least 40 ECTS credits during the academic year (see Table B1 in the Appendix B for the means and standard deviations of all degree performance indicators by macro-area). A different picture emerges from Fig. 2, which shows the geographical distribution for student

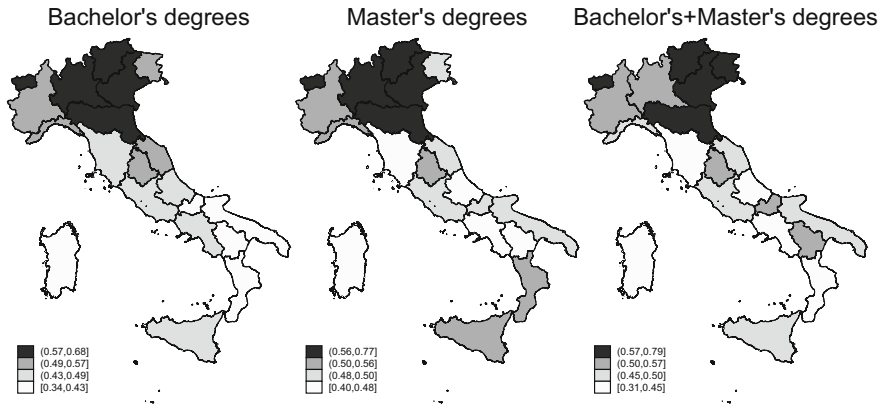


Fig. 1 Student progression

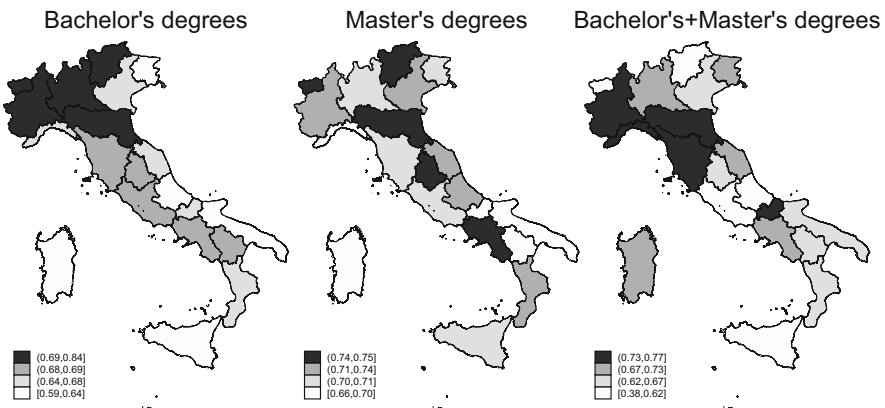


Fig. 2 Student satisfaction

satisfaction with their chosen degree—measured as the percentage of graduates who would enrol again in the same degree programme at their university. We cannot identify a clear pattern between regions in different macro-areas of the country in this case, but we observe large differences across regions by degree level. For example, students who enrolled in bachelor's degrees in Trentino-Alto Adige are much more satisfied with their choice than students enrolled in combined bachelor's/master's degrees in the same region. The only region that is always ranked at the top for student satisfaction is Emilia-Romagna.<sup>7</sup>

<sup>7</sup> We acknowledge that in the model estimation there are some timing issues. Indeed, our models are estimated using contemporaneous measures of the dependent and explanatory variables. On the grounds that degree-programme features (e.g. language of instruction, type of access etc.) or the composition of the teaching body frequently change, we expect our regressors to be

### 5.3 Explanatory Variables

The explanatory variables used in the empirical analysis come from two main sources. The first is ANVUR, for the indicators that are used as degree-programme-level inputs. The second source is the degree-programme cards (namely, *Scheda SUA-CdS*), the completion of which is made mandatory by the national system of QA and which gather a wealth of information on degree programmes.<sup>8</sup>

*Contextual Factors* As we anticipated, since comparisons of ANVUR indicators should be made with degrees in the same degree class and geographic macro-area, we include interaction terms defined at this level in the models (i.e. degree class group by geographic macro-area level). The four macro-areas are North-West, North-East, Centre, South and Islands (*area4* variable). Including these fixed effects purges the ANVUR output indicators of factors that depend on the degree subject or geography, namely the geographical location of the university branch supplying the degree. The degree class is provided by SUA-CdS cards and the macro-area by ANVUR. Degree classes are aggregated into the following 14 groups using the classification provided by the National University Council (CUN): Mathematics and Informatics; Physics; Chemistry; Earth Sciences; Biology; Medicine; Agricultural and Veterinary Sciences; Civil Engineering and Architecture; Industrial and Information Engineering; Antiquities, Philology, Literary Studies, Art History; History, Philosophy, Pedagogy and Psychology; Law Studies; Economics and Statistics; Political and Social Sciences.

We have computed a measure of the potential level of spatial competition for each degree programme, namely the number of programmes in the same broad degree subject and geographic macro-area. Previous research by Cattaneo et al. (2017) has shown that competition among universities, measured through geographical proximity and similarity of educational supply (in terms of subject groups), affects the number of student enrolments. Similarly, we might expect better incentives towards improvement in highly competitive geographical contexts.

*Non-teaching-Personnel Degree-Programme Features* Selective admission degrees are more likely to perform better than non-selective degrees in several dimensions. Academic preparedness is indeed a key determinant of both student dropout and academic progression (Arulampalam et al., 2004a, 2004b; Jia and Maloney 2015). On top of ‘cream-skimming effects’, the concentration of more able individuals induced by selective admission policies may also spur positive peer group effects (see the literature review above). For this reason, we include dichotomous indicators

---

affected by substantial measurement error, which may create an attenuation bias in the relations of interest. These issues are probably more severe for the outcomes that are observed with a delay, namely graduation time and satisfaction, than for first-year progression indicators. Lagging the independent variables is not feasible given the short time span covered by our data (e.g. we would lose three years of data for first-level degrees).

<sup>8</sup> These data are publicly available on the University website, <https://www.universitaly.it/>.

for the type of admission, namely, programmed number degrees (or *numerus clausus* for brevity),<sup>9</sup> and entry requirement assessment, respectively, while the control group is open admission degrees. In the first case, there is a fixed maximum number of students who can access a given degree. In the second case, the number is not fixed, but entry requirements are assessed through a test or an application package and an interview. While in many cases for bachelor's degrees, these entry requirements (e.g. the score on an entry test) are not binding for students, i.e. if they do not meet the entry requirements, they can still enrol in the degree with some academic debits, in master's degrees they are binding, entailing very different policies for the two levels of degrees.

A second proxy of degree-programme selectivity is the official teaching language. We include in the model an indicator for degrees completely or partially taught in English. This is a feature that we expect to potentially affect not only student progression—since degrees using English as the language of instruction attract better students, on average—but potentially also the satisfaction associated with the degree. International degrees may indeed attract more foreign students and enhance the university experience.

Another dimension we consider is the degree of multidisciplinary of degree programmes. This degree feature is captured by an indicator for inter-class degrees, that is, degrees spanning different degree classes. It captures the potential advantages/disadvantages of knowledge and curriculum specialization vs diversification. These degrees generally require student proficiency in quite different subjects.

Aspects related to the degree programme's governance may also affect performance. Agasisti et al. (2019), for instance, demonstrate that the composition and the role of the quality insurance committee (QAC) instituted at the university level affects the success of higher education institutions in pursuing effective quality assurance policies. Since we do not have variables measuring the volume of activity of the QA group of each degree programme, or its composition, we use the size of the QA group (i.e. the number of participating teachers and students) as a proxy. The QA group includes the members of the Review Team (*Gruppo del Riesame*).

These data come from SUA-CdS.

*Teaching Personnel Degree-Programme Features* These are variables capturing characteristics of the teaching body that can potentially affect teaching quality.

We include the percentage of personnel in each academic role, and more specifically the percentage of full professors, associate professors, open-ended researchers, temporary researchers (both in tenure track and not in tenure track), other teaching personnel and of external teachers. At first glance, it is not clear what to expect. On the one hand, more experienced teachers (typically full and associate professors)

---

<sup>9</sup> There are two types of programmed number degrees, those with programmed numbers determined at the national level (mainly Medicine and Surgery, Dentistry, Veterinary, Medical Professions, Architecture, Primary Education Sciences degrees) and those with numbers determined at the local level.

may be better teachers thanks to learning by doing, and this could positively affect all output indicators. Moreover, junior personnel are rarely formally assessed on teaching quality but more often on research performance, and for this reason as well we might expect little focus on teaching (De Philippis 2021). On the other hand, full professors in particular have fewer career concerns and are often more engaged in paid consultancy outside of the university compared to junior personnel, and the time and effort they devote both to teaching and research activities may be lower than that of the latter (Muscio et al. 2017). Furthermore, junior personnel may be more aware of the recent developments in the profession/subject, which can be incorporated into their teaching, while more senior teachers may use outdated syllabi. For this reason, the sign of the relation between seniority (or academic qualification, in our case) and teaching quality is ambiguous and must be empirically assessed.<sup>10</sup>

Another proxy of teaching quality is the consistency between the scientific sectors (SSD) in which teachers are recruited (mainly corresponding to their research field) and the scientific sector of the course they teach. We might expect a positive effect in virtue of the strong specialization of academic knowledge and the potential complementarity between research and teaching activities, especially in master's and combined bachelor's/master's degrees. ANVUR provides a useful indicator for this purpose: the percentage of structured (i.e. non-temporary) teaching personnel that belong to scientific sectors that are core or characterize the degree programme for which they are 'reference teachers'. Indeed, the Ministry of University requires a given number of teachers to be considered as reference teachers in each degree programme. The main rationale is to prevent an excessive expansion and fragmentation of the higher education supply. Having reference teachers that are not in the main subject fields of a degree may imply a bad match between the teaching staff and the content they have to teach.

For master's degrees, ANVUR provides an interesting indicator that allows for a direct test of potential research–teaching complementarity (see, for instance, De Philippis 2021; Rodríguez & Rubio 2016; Artés et al. 2017, and Palali et al. 2018). This is the summation of the indicator of research quality of the university for each SSD assessed through the last Research Evaluation Exercise (VQR 2011–2014), where each SSD is weighted by the ECTS of courses included in the degree-programme syllabus. It is worth mentioning that this is not an indicator of the research performance of personnel providing teaching services in the given degree programme, which is not provided by ANVUR, but rather the average research performance of teachers in the scientific sectors prevailing in that programme.

---

<sup>10</sup> Figlio et al. (2015), for instance, find that first-year students at Northwestern University learn more from contingent teachers than from tenure-track/tenured professors and relate this to the fact that the bottom quarter of the tenure-track/tenured faculty has lower 'value added' than their contingent counterparts. Similar findings are reported for China by Tian et al. (2019).

#### 5.4 *Creation of the Linked ANVUR-SUA CdS Database*

The dataset used in the empirical analysis was built starting from a dataset of ANVUR indicators (provided by MIUR) and containing information on 37 indicators at the level of the degree programme (*corso di studi*, CdS), from which we selected our variables related to degree-programme performance in terms of student progression and time-to-degree. In the original dataset, ANVUR indicators are available for 2290 degree programmes in 92 Italian universities (public, private and online) over the 2013–2018 period.

These data have been merged with those extracted from the degree-programme cards (SUA-CdS), a set of information at the degree-programme level regarding the structure and characteristics of each degree (e.g. duration, procedure for admission, language of teaching, academic staff, tutors, persons in charge of the quality process etc.) used to create the main explanatory variables. Unfortunately, the two sets of data identify degree programmes using different coding systems, and thus we use the triplet name of university—name of degree programme—duration to match the observations. In a few cases—both in ANVUR and in the SUA-CdS data—we found multiple observations for the triplet (17.5% in the ANVUR data, mainly due to several university branches observed in some degree programmes (CdS), especially in the field of medicine and nursing, and 2.3% in SUA-CdS, in most cases due to a ‘double’ version of the same CdS, e.g. in Italian and in English), which did not allow for a one-to-one match. We decided to handle this problem by identifying the CdS code associated with the largest number of students enrolled (i.e. the ‘head branch’) for the ANVUR indicators, that is, the highest value of the ANVUR indicator ‘iC00d’, and then keeping only the observations identified by the selected code. To the univocal triplet name of university–name of CdS–duration in the ANVUR data we linked data from SUA-CdS, keeping all of the (few) multiple observations. Overall, we managed to find correspondences for almost 96% of univocal observations from the ANVUR dataset. Unmatched data mainly refer to degree programmes that are not present in all of the years spanned by the dataset and are probably affected by changes in the educational supply of universities over time. We merged to this dataset some variables related to the type of high school attended (school track) by new enrolled students and their final secondary school grade, to be used as control variables. The latter were provided by the Ministry of University and are built using data from the National University Student Registry (ANS, *Anagrafe Nazionale degli Studenti e dei Laureati*).

This merged database is used in our empirical analysis. Sample summary statistics for the merged ANVUR-SUA CdS database are shown in Table 1. The summary statistics do not necessarily correspond to those in the estimation samples, the composition of which varies according to the dependent variables.

**Table 1** Sample summary statistics

Variables	Bachelor's degrees		Master's degrees		Combined Bachelor's /Master's degrees	
	Mean	SD	Mean	SD	Mean	SD
<i>Dependent variables</i>						
iC01 (ECTS40)	0.50	0.22	0.52	0.19	0.54	0.19
iC13 (ECTS <sub>1</sub> )	0.45	0.27	0.54	0.29	0.52	0.30
iC15 (ECTS20 <sub>1</sub> )	0.52	0.30	0.67	0.34	0.60	0.34
iC16 (ECTS40 <sub>1</sub> )	0.36	0.27	0.42	0.28	0.43	0.31
iC02 (GRAD)	0.51	0.23	0.58	0.25	0.49	0.26
iC18 (ENR)	0.67	0.18	0.72	0.18	0.67	0.18
iC25 (SATI)	0.86	0.16	0.86	0.18	0.86	0.17
<i>Explanatory variables</i>						
Type of access: <i>numerus clausus</i>	0.21	0.41	0.04	0.19	0.54	0.50
Type of access: entry test	0.29	0.45	0.14	0.35	0.24	0.43
English or multi-language degree	0.01	0.09	0.11	0.31	0.02	0.13
Inter-class degree	0.02	0.15	0.03	0.18		
Percentage of PO	0.16	0.13	0.23	0.15	0.21	0.15
Percentage of PA	0.27	0.16	0.31	0.16	0.30	0.16
Percentage of RD	0.04	0.08	0.04	0.08	0.03	0.06
Percentage of RU	0.22	0.15	0.20	0.14	0.22	0.14
Percentage of other	0.01	0.05	0.01	0.05	0.01	0.04
Percentage of external	0.29	0.23	0.21	0.19	0.23	0.22
Size of QA group	5.54	2.54	5.28	2.29	6.12	2.98
No. of students per tutor (/10, top coded)	2.90	3.76	1.76	2.12	2.19	3.01
Percent of teachers in core or characterizing SSD	0.94	0.12	0.89	0.16	0.98	0.05
VQR score teachers			0.98	0.21		
No. of degrees in the CUN area and area4 (/10)	5.26	4.33	3.78	2.27	0.62	0.69
Academic track (%)	0.59	0.19	0.64	0.18	0.74	0.17
Secondary school mark 90 (%)	0.21	0.13	0.28	0.16	0.32	0.16

*Note:* SD stands for standard deviation. CUN areas are broadly defined subject groups and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. The sample summary statistics are only reported for the estimation sample in column (1) of Table 2 and Table 3 for bachelor's and master's degrees, respectively. Summary statistics for other samples are similar. See Sect. 5 for variable definitions

## 6 Results

In this section, we comment on the main results of our regression analysis by level of degree (bachelor's, master's and combined bachelor's/master's).

## 6.1 Student Progression and Time-to-Degree

The estimates of the models of student progression for bachelor's degrees are reported in Table 2. Our main finding is that lower shares of junior personnel and more tutors are associated with a faster progression of students. Furthermore, research quality is positively correlated with student progression in master's degrees.

Selective admission degree programmes generally display better progression indicators. The percentage of regular (i.e. regularly attending) students achieving at least 40 ECTS per year is 25 percentage points higher in degrees with *numerus clausus*. However, no other advantage emerges for the other performance indicators, which may be partly due to the low number of degrees with this type of admission. A possible caveat is that the number of students admitted may be below the *numerus clausus*, which is therefore non-binding. In such a case, there would be no difference with courses assessing the entry requirements. Unfortunately, SUA-CdS data do not provide additional information on the number of places available to measure whether or not and to what extent it is binding. In contrast, in degree programmes with access subject to a non-binding entry test or entry requirements, regular students are 7.3 percentage points (pp hereafter) more likely to pass at least 40 ECTS during the academic year than students in open-access degrees, an advantage that is also displayed when focusing on first-year students only (6.6 pp). Similarly, the probability of passing at least 20 ECTS in the first year is 6.4 pp higher. The average percentage of ECTS passed over the total number of ECTS to be passed in the first year is 6.9 pp higher. Thus, our analysis shows that entry tests assessing student preparedness—even when they are not binding for enrolment—may convey valuable information by signalling to potential students whether they are making the right choice and are indeed always positively associated with better student progression indicators. Yet, quite surprisingly, students in degrees with entry tests are less likely to graduate in the normal duration: the percentage of graduates within the legal duration is 1.8 pp lower. A similar penalty in graduation time is found for degrees with *numerus clausus*, although it is statistically non-significant. A possible explanation is that degrees with entry tests are also more academically demanding than open-access degrees. As a consequence, students may repeat exams to increase their GPA (indeed, in Italy, there are several exam sessions—generally 5–6 per year—and students can refuse grades and retake exams if they are not happy with their exam results). Alternatively, they can devote more time to their final thesis, in an attempt to increase their GPA. In both cases, this would lead to an increase in the graduation time. Unfortunately, our data do not allow us to test these hypotheses.

As expected, degree programmes taught in English perform better in all student progression indicators. The advantages with respect to degrees using Italian as the language of instruction are sizable: 16.7 pp for the percentage of regular students achieving at least 40 ECTS, 17.8 pp for the percentage of ECTS achieved in the first year over the total number achievable, 13 (18.8) pp in the percentage of students achieving at least 20 (40) ECTS in the first year and a 18.7 pp higher probability of



Table 2 Student progression—bachelor's degrees

	(1)	(2)	(3)	(4)	(5)
Variables	iC01 ECTS40	iC13 ECTS <sub>1</sub>	iC15 ECTS20 <sub>1</sub>	iC16 ECTS40 <sub>1</sub>	iC02 GRAD
Type of access: <i>numerus clausus</i>	0.247* (0.135)	0.054 (0.126)	-0.033 (0.135)	0.091 (0.167)	-0.076 (0.175)
Type of access: entry test	0.073*** (0.004)	0.069*** (0.004)	0.064*** (0.004)	0.066*** (0.005)	-0.018*** (0.005)
English or multi-language degree	0.167*** (0.016)	0.178*** (0.016)	0.130*** (0.016)	0.188*** (0.020)	0.187*** (0.022)
Inter-class degree	0.018 (0.012)	-0.002 (0.012)	-0.011 (0.013)	0.005 (0.016)	0.010 (0.016)
Percentage of PO	-0.073*** (0.013)	-0.020 (0.012)	-0.013 (0.013)	-0.017 (0.016)	-0.119*** (0.016)
Percentage of PA	-0.051*** (0.011)	-0.042*** (0.010)	-0.048*** (0.011)	-0.059*** (0.013)	-0.132*** (0.013)
Percentage of RD	-0.112*** (0.020)	-0.056*** (0.019)	-0.120*** (0.021)	-0.076*** (0.025)	0.047* (0.024)
Percentage of RU	-0.086*** (0.012)	-0.058*** (0.011)	-0.053*** (0.012)	-0.081*** (0.015)	-0.116*** (0.015)
Percentage of other	-0.411*** (0.033)	-0.329*** (0.031)	-0.422*** (0.033)	-0.359*** (0.041)	0.058 (0.037)
Size of QA group	-0.002*** (0.001)	-0.001* (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.005*** (0.001)

(continued)

Table 2 (continued)

	(1)	(2)	(3)	(4)	(5)
Variables	iC01 ECTS40	iC13 ECTS <sub>1</sub>	iC15 ECTS20 <sub>1</sub>	iC16 ECTS40 <sub>1</sub>	iC02 GRAD
No. of students per tutor (/10, top coded)	-0.002*** (0.000)	-0.001** (0.000)	-0.002*** (0.000)	-0.002*** (0.001)	0.001* (0.001)
Percent of teachers in core or characterizing SSD	-0.001 (0.017)	0.013 (0.016)	0.013 (0.017)	0.039* (0.021)	-0.069*** (0.021)
No. of degrees in the CUN area and area4 (/10)	0.027*** (0.008)	0.025*** (0.008)	0.036*** (0.008)	0.035*** (0.010)	0.008 (0.008)
Observations	9143	9057	9080	9079	10,322
R-squared	0.630	0.794	0.808	0.642	0.410
Degree class × <i>area4</i> FEs	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES
Mean of dep. var.	0.505	0.449	0.517	0.356	0.510

Note: \*, \*\* and \*\*\* indicates statistical significance at the 10%, 5% and 1% levels, respectively. Omitted reference categories for categorical variables are open-access degree programmes, for type of admission, and the percentage of external personnel, for the composition of the teaching body. Dependent variables are indicated in the column headings (the original name of the ANVUR indicator and the name used in our analysis). CUN areas are broadly defined subject groups, and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. VQR stands for the 2011–2014 research evaluation exercise; PO, PA, RU, RD, and other stand for full, associate, open-ended researchers, temporary researchers and other teaching personnel (e.g. PhD students, postdocs etc.), respectively

graduating on time. As already mentioned, this may be related to the better academic preparedness and motivation of students enrolled in degrees taught in English.

The composition of the teaching body is also significantly related to student progression. Compared to the current composition, increasing the share of full professors by 10 pp (reducing the percentage of external personnel) is associated with a decrease of 0.73 pp in the percentage of regular students achieving at least 40 ECTS. A similar significant penalty is observed for only one another indicator of student progression, namely the percentage of students graduating on time, which is 1.19 pp lower. Negative gaps are also associated with the percentages of associate professors and both temporary and open-ended researchers. The only differences between the latter two groups are a smaller negative coefficient for the percentage of students obtaining at least 20 ECTS (in the first year) for open-ended researchers, and a larger gap for on-time graduation, for which the coefficient of temporary researchers (compared to external personnel) becomes positive. Quite remarkable are the performance penalties suffered by degree programmes that employ more personnel classified in the residual 'other' category, including, *inter alia*, junior personnel such as PhD students and postdocs. Increasing the percentage accounted for by this last group by 10 pp is associated with a 4.11, 3.29, 4.22 and 3.59 pp decrease in the percentages of regular students obtaining at least 40 ECTS, in the number of ECTS over the total achievable in the year, and the percentage of students obtaining at least 20 ECTS and 40 ECTS in the first year, respectively. Without any further supporting information, it is difficult to interpret these coefficients. As for the penalty associated with junior personnel, a potential explanation could be the lack of teaching experience and adequate incentives or motivation to teach, since at least in the first stages of their careers, tenure and promotion mainly depend on research performance. Often times, this junior staff is employed full-time on research and only teaches to integrate their income or because more senior staff ask them to do teaching support activities. On average, it would not be too surprising to find that motivation to teach for this specific group can be rather low (especially for PhD students who still have to complete their studies). A similar argument holds for temporary researchers, who are still under 'probation' and whose likelihood of entering tenure tracks and achieving tenure mainly depends on their research activities. As for the better student progression associated with external personnel, it is difficult to find a clear-cut interpretation since it includes both teaching personnel from other universities and professionals, who could possess very different levels of experience and motivation for teaching. A possible interpretation of their positive results on student progression is that since they do it on a voluntary basis and receive teaching contracts, they may be more motivated, perhaps also because of the extra money they receive for teaching. Good performance in teaching may also be a pre-condition for the renewal of their contract. On a more negative note, external personnel may be less interested in how much students learn and may apply lower standards to reduce their workload (e.g. time to mark exams), increasing in this way the pace of student progression.

The size of the QA group is negatively associated with the percentage of regular students achieving at least 40 ECTS ( $-0.2$  pp for a one-unit increase), the percentage

of first-year credits achieved ( $-0.1$  pp for a one-unit increase) and the percentage of on-time graduates ( $-0.5$  pp for a one-unit increase). A possible interpretation is that making the group larger produces a dilution of the individual effort and responsibility, creating incentives to free ride, or greater group heterogeneity could make it more difficult to have a clear orientation of governance (coordination problems). Other interpretations are possible, however. Given individual time constraints, time devoted to administration by members of the QA group is subtracted from research and teaching, therefore potentially penalizing student results. Moreover, the negative association may also capture reverse causation because heads of degree programmes not performing well may devote more staff to QA activities.

Tutor teachers seem to be a useful resource for improving student progression. The number of students per tutor teacher is negatively associated with all indicators except on-time graduation. Penalties associated with a 10-student increase vary within the range of  $-0.2$  pp and  $-0.1$  pp depending on the indicator, with an unexpected marginally significant positive association of  $0.1$  pp for the percentage of students graduating on time. The negative coefficient for the number of students per tutor teacher confirms that student support is effective and advisable.

Quite surprisingly, the percentage of 'reference' teachers in core or characterizing SSD for a degree programme turns out to be negatively associated with on-time graduation. This may reflect higher teaching standards (e.g. higher exam fail rates) applied to these courses compared to ancillary courses for the particular degree programme, which do not affect first-year progression but still have an effect on the time needed for degree completion.

The intensity of spatial competition seems to be positively associated with student progression, with gains of 2.7, 2.5, 3.6 and 3.5 pp in the percentages of regular students obtaining at least 40 ECTS in the academic year, and of students obtaining 40 and 20 ECTS in the first year, respectively, for a 10-unit increase in the number of courses in the same CUN subject group, duration (i.e. degree level) and geographic macro-area.

Table 3 reports the estimates for master's degrees. Many effects are consistent with those found for bachelor's degrees, and are not commented here.

Similar to what we found for bachelor's degrees, degree programmes with an assessment of entry requirements seem to perform better in all progression indicators except graduation time, compared to degrees with a *numerus clausus* (the comparison group). The premia are 4.2, 4, 2.5 and 4.9 pp on the percentages of regular students achieving at least 40 ECTS in the academic year, and of students obtaining 20 and 40 ECTS in the first year, respectively. In contrast, degrees with a programmed number have a 1.4 pp higher percentage of students graduating on time. However, unlike for bachelor's degrees, entry requirements are generally binding for master's degrees, so a *numerus clausus* approach means higher selectivity only on the grounds that the number of students willing to enrol exceeds the programmed number. Another possible difference is that admission to selective degrees is generally made on the basis of standardized tests, given the very high number of applicants (programmed numbers are generally introduced because demand is much higher than the number of places available), so as to make the

**Table 3** Student progression—master’s degrees

Variables	(1)	(2)	(3)	(4)	(5)
	iC01 ECTS40	iC13 ECTS <sub>1</sub>	iC15 ETCS20 <sub>1</sub>	iC16 ECTS40 <sub>1</sub>	iC02 GRAD
Type of access: entry test	0.042*** (0.005)	0.040*** (0.004)	0.025*** (0.004)	0.049*** (0.006)	-0.014** (0.007)
English or multi-language degree	0.070*** (0.006)	0.057*** (0.005)	0.027*** (0.005)	0.077*** (0.007)	0.090*** (0.007)
Inter-class degree	0.030** (0.012)	0.025** (0.010)	0.007 (0.010)	0.032** (0.014)	-0.004 (0.015)
Percentage of PO	-0.010 (0.014)	0.003 (0.011)	0.001 (0.011)	-0.031* (0.016)	-0.132*** (0.017)
Percentage of PA	-0.038*** (0.013)	0.003 (0.011)	0.009 (0.010)	-0.007 (0.015)	-0.101*** (0.015)
Percentage of RD	-0.070*** (0.024)	0.027 (0.020)	-0.049** (0.020)	0.036 (0.029)	0.076** (0.030)
Percentage of RU	-0.060*** (0.015)	-0.046*** (0.012)	-0.006 (0.012)	-0.055*** (0.018)	-0.098*** (0.019)
Percentage of other	-0.189*** (0.036)	-0.192*** (0.030)	-0.321*** (0.029)	-0.161*** (0.043)	0.402*** (0.046)
Size of QA group	-0.001 (0.001)	-0.001** (0.001)	0.000 (0.001)	0.000 (0.001)	-0.007*** (0.001)
No. of students per tutor (/10, top coded)	-0.004*** (0.001)	0.000 (0.001)	-0.002** (0.001)	-0.000 (0.001)	-0.003*** (0.001)

(continued)

Table 3 (continued)

	(1)	(2)	(3)	(4)	(5)
Variables	iC01 ECTS40	iC13 ECTS <sub>1</sub>	iC15 ECTS20 <sub>1</sub>	iC16 ECTS40 <sub>1</sub>	iC02 GRAD
Percent of teachers in core or characterizing SSD	-0.021 (0.013)	-0.017 (0.011)	0.021** (0.010)	-0.016 (0.015)	-0.038** (0.016)
VQR score teachers	0.094*** (0.010)	0.061*** (0.008)	0.060*** (0.008)	0.103*** (0.011)	0.047*** (0.012)
No. of degrees in same CUN area and area4 (/10)	-0.000 (0.007)	0.003 (0.005)	-0.010* (0.005)	0.013* (0.008)	0.007 (0.007)
Observations	8889	8837	8884	8882	10,408
R-squared	0.432	0.838	0.891	0.658	0.455
Degree class × area4 FEs	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES
Mean of dep. var.	0.520	0.536	0.673	0.424	0.576

Note: \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1%, levels, respectively. Omitted reference categories for categorical variables are *numerus clausus* degree programmes, for type of admission, and the percentage of external personnel, for the composition of the teaching body. Dependent variables are indicated in the column headings (the original name of the ANVUR indicator and the name used in our analysis). CUN areas are broadly defined subject groups, and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. VQR stands for the 2011–2014 research evaluation exercise; PO, PA, RU, RD and other stand for full, associate, open-ended researchers, temporary researchers and other teaching personnel (e.g. PhD students, postdocs etc.), respectively

selection process less cumbersome. In contrast, in courses featuring the assessment of entry requirements, selection is often based on the evaluation of application packages and interviews. Thus, the differences in the outcomes for the two types of degree programmes may simply signal that selection based on standardized tests is less able to screen for the best potential students. An equally plausible explanation, however, is that more selective degrees are tougher.

In master's degrees as well, degree programmes taught in English fare much better than programmes taught in Italian on all performance indicators. Not surprisingly, effect magnitudes are smaller than those observed in bachelor's degrees, since students have already undergone a process of selection during their undergraduate education, but they are still remarkable. To take just a few examples, degrees taught in English have a 7.7 pp higher percentage of students obtaining at least 40 ECTS in the first year and a 9 pp higher percentage of on-time graduates.

Multidisciplinary seems to pay in terms of student progression. Inter-class degrees have an advantage of 3, 2.5 and 3.2 pp for the percentage of regular students with at least 40 ECTS, the percentage of ECTS achieved over the total in the academic year, and of students with at least 40 ECTS in the first year, respectively. A possible explanation is that given the peculiarity of these degrees, which require heterogeneous interests and abilities, students enrolled in these programmes may be highly motivated.

As we observed for bachelor's degrees, the composition of the teaching body is important for student progression in master's degrees as well. Significant penalties for some categories of structured personnel emerge compared to external personnel. The largest penalties are associated with the 'other' category, which displays a positive premium on the percentage of on-time graduation, however (4 pp associated with a 10-pp increase in the percentage of 'other' personnel), a positive association with graduation time that is shared with the group of temporary researchers (0.76 pp associated with a 10-pp increase in the percentage of temporary researchers). We have already commented on the possible explanations for these effects for bachelor's degrees.

The negative association between the size of the QA group and the indicators of student progression that was observed for bachelor's degrees is confirmed for master's degrees, at least for the percentage of ECTS achieved (a 0.1 pp decrease for a one-unit increase in the QA group) and for on-time graduation (-0.7 pp).

The analysis for master's degrees also confirms the valuable role of tutor teachers. A higher student-tutor teacher ratio is associated with slower student progression, with significant effects on the percentage of regular students obtaining at least 40 ECTS, obtaining at least 20 ECTS in the first year and graduating on time.

We find ambiguous results for the percentage of 'reference' teachers in core or characterizing SSD: although it is positively associated with the percentage of students achieving at least 20 ECTS in the first year, it is negatively associated with the percentage of students completing their studies within the normal duration (-0.38 pp associated with a 10-pp increase in the explanatory variable).

Our results point to some form of complementarity between teaching quality and research quality, as measured by the research assessment (VQR) results of the personnel teaching in a degree programme. A one-point increase in the VQR score (in our estimation sample ranging between zero and 1.81) is associated with increases of 9.4, 6.1, 6, 10.3 and 4.7 pp in the percentage of regular students achieving at least 40 ECTS per year, of achieved ECTS over the total in the first year, of students achieving at least 20 and 40 ECTS in the first year and of on-time graduates, respectively. This association is further explored in chapter “The Relationship Between Teaching and Research in the Italian University System” of this book. These are intriguing results that deserve further investigation, possibly using individual-level data. Indeed, as we have stressed in the literature review, results on the complementarity between university teaching and research are quite mixed and research on this issue is still sparse.

Unlike for bachelor’s degrees, for master’s degrees we find a very limited scope for positive returns from spatial competition on student progression.

Finally, the results for combined bachelor’s/master’s degrees<sup>11</sup> are reported in Table 4.

We find results that are generally consistent with those for master’s degrees, for instance, in terms of entry requirements, the composition of teaching personnel and the size of the QA group. A striking difference is the negative premia for almost all student progression indicators suffered by degree programmes using English as the language of instruction. The percentage of students achieving at least 40 ECTS in the first year, for instance, is 11.8 pp lower. A possible explanation may be that the level of academic preparedness of foreign students might be below that of Italian students, on average, because the share of foreign students is likely to be larger in the degrees taught in English. In other words, the selection criteria applied by universities may be less effective in screening the best foreign students in combined bachelor’s/master’s degrees.

A higher percentage of teachers in the ‘core’ subjects of a degree programme is positively associated with both the percentage of regular students achieving at least 40 ECTS and the percentage of on-time graduates. Finally, unlike for master’s degrees, the intensity of spatial competition turns out to be positively related to student progression.

## 6.2 *Student Satisfaction*

Due to the progressive establishment of quasi-markets in education, universities compete for students. With increasing competition, student satisfaction becomes key for the success of degree programmes, for instance, to attract students overall or

---

<sup>11</sup> These degrees are available only for a subset of academic fields such as medicine, law, chemistry, veterinary science and architecture.



**Table 4** Student progression—combined bachelor's/master's degrees

Variables	(1) iC01 ECTS40	(2) iC13 ECTS <sub>1</sub>	(3) iC15 ETCS20 <sub>1</sub>	(4) iC16 ECTS40 <sub>1</sub>	(5) iC02 GRAD
Type of access = 3, entry test	0.062*** (0.017)	0.020 (0.018)	0.032* (0.019)	0.028 (0.027)	-0.047* (0.026)
English or multi-language degree	-0.062** (0.026)	-0.054** (0.027)	-0.048* (0.029)	-0.118*** (0.041)	0.086 (0.056)
Percentage of PO	-0.126*** (0.027)	-0.111*** (0.028)	-0.067** (0.028)	-0.191*** (0.041)	-0.164*** (0.043)
Percentage of PA	-0.129*** (0.025)	-0.089*** (0.026)	-0.076*** (0.027)	-0.150*** (0.039)	-0.212*** (0.039)
Percentage of RD	-0.163*** (0.057)	-0.131** (0.059)	-0.103* (0.061)	-0.183** (0.088)	0.096 (0.091)
Percentage of RU	-0.112*** (0.028)	-0.064** (0.029)	-0.015 (0.030)	-0.121*** (0.043)	-0.101** (0.046)
Percentage of other	-0.068 (0.085)	-0.252*** (0.089)	-0.526*** (0.091)	-0.403*** (0.133)	0.746*** (0.118)
Size of QA group	-0.003** (0.001)	0.002 (0.001)	0.000 (0.001)	0.002 (0.002)	-0.003* (0.002)
No. of students per tutor (/10, top coded)	0.002* (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.002)	0.001 (0.002)
Percent of teachers in core or characterizing SSD	0.216*** (0.078)	0.098 (0.082)	0.041 (0.084)	0.178 (0.122)	0.367** (0.168)
No. of degrees in same CUN area and area4 (/10)	0.066 (0.061)	0.312*** (0.064)	0.285*** (0.065)	0.393*** (0.095)	-0.163 (0.128)
Observations	1346	1347	1345	1346	1262
R-squared	0.619	0.832	0.860	0.658	0.481
Degree class × area4 FEs	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES
Mean of dep. var.	0.536	0.520	0.601	0.426	0.488

*Note:* \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Omitted reference categories for categorical variables are *numerus clausus* degree programmes, for type of admission, and the percentage of external personnel, for the composition of the teaching body. Dependent variables are indicated in the column headings (the original name of the ANVUR indicator and the name used in our analysis). CUN areas are broadly defined subject groups, and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. VQR stands for the 2011–2014 research evaluation exercise; PO, PA, RU, RD and other stand for full, associate, open-ended researchers, temporary researchers and other teaching personnel (e.g. PhD students, postdocs etc.), respectively

highly qualified students more specifically. Although we expect student satisfaction to partly reflect the speed of their careers, i.e. progression, many more elements related to their overall ‘university experience’ enter into this judgement.

In Table 5, we explore the correlates of student satisfaction for all degree levels. Specifically, columns (1) and (2) refer to bachelor’s degrees, columns (3) and (4) to master’s degrees and the remaining to combined bachelor’s/master’s degrees. We find that degree selectivity is not necessarily a synonym for higher student satisfaction in bachelor’s degrees (quite the opposite, in fact). Bachelor’s degrees with entry requirements have a 1.2 pp lower percentage of graduates who state that they would re-enrol in the same programme, compared to open-access programmes. Although the estimated coefficient for *numerus clausus* programmes is non-significant, presumably owing to the low number of bachelor’s degrees with this admission type, it is negative and sizeable in the first column. As for master’s degrees, in contrast, we find a negative gap of  $-2.2$  pp with respect to selective degrees (in this case, *numerus clausus* is the omitted category since open admission is not allowed in master’s degrees). Master’s degrees with entry requirements also score  $-1.3$  pp lower for the percentage of graduates that declare being generally satisfied with their degree programmes compared to selective degrees.

Quite surprisingly, degree programmes in which lecturing is in English score worse in terms of student satisfaction, irrespective of the degree level. The negative penalties are very high in bachelor’s degrees, where the percentage of students who declare that they would re-enrol in the same programme is 15.1 pp lower and those who report being generally satisfied is 17.2 pp lower. In addition to students in degrees taught in English having higher expectations (being more able, on average), another possible reading of this result is that highly internationalized degrees also attract more foreign students who, bearing higher educational costs, demand higher educational standards. Finally, a certain degree of dissatisfaction, especially among international students, may be caused by the teaching staff not always being adequately able to speak English properly. Indeed, although universities face a pressure to increase their teaching supply in English in order to attract foreign students, they may lack the personnel able to do it, given the low level of internationalization of the teaching staff.

Both the results for admission criteria and the language of instruction are quite interesting and point towards a potential tension for universities between selecting top-level applicants—i.e. ‘cream skimming’ and therefore increasing their performance indicators related to student progression—and the need to ensure them a top-quality education, with the risk of having unsatisfied students. Students in more selective programmes are likely to develop higher expectations that may not be met by the degree programmes.

As for the composition of the teaching body, we find results that are not consistent across degree levels. Indeed, the prevalence of more senior teaching staff, namely full and associate professors, is generally negatively associated with satisfaction indicators in bachelor’s degrees, but the relationship is positive for master’s degrees. A possible explanation is that more experienced teachers and researchers prefer to teach more advanced material, and their motivation and effort may be higher

**Table 5** Student satisfaction

Variables	(1) iC18 ENR	(2) iC25 SATI	(3) iC18 ENR	(4) iC25 SATI	(5) iC18 ENR	(6) iC25 SATI
Type of access: <i>numerus clausus</i>	-0.187 (0.157)	-0.028 (0.138)	-	-	-	-
Type of access: entry test	-0.012** (0.005)	-0.006 (0.005)	-0.022*** (0.007)	-0.013** (0.006)	-0.013 (0.026)	-0.007 (0.025)
English or multi-language degree	-0.151*** (0.020)	-0.172*** (0.017)	-0.045*** (0.007)	-0.029*** (0.006)	-0.114** (0.054)	-0.078 (0.053)
Inter-class degree	-0.186*** (0.017)	-0.246*** (0.015)	-0.255*** (0.015)	-0.309*** (0.014)	-	-
Percentage of PO	-0.047*** (0.018)	-0.033** (0.016)	0.033* (0.017)	0.035** (0.016)	-0.015 (0.042)	0.009 (0.041)
Percentage of PA	-0.046*** (0.013)	-0.037*** (0.012)	0.027* (0.015)	0.041*** (0.014)	-0.043 (0.036)	-0.041 (0.036)
Percentage of RD	0.109*** (0.025)	-0.014 (0.022)	0.007 (0.028)	-0.054** (0.026)	0.115 (0.088)	-0.080 (0.087)
Percentage of RU	-0.048*** (0.016)	-0.036*** (0.014)	0.010 (0.020)	0.033* (0.018)	-0.132*** (0.045)	-0.143*** (0.044)
Percentage of other	0.120*** (0.039)	-0.007 (0.034)	0.063 (0.047)	-0.060 (0.042)	-0.714*** (0.127)	-0.796*** (0.124)
Size of QA group	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.001 (0.001)	0.002 (0.002)	0.002 (0.002)
No. of students per tutor (/10, top coded)	-0.001* (0.001)	-0.001 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.003 (0.002)	0.004* (0.002)

(continued)

Table 5 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	iC18 ENR	iC25 SATI	iC18 ENR	iC25 SATI	iC18 ENR	iC25 SATI
Percent of teachers in core or characterizing SSD	0.211*** (0.022)	0.214*** (0.020)	0.076*** (0.017)	0.099*** (0.016)	-0.080 (0.168)	-0.121 (0.164)
VQR score teachers			0.093***	0.133***		
			(0.012)	(0.011)		
No. of degrees in same CUN area and area4 (/10)	0.024*** (0.009)	0.033*** (0.008)	0.005 (0.011)	0.015 (0.010)	-0.028 (0.155)	-0.015 (0.152)
Degree level	BA	BA	MA	MA	BA+MA	BA+MA
Observations	7323	7285	7303	7280	1028	1024
R-squared	0.274	0.254	0.279	0.337	0.227	0.183
Degree class × <i>area4</i> FEs	YES	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES	YES
Mean of dep. var.	0.668	0.860	0.716	0.862	0.672	0.863

Note: \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Omitted reference categories for categorical variables are open-access and *numerus clausus* degree programmes, for type of admission and bachelor's vs. master's and combined bachelor's/master's degrees, respectively, and the percentage of external personnel, for the composition of the teaching body. Dependent variables are indicated in the column headings (the original name of the ANVUR indicator and the name used in our analysis). CUN areas are broadly defined subject groups, and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. VQR stands for the 2011–2014 research evaluation exercise; PO, PA, RU, RD and other stand for full, associate, open-ended researchers, temporary researchers and other teaching personnel (e.g. PhD students, postdocs etc.), respectively

when they teach in master's degrees. Or alternatively, master's degree students may appreciate advanced and more difficult teaching material compared to their bachelor's degree peers. It is worth mentioning that at least in bachelor's degrees, we do not observe the same penalty noted for student progression in programmes employing a larger share of junior personnel: student satisfaction measured by the percentage of those who would re-enrol turns out to be higher (+1.2 pp for a 10 pp increase in the percentage of 'other personnel').

We do not find effects for the size of the QA group or for tutors.

The percentage of teachers in 'core' subjects is instead very strongly associated with student satisfaction, especially in bachelor's degrees. Increasing by 10 pp the percentage of teachers in core SSD is associated with premia of 2.11 pp for the percentage of students who would re-enrol and of 2.14 pp for the percentage of students who declare being generally satisfied. This may point to the fact that a higher degree specialization or a better match between teachers and the subject fields in which they teach may increase student satisfaction.

Degree programmes employing teaching staff that perform well in research generally display higher student satisfaction. Master's degrees with a one-point higher VQR score have 9.3 and 13.3 pp higher percentages of students who would re-enrol and who are satisfied with the degree, respectively.

Finally, spatial competition appears to be positively associated with student satisfaction only in bachelor's degrees. A possible explanation is that master's degrees may be quite specialized and be subject to less spatial competition and thus have fewer incentives to improve compared to bachelor's degrees.

## **7 Robustness Checks: Controlling for the Quality of Student Intake**

Up to now, we have excluded from the regression model measures of the 'quality' of student intake. However, as we have argued, variables such as the type of access or the language of instruction partly proxy for this. In Tables C1–C4 in the Appendix C, we have re-estimated all models in the main text including two additional variables provided by the Ministry of Education and computed on National Student Registry data: the percentage of students coming from the academic secondary school track and the percentage of newly enrolled students with a final secondary school mark of 90 or greater (out of a range of 60–100). These are indicators of student ability and are partly correlated with student family background, since high socio-economic status students are more likely to enrol in the academic track compared to the technical and vocational tracks (i.e. the Italian system of upper secondary education is characterized by three broad tracks, and the academic track is the one generally chosen by students who plan to enrol in tertiary education).

Consistent with the past literature, the estimates show that the two proxies of student academic preparedness at entry into HE are strong positive predictors

of student progression at all degree levels. Yet, the coefficients on the degree-programme variables are generally not affected. Quite interestingly, more able students—namely, those coming from the academic track—appear to also be ‘pickier’ or have higher expectations, i.e. a higher concentration of these is associated with lower levels of student satisfaction at the bachelor’s degree level. Interestingly, after controlling for students’ entry qualifications, the coefficient on entry test admission ceases to be statistically significant for bachelor’s degrees, while we still find a satisfaction penalty for master’s degrees. Conversely, lower average student satisfaction is still observed for degrees taught in English, for both bachelor’s and master’s degrees.

## 8 Concluding Remarks

The expansion of the number of university graduates in the population is one of the key objectives set by the EU. The ‘Education and Training 2020’ (ET 2020) work programme set an ambitious target: ‘The share of 30–34 year-olds with tertiary educational attainment should be at least 40%.’<sup>12</sup> One way of achieving this goal is to lower student dropout from higher education and ensuring satisfactory student progression. In order to reach this goal, many countries have established quality assurance systems for higher education.

The existence of QA systems coupled with the higher competition for students (the quasi-market in higher education) has led Italian universities to devote increasing attention to the quality of the degree programmes they offer. Yet, a systematic analysis of the degree-programme correlates of student dropout and progression is still lacking. In this chapter, we leverage the very rich set of indicators built by the National Agency for the Evaluation of the University System and Research (ANVUR) at the degree level and seek to fill this gap by merging degree-programme-level information gathered by the programme cards (*Scheda SUA-CdS*) with ANVUR degree-programme performance indicators. To the best of our knowledge, this is the first analysis using degree programmes as the unit of observation and data on the complete Italian university supply (for the years 2013–2018).

Our empirical analysis identifies several degree-programme characteristics associated with student dropout and progression.

Bachelor’s degree programmes with entry requirements generally have better student progression indicators than those with open admission policies, except for graduation times. Interestingly, programmes with this type of admission also exhibit better progression indicators with respect to the more selective (on paper) master’s degree programmes with *numerus clausus* policies. Higher selectivity is often negatively associated with student satisfaction in bachelor’s degrees, however,

---

<sup>12</sup> <https://ec.europa.eu/eurostat/web/education-and-training/eu-benchmarks>.

presumably owing to the higher expectations of enrolled students, while it is positively associated with student satisfaction in master's degrees. A positive association with student progression is also observed for programmes taught in English (except in combined bachelor's/master's degrees), while a penalty on student satisfaction generally emerges for degrees not taught in Italian, irrespective of the degree level. We put forward that this may be partly due to the fact part of the teaching body lacks adequate proficiency in English.

Degree-programme performance is affected by the composition of the teaching body, with programmes employing external teachers generally showing better progression indicators but not necessarily higher average student satisfaction. Programmes where more junior personnel (e.g. PhD students or postdocs) accounts for a larger proportion of the teaching body display slower student progression at all degree levels but higher on-time graduation rates in master's and combined bachelor's/master's degrees. We argue that this can be explained by the different teaching incentives and motivations of external and internal junior vs senior teaching staff.

Tutor teachers appear to be a valuable resource to support students' academic careers and are generally associated with better progression indicators. Yet, those premia are not reflected in average student satisfaction.

A higher proportion of teachers in the 'core' subject groups of degree programmes is associated with a higher percentage of students graduating on time for combined bachelor's/master's degrees, whereas the effect turns out to be negative for bachelor's and master's degrees. Counterintuitively, the same variable is associated with higher student satisfaction in bachelor's and master's degrees.

Our analysis points to some complementarity between research quality and teaching quality at advanced levels of tertiary education. Master's degree programmes whose teaching body performed well in the last Italian research evaluation exercise (2011–2014) perform better in terms of both student progression and student satisfaction.

Finally, the geographical concentration of degree programmes in the same broadly defined subject groups, as a proxy of spatial competition, is positively correlated with student progression in bachelor's and combined bachelor's/master's degrees and student satisfaction in bachelor's degrees, suggesting that higher competitive pressure may push higher education institutions to improve the quality of the educational services they provide.

Although the richness of our data allows us to uncover many interesting associations between the characteristics of degree programmes and student progression and dropout, this study has a descriptive nature, and without further research these associations cannot necessarily be attributed a causal interpretation. Our work nonetheless provides some interesting insights that could represent a starting point for future research.

**Acknowledgments** We thank ANVUR (National Agency for the Evaluation of the University System and Research) for providing some of the data used in this study. Special thanks go to Giampiero D'Alessandro for his help with the data. Comments received at the ANVUR 'III

Concorso idee di ricerca' (Rome, November 2019) are gratefully acknowledged. ANVUR does not bear any responsibility for the content of this report. The usual disclaimer applies.

## Appendix A: List of Abbreviations and Acronyms Used in the Chapter

CEV	<i>Commissione di Esperti della Valutazione</i> (Evaluation Expert Committee)
CRUI	<i>Conferenza dei Rettori delle Università Italiane</i> (Conference of the Rectors of Italian Universities)
CUN	<i>Consiglio Universitario Nazionale</i> (National University Council)
DID	Difference in differences
EHEA	European Higher Education Area
ENQA	European Network of Quality Assurance Agencies
ESG	European Standards and Guidelines
ET 2020	Education and Training 2020
EU	European Union
GPA	Grade point average
HE	Higher education
HEI	Higher education institution
L	First-level degree
L&T	Learning and teaching
LM	Second-level degree
LMCU	Unique-cycle degree
MIUR	<i>Ministero dell'Istruzione, dell'Università e della Ricerca</i> (Ministry of Education, Universities and Research)
NPM	New public management
OECD	Organization for Economic Co-operation and Development
PA	<i>Professore Associato</i> (associate professor)
PhD	Doctor of philosophy
PI	Performance indicator
PO	<i>Professore Ordinario</i> (full professor)
QA	Quality assurance
QAC	Quality Assurance Council
RD	<i>Ricercatore a tempo Determinato</i> (temporary contract researcher)
RDD	Regression discontinuity design
RU	<i>Ricercatore Universitario</i> (open-ended contract researcher)



## Appendix B: Macro-Area Differences in the Values of the ANVUR Performance Indicators

**Table B.1** Mean of degree-level performance indicators by geographic macro-area

Variables	North-West		North-East		Centre		South and Islands	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Panel (a) Bachelor's degrees								
<i>Performance Indicators</i>								
iC01 (ECTS40)	0.57	0.22	0.60	0.19	0.47	0.21	0.43	0.20
iC13 (ECTS <sub>1</sub> )	0.48	0.29	0.52	0.30	0.42	0.26	0.41	0.25
iC15 (ECTS20 <sub>1</sub> )	0.53	0.31	0.57	0.31	0.49	0.30	0.49	0.29
iC16 (ECTS40 <sub>1</sub> )	0.40	0.29	0.44	0.28	0.32	0.26	0.30	0.25
iC02 (GRAD)	0.61	0.19	0.57	0.19	0.50	0.22	0.41	0.23
iC18 (ENR)	0.69	0.17	0.68	0.16	0.68	0.20	0.64	0.19
iC25 (SATI)	0.87	0.14	0.88	0.12	0.85	0.18	0.85	0.16
Panel (b) Master's degrees								
<i>Performance Indicators</i>								
iC01 (ECTS40)	0.56	0.18	0.58	0.17	0.48	0.18	0.48	0.19
iC13 (ECTS <sub>1</sub> )	0.57	0.30	0.59	0.30	0.51	0.27	0.49	0.26
iC15 (ECTS20 <sub>1</sub> )	0.69	0.35	0.71	0.35	0.64	0.33	0.66	0.34
iC16 (ECTS40 <sub>1</sub> )	0.47	0.29	0.49	0.30	0.38	0.26	0.37	0.26
iC02 (GRAD)	0.68	0.21	0.62	0.23	0.53	0.25	0.51	0.26
iC18 (ENR)	0.72	0.18	0.73	0.16	0.71	0.20	0.71	0.18
iC25 (SATI)	0.86	0.17	0.88	0.14	0.84	0.21	0.87	0.17
Panel (c) Combined								
Bachelor's/master's degrees								
<i>Performance Indicators</i>								
iC01 (ECTS40)	0.60	0.19	0.65	0.15	0.51	0.18	0.47	0.18
iC13 (ECTS <sub>1</sub> )	0.54	0.30	0.59	0.31	0.49	0.30	0.50	0.29
iC15 (ECTS20 <sub>1</sub> )	0.62	0.34	0.65	0.38	0.59	0.34	0.58	0.33
iC16 (ECTS40 <sub>1</sub> )	0.47	0.32	0.53	0.31	0.38	0.32	0.39	0.29
iC02 (GRAD)	0.56	0.22	0.53	0.23	0.51	0.27	0.42	0.26
iC18 (ENR)	0.73	0.13	0.69	0.15	0.65	0.21	0.65	0.19
iC25 (SATI)	0.89	0.13	0.88	0.11	0.82	0.23	0.87	0.16

*Note:* The table reports the mean and standard deviation (SD) of the degree-programme performance indicators used in the empirical analysis. See Sect. 5.2 for variable definitions

## Appendix C: Robustness Checks with Secondary School Diploma Type and Secondary School Graduation Mark

**Table C.1** Student progression—bachelor's degrees

Variables	(1) iC01 ECTS40	(2) iC13 ECTS <sub>1</sub>	(3) iC15 ETCS20 <sub>1</sub>	(4) iC16 ECTS40 <sub>1</sub>	(5) iC02 GRAD
Type of access: <i>numerus clausus</i>	0.214* (0.128)	0.027 (0.122)	-0.054 (0.132)	0.060 (0.163)	-0.082 (0.175)
Type of access: entry test	0.049*** (0.004)	0.050*** (0.004)	0.047*** (0.004)	0.044*** (0.005)	-0.022*** (0.005)
English or multi-language degree	0.160*** (0.015)	0.170*** (0.015)	0.125*** (0.016)	0.180*** (0.020)	0.188*** (0.022)
Inter-class degree	0.007 (0.012)	-0.009 (0.012)	-0.018 (0.012)	-0.006 (0.015)	0.006 (0.016)
Percentage of PO	-0.095*** (0.013)	-0.038*** (0.012)	-0.029** (0.013)	-0.038** (0.016)	-0.125*** (0.016)
Percentage of PA	-0.059*** (0.010)	-0.049*** (0.010)	-0.053*** (0.011)	-0.066*** (0.013)	-0.135*** (0.013)
Percentage of RD	-0.095*** (0.020)	-0.045** (0.019)	-0.110*** (0.020)	-0.062** (0.025)	0.055** (0.024)
Percentage of RU	-0.073*** (0.011)	-0.046*** (0.011)	-0.044*** (0.012)	-0.068*** (0.014)	-0.113*** (0.015)
Percentage of other	-0.349*** (0.032)	-0.276*** (0.031)	-0.375*** (0.033)	-0.301*** (0.041)	0.066* (0.037)
Size of QA group	-0.002*** (0.001)	-0.001** (0.001)	-0.001* (0.001)	-0.001 (0.001)	-0.005*** (0.001)
No. of students per tutor (/10, top coded)	-0.001*** (0.000)	-0.001* (0.000)	-0.001*** (0.000)	-0.001*** (0.001)	0.001** (0.001)
Percent of teachers in core or characterizing SSD	-0.026 (0.016)	-0.007 (0.015)	-0.003 (0.016)	0.016 (0.020)	-0.074*** (0.021)
No. of degrees in the same CUN area and area4 (/10)	0.026*** (0.008)	0.025*** (0.008)	0.037*** (0.008)	0.036*** (0.010)	0.008 (0.008)
Academic track (%)	0.193*** (0.011)	0.162*** (0.010)	0.148*** (0.011)	0.173*** (0.014)	0.042*** (0.014)

(continued)

**Table C.1** (continue)

Secondary school mark $\geq 90$ (%)	0.241*** (0.014)	0.183*** (0.014)	0.134*** (0.015)	0.218*** (0.018)	0.042** (0.018)
Observations	9122	9052	9075	9074	10,253
R-squared	0.664	0.808	0.816	0.659	0.412
Degree class $\times$ <i>area4</i> FEs	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES
Mean of dep. var.	0.505	0.449	0.517	0.356	0.509

*Note:* \* \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Omitted reference categories for categorical variables are open-access degree programmes, for type of admission, and the percentage of external personnel, for the composition of the teaching body. Dependent variables are indicated in the column headings (the original name of the ANVUR indicator and the name used in our analysis). CUN areas are broadly defined subject groups, and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. VQR stands for the 2011–2014 research evaluation exercise; PO, PA, RU, RD and other stand for full, associate, open-ended researchers, temporary researchers and other teaching personnel (e.g. PhD students, postdocs etc.), respectively

**Table C.2** Student progression—master’s degrees

Variables	(1) iC01 ECTS40	(2) iC13 ECTS <sub>1</sub>	(3) iC15 ETCS20 <sub>1</sub>	(4) iC16 ECTS40 <sub>1</sub>	(5) iC02 GRAD
Type of access: entry test	0.040*** (0.005)	0.038*** (0.004)	0.023*** (0.004)	0.047*** (0.006)	-0.011 (0.007)
English or multi-language degree	0.082*** (0.006)	0.065*** (0.005)	0.036*** (0.005)	0.087*** (0.007)	0.086*** (0.007)
Inter-class degree	0.025** (0.012)	0.020* (0.010)	0.002 (0.010)	0.026* (0.014)	-0.001 (0.015)
Percentage of PO	-0.022 (0.014)	-0.007 (0.011)	-0.011 (0.011)	-0.043*** (0.016)	-0.126*** (0.017)
Percentage of PA	-0.038*** (0.013)	0.002 (0.010)	0.008 (0.010)	-0.009 (0.015)	-0.099*** (0.015)
Percentage of RD	-0.063*** (0.024)	0.033* (0.020)	-0.043** (0.019)	0.043 (0.029)	0.069** (0.030)
Percentage of RU	-0.059*** (0.015)	-0.048*** (0.012)	-0.006 (0.012)	-0.056*** (0.018)	-0.097*** (0.019)
Percentage of other	-0.176*** (0.036)	-0.187*** (0.030)	-0.312*** (0.029)	-0.154*** (0.042)	0.409*** (0.046)
Size of QA group	-0.001 (0.001)	-0.002** (0.001)	0.000 (0.001)	-0.000 (0.001)	-0.007*** (0.001)
No. of students per tutor (/10, top coded)	-0.004*** (0.001)	-0.000 (0.001)	-0.002*** (0.001)	-0.000 (0.001)	-0.003*** (0.001)

(continued)

**Table C.2** (continued)

Variables	(1) iC01 ECTS40	(2) iC13 ECTS <sub>1</sub>	(3) iC15 ETCS20 <sub>1</sub>	(4) iC16 ECTS40 <sub>1</sub>	(5) iC02 GRAD
Percent of teachers in core or characterizing SSD	-0.025*	-0.020*	0.017*	-0.020	-0.033**
	(0.013)	(0.011)	(0.010)	(0.015)	(0.016)
VQR score teachers	0.086***	0.052***	0.052***	0.094***	0.053***
	(0.010)	(0.008)	(0.008)	(0.011)	(0.012)
No. of degrees in the same CUN area and area4 (/10)	-0.001	0.003	-0.010*	0.013*	0.008
	(0.007)	(0.005)	(0.005)	(0.008)	(0.007)
Academic track (%)	0.086***	0.053***	0.071***	0.065***	-0.004
	(0.012)	(0.010)	(0.010)	(0.015)	(0.015)
Secondary school mark $\geq$ 90 (%)	0.042***	0.087***	0.076***	0.101***	-0.144***
	(0.015)	(0.013)	(0.012)	(0.018)	(0.019)
Observations	8880	8837	8884	8882	10,369
R-squared	0.437	0.840	0.892	0.661	0.459
Degree class $\times$ area4 FEs	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES
Mean of dep. var.	0.520	0.536	0.673	0.424	0.576

*Note:* \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Omitted reference categories for categorical variables are *numerus clausus* degree programmes, for type of admission, and the percentage of external personnel, for the composition of the teaching body. Dependent variables are indicated in the column headings (the original name of the ANVUR indicator and the name used in our analysis). CUN areas are broadly defined subject groups, and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. VQR stands for the 2011–2014 research evaluation exercise; PO, PA, RU, RD and other stand for full, associate, open-ended researchers, temporary researchers and other teaching personnel (e.g. PhD students, postdocs etc.), respectively

**Table C.3** Student progression—combined bachelor's/master's degrees

Variables	(1) iC01 ECTS40	(2) iC13 ECTS <sub>1</sub>	(3) iC15 ETCS20 <sub>1</sub>	(4) iC16 ECTS40 <sub>1</sub>	(5) iC02 GRAD
Type of access: entry test	0.035**	-0.006	0.014	-0.009	-0.054**
	(0.017)	(0.018)	(0.019)	(0.027)	(0.027)
English or multi-language degree	-0.013	0.006	-0.000	-0.046	0.077
	(0.027)	(0.028)	(0.030)	(0.043)	(0.057)
Percentage of PO	-0.133***	-0.117***	-0.070**	-0.199***	-0.162***
	(0.026)	(0.027)	(0.028)	(0.040)	(0.043)
Percentage of PA	-0.115***	-0.073***	-0.064**	-0.129***	-0.213***
	(0.024)	(0.025)	(0.026)	(0.038)	(0.039)

(continued)

**Table C.3** (continued)

Percentage of RD	-0.127** (0.055)	-0.087 (0.058)	-0.073 (0.060)	-0.126 (0.087)	0.085 (0.091)
Percentage of RU	-0.103*** (0.027)	-0.057** (0.028)	-0.011 (0.029)	-0.109** (0.042)	-0.098** (0.046)
Percentage of other	-0.027 (0.083)	-0.209** (0.087)	-0.495*** (0.090)	-0.348*** (0.130)	0.762*** (0.119)
Size of QA group	-0.002** (0.001)	0.002* (0.001)	0.001 (0.001)	0.003* (0.002)	-0.003* (0.002)
No. of students per tutor (/10, top coded)	0.002** (0.001)	0.001 (0.001)	0.001 (0.001)	0.002 (0.002)	0.002 (0.002)
Percent of teachers in core or characterizing SSD	0.193** (0.076)	0.060 (0.080)	0.015 (0.083)	0.140 (0.119)	0.397** (0.169)
No. of degrees in the same CUN area and area4 (/10)	0.066 (0.059)	0.309*** (0.062)	0.281*** (0.064)	0.395*** (0.093)	-0.158 (0.128)
Academic track (%)	0.090*** (0.032)	0.159*** (0.033)	0.119*** (0.035)	0.152*** (0.050)	-0.062 (0.054)
Secondary school mark ≥ 90 (%)	0.220*** (0.031)	0.156*** (0.032)	0.096*** (0.034)	0.283*** (0.048)	0.142*** (0.054)
Observations	1346	1347	1345	1346	1258
R-squared	0.641	0.840	0.863	0.673	0.483
Degree class × area4 FEs	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES
Mean of dep. var.	0.536	0.520	0.601	0.426	0.489

*Note:* \*, \*\* and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Omitted reference categories for categorical variables are *numerus clausus* degree programmes, for type of admission, and the percentage of external personnel, for the composition of the teaching body. Dependent variables are indicated in the column headings (the original name of the ANVUR indicator and the name used in our analysis). CUN areas are broadly defined subject groups, and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. VQR stands for the 2011–2014 research evaluation exercise; PO, PA, RU, RD and other stand for full, associate, open-ended researchers, temporary researchers and other teaching personnel (e.g. PhD students, postdocs etc.), respectively

Table C.4 Student satisfaction

Variables	(1) iC18 ENR	(2) iC25 SATI	(3) iC18 ENR	(4) iC25 SATI	(5) iC18 ENR	(6) iC25 SATI
Type of access: <i>numerus clausus</i>	-0.182 (0.157)	-0.023 (0.138)	-	-	-	-
Type of access: entry test	-0.007 (0.005)	-0.002 (0.005)	-0.020*** (0.007)	-0.013** (0.006)	-0.023 (0.026)	-0.013 (0.025)
English or multi-language degree	-0.152*** (0.020)	-0.173*** (0.017)	-0.046*** (0.007)	-0.032*** (0.007)	-0.097* (0.055)	-0.063 (0.054)
Inter-class degree	-0.183*** (0.017)	-0.243*** (0.015)	-0.252*** (0.015)	-0.307*** (0.014)		
Percentage of PO	-0.045** (0.018)	-0.030* (0.016)	0.034* (0.017)	0.036** (0.016)	-0.014 (0.042)	0.011 (0.041)
Percentage of PA	-0.046*** (0.013)	-0.036*** (0.012)	0.026* (0.015)	0.040*** (0.014)	-0.038 (0.036)	-0.039 (0.036)
Percentage of RD	0.109*** (0.025)	-0.013 (0.022)	0.004 (0.028)	-0.057** (0.026)	0.132 (0.088)	-0.070 (0.087)
Percentage of RU	-0.049*** (0.016)	-0.036*** (0.014)	0.013 (0.020)	0.033* (0.018)	-0.138*** (0.045)	-0.149*** (0.044)
Percentage of other	0.103*** (0.039)	-0.022 (0.034)	0.066 (0.047)	-0.062 (0.042)	-0.705*** (0.128)	-0.795*** (0.124)
Size of QA group	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.001 (0.001)	0.002 (0.002)	0.002 (0.002)
No. of students per tutor (/10, top coded)	-0.001** (0.001)	-0.001* (0.001)	0.000 (0.001)	0.000 (0.001)	0.003 (0.002)	0.004* (0.002)

(continued)

**Table C.4** (continued)

Percent of teachers in core or characterizing SSD	0.218*** (0.023)	0.221*** (0.020)	0.079*** (0.017)	0.103*** (0.016)	-0.134 (0.171)	-0.167 (0.167)
VQR score teachers			0.096*** (0.012)	0.137*** (0.011)		
No. of degrees in the same CUN area and area4 (10)	0.025*** (0.009)	0.034*** (0.008)	0.008 (0.011)	0.016 (0.010)	-0.037 (0.155)	-0.018 (0.152)
Academic track (%)	-0.047*** (0.015)	-0.050*** (0.013)	-0.005 (0.015)	-0.029*** (0.014)	0.102* (0.053)	0.076 (0.052)
Secondary school mark $\geq 90$ (%)	-0.022 (0.019)	-0.012 (0.017)	-0.051*** (0.019)	-0.020 (0.018)	0.030 (0.052)	-0.006 (0.050)
Degree Level	BA	BA	MA	MA	BA+MA	BA+MA
Observations	7281	7243	7280	7257	1024	1020
R-squared	0.276	0.257	0.282	0.340	0.232	0.187
Degree class $\times$ area4 FEs	YES	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES	YES
Mean of dep. var.	0.668	0.860	0.716	0.862	0.672	0.863

*Note:* \*, \*\*, and \*\*\* indicate statistical significance at the 10%, 5% and 1% levels, respectively. Omitted reference categories for categorical variables are open-access and *numerus clausus* degree programmes, for type of admission and bachelor's vs. master's and combined bachelor's/master's degrees, respectively, and the percentage of external personnel, for the composition of the teaching body. Dependent variables are indicated in the column headings (the original name of the ANVUR indicator and the name used in our analysis). CUN areas are broadly defined subject groups, and *area4* is a geographic indicator (North-West, North-East, Centre, South and Islands). SSDs are narrowly defined scientific areas. VQR stands for the 2011–2014 research evaluation exercise; PO, PA, RU, RD and other stand for full, associate, open-ended researchers, temporary researchers and other teaching personnel (e.g. PhD students, postdocs etc.), respectively

## Appendix D: ANVUR Indicators

**Table D.1** Indicators on teaching (level of the entire university + degree programme)

Code	Indicator description	Domain and area
iC01	Percentage of regularly attending students who have earned at least 40 ECTS credits during the academic year	Learning outcomes: student success rate
iC02	Percentage of students (undergraduate and postgraduate) graduating within the natural programme duration	Learning outcomes: student success rate
iC03	Percentage of undergraduate first-year students coming from other regions	Environment: student body composition
iC04	Percentage of first-year postgraduate students who graduated from another university	Environment: student body composition
iC05	Ratio between regular enrolled students and members of the academic staff	Environment: teaching resources
iC06	Percentage of graduates employed within one year of obtaining their qualification	Learning outcomes (first-level degrees): employability
iC07	Percentage of graduates employed within three years of obtaining their qualification	Learning outcomes: employability
iC08	Percentage of tenured faculty members who belong to scientific disciplinary sectors (SSD) characterizing the degree programme (in which they teach)	Environment: academic staff composition
iC09	Indicator of research quality (based on VQR results) of academic staff teaching in the degree programme (second-level degrees only)	Teaching processes: quality of the staff



**Table D.2** Indicators of internationalization (level of the entire university + degree programme)

Code	Indicator description	Domain and area
iC10	Percentage of ECTS credits earned abroad by regularly enrolled students over the total of ECTS credits earned by students enrolled during the natural duration of the programme	Learning outcomes: student success rate
iC11	Percentage of regular graduates (within the normal degree programme duration) that have earned at least 12 ECTS credits abroad	Learning outcomes: student success rate
iC12	Percentage of students enrolled in the first year (undergraduates and postgraduates) who earned their qualifications abroad	Environment: student body composition

**Table D.3** Further indicators for the evaluation of teaching (degree-programme level)

Code	Indicator description	Domain and area
iC13	Percentage of ECTS credits achieved in the first year over total ECTS credits to be achieved in the first year	Learning outcomes: student success rate
iC14	Percentage of students who continue on to the second year in the same degree programme	Learning outcomes: student success rate
iC15	Percentage of students who continue on to the second year in the same degree programme having earned at least 20 ECTS credits in the first year	Learning outcomes: student success rate
iC16	Percentage of students who continue on to the second year in the same degree programme having earned at least 40 ECTS credits in the first year	Learning outcomes: student success rate
iC17	Percentage of students (undergraduate and postgraduate) graduating within 1 year beyond the normal duration of the degree programme	Learning outcomes: student success rate
iC18	Percentage of graduates who would re-enrol in the same degree programme again	Learning outcomes: student satisfaction
iC19	Hours of teaching by tenured faculty members on the total of teaching hours	Environment: teaching resources
iC20	Ratio between tutors and enrolled students (for online degree programmes)	Environment: teaching resources

**Table D.4** Further indicators for testing (degree-programme level)

Code	Indicator description	Domain and area
iC21	Percentage of students who continue in the second year, whether in the same university and the same degree programme or not	Learning outcomes: student success rate
iC22	Percentage of freshmen (undergraduate and postgraduate) who graduate within the normal duration for the degree programme	Learning outcomes: student success rate
iC23	Percentage of freshmen (undergraduate and postgraduate) who continue their second year in a different degree programme within the same university	Learning outcomes: student success rate
iC24	Percentage of dropouts after the first academic year	Learning outcomes: student success rate
iC25	Percentage of graduates generally satisfied with their degree programme	Learning outcomes: student satisfaction
iC26	Percentage of graduates employed within one year of obtaining their qualification (second-level and unique-cycle degrees)	Learning outcomes: employability
iC27	Ratio between enrolled students and the total number of academic staff (weighted for teaching hours)	Environment: teaching resources
iC28	Ratio between first-year students and the total number of first-year lecturers (weighted by teaching hours)	Environment: teaching resources
iC29	Ratio between tutors with a doctorate degree and enrolled students (for online programmes)	Environment: teaching resources

## References

- Adamopoulou, E., & Tanzi, G. M. (2017). Academic drop-out and the great recession. *Journal of Human Capital*, 11(1), 35–71.
- Agasisti, T., Barbato, G., Dal Molin, M., & Turri, M. (2019). Internal quality assurance in universities: Does NPM matter? *Studies in Higher Education*, 44(6), 960–977.
- Aina, C., Baici, E., & Casalone, G. (2011). Time-to-degree: Students' abilities, university characteristics or something else? Evidence from Italy. *Education Economics*, 19(3), 311–325.
- Aina, C. (2013). Parental background and university dropout in Italy. *Higher Education*, 65(4), 437–456.
- Aina, C., Baici, E., Casalone, G., & Pastore, F. (2021). The determinants of university dropout: A review of the socio-economic literature. *Socio-Economic Planning Sciences*, 101102.
- Alon, S. (2007). The influence of financial aid in leveling group differences in graduating from elite institutions. *Economics of Education Review*, 26(3), 296–311.
- Alon, S., & Tienda, M. (2005). Assessing the mismatch hypothesis: Differences in college graduation rates by institutional selectivity. *Sociology of Education*, 78(4), 294–315.
- Andreani, M., Russo, D., Salini, S., & Turri, M. (2020). Shadows over accreditation in higher education: Some quantitative evidence. *Higher Education*, 79(4), 691–709.

- Andrews, R. J., Li, J., & Lovenheim, M. F. (2016). Quantile treatment effects of college quality on earnings. *Journal of Human Resources*, *51*(1), 200–238.
- Andrews, R. J., Imberman, S. A., & Lovenheim, M. F. (2020). Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, *74*, 101923.
- Anelli, M. (2018). Returns to elite university education: a quasi-experimental analysis. *Journal of the European Economic Association*, *18*(6), 2824–2868.
- Angrist, J., Lang, D., & Oreopoulos, P. (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, *1*(1), 136–163.
- Arendt, J. N. (2013). The effect of public financial aid on dropout from and completion of university education: Evidence from a student grant reform. *Empirical Economics*, *44*(3), 1545–1562.
- Artés, J., Pedraja-Chaparro, F., & del Mar Salinas-Jiménez, M. (2017). Research performance and teaching quality in the Spanish higher education system: Evidence from a medium-sized university. *Research Policy*, *46*(1), 19–29.
- Arulampalam, W., Naylor, R. A., & Smith, J. P. (2004a). A hazard model of the probability of medical school drop-out in the UK. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *167*(1), 157–178.
- Arulampalam, W., Naylor, R. A., & Smith, J. P. (2004b). Factors affecting the probability of first year medical student dropout in the UK: A logistic analysis for the intake cohorts of 1980–1992. *Medical Education*, *38*(5), 492–503.
- Barro, R. J. (2001). Human capital and growth. *American Economic Review*, *91*(2), 12–17.
- Baryla, Jr., E. A., & Dotterweich, D. (2001). Student migration: Do significant factors vary by region? *Education Economics*, *9*(3), 269–280.
- Becker, G. S. (1994). *Human capital: A theoretical and empirical analysis with special reference to education* (3rd Ed). NBER Books, National Bureau of Economic Research.
- Belloc, F., Maruotti, A., & Petrella, L. (2010). University drop-out: An Italian experience. *Higher Education*, *60*(2), 127–138.
- Bennet, R. (2009). Determinants of undergraduate student drop out rates in a university business studies department. *Journal of Further and Higher Education*, *27*(2), 123–141.
- Biancardi, D., & Bratti, M. (2019). The effect of introducing a research evaluation exercise on student enrolment: Evidence from Italy. *Economics of Education Review*, *69*, 73–93.
- Black, D. A., & Smith, J. A. (2004). How robust is the evidence on the effects of college quality? Evidence from matching. *Journal of Econometrics*, *121*(1–2), 99–124.
- Bleiklie, I., & Michelsen, S. (2013). Comparing HE policies in Europe. *Higher Education*, *65*(1), 113–133.
- Bondonio, D., & Berton, F. (2018). The impact of degree duration on higher education participation: Evidence from a large? Scale natural experiment. *Oxford Bulletin of Economics and Statistics*, *80*(5), 905–930.
- Bound, J., & Turner, S. (2007). Cohort crowding: How resources affect collegiate attainment. *Journal of Public Economics*, *91*(5–6), 877–899.
- Bound, J., Lovenheim, M. F., & Turner, S. (2010). Why have college completion rates declined? An analysis of changing student preparation and collegiate resources. *American Economic Journal: Applied Economics*, *2*(3), 129–157.
- Bound, J., Lovenheim, M. F., & Turner, S. (2012). Increasing time to baccalaureate degree in the United States. *Education Finance and Policy*, *7*(4), 375–424.
- Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton: Princeton University Press.
- Bradley, S., & Migali, G. (2019). The effects of the 2006 tuition fee reform and the Great Recession on university student dropout behaviour in the UK. *Journal of Economic Behavior & Organization*, *164*, 331–356.
- Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, *41*, 71–88.

- Braga, M., Paccagnella, M., & Pellizzari, M. (2016). The impact of college teaching on students' academic and labor market outcomes. *Journal of Labor Economics*, 34(3), 781–822.
- Bratti, M., & Verzillo, M. (2019). The “gravity” of quality: research quality and the attractiveness of universities in Italy. *Regional Studies*, 53(10), 1385–1396.
- Brownback, A., & Sadoff, F. (2020). Improving college instruction through incentives. *Journal of Political Economy*, 128(8), 2925–2972.
- Cappellari, L., & Lucifora, C. (2009). The “Bologna Process” and college enrolment decisions. *Labour Economics*, 16(6), 638–647.
- Cardoso, S., Rosa, M. J., & Stensaker, B. (2016). Why is quality in higher education not achieved? The view of academics. *Assessment & Evaluation in Higher Education*, 41(6), 950–965.
- Carrieri, V., D’Amato, M., & Zotti, R. (2015). On the causal effects of selective admission policies on students’ performances: Evidence from a quasi-experiment in a large Italian university. *Oxford Economic Papers*, 67(4), 1034–1056.
- Cattaneo, M., Malighetti, P., Meoli, M., & Paleari, S. (2017). University spatial competition for students: The Italian case. *Regional Studies*, 51(5), 750–764.
- Cheng, N. S. (2015). A comparison of compliance and aspirational accreditation models: recounting a university’s experience with both a Taiwanese and an American accreditation body. *Higher Education*, 70(6), 1017–1032.
- Christopher, H., & Hood, C. (2006). Gaming in targetworld: The targets approach to managing British public services. *Public Administration Review*, 66(4), 515–521.
- Ciani, E., & Mariani, V. (2014). How the labour market evaluates Italian universities. Bank of Italy Occasional Paper, 247.
- De Angelis, I., Mariani, I., & Torrini, R. (2017). New evidence on interregional mobility of students in tertiary education: The case of Italy. *Politica economica*, 33(1), 73–96.
- Deming, D. J., Yuchtman, N., Abulafi, A., Goldin, C., & Katz, L. F. (2016). The value of post-secondary credentials in the labor market: An experimental study. *American Economic Review*, 106(3), 778–806.
- Denning, J. T. (2019). Born under a lucky star financial aid, college completion, labor supply, and credit constraints. *Journal of Human Resources*, 54(3), 760–784.
- De Philippis, M. (2021). Multi-task agents and incentives: The case of teaching and research for university professors. *The Economic Journal*, 131(636), 1643–1681.
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1999). An event history model of student departure. *Economics of Education Review*, 18(3), 375–390.
- Dill, D. D., & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education*, 49(4), 495–533.
- Di Pietro, G. (2004). The determinants of university dropout in Italy: A bivariate probability model with sample selection. *Applied Economic Letters*, 11(3), 187–191.
- Di Pietro, G. (2006). Regional labour market conditions and university dropout rates: Evidence from Italy. *Regional Studies*, 40(6), 617–630.
- Di Pietro, G., & Cutillo, A. (2006). University quality and labour market outcomes in Italy. *Labour*, 20(1), 37–62.
- Di Pietro, G., & Cutillo, A. (2008). Degree flexibility and university drop-out: The Italian experience. *Economics of Education Review*, 27(5), 546–555.
- Dynarski, S., & Scott-Clayton, J. (2013). Financial Aid Policy: Lessons from research. *The Future of Children*, 23(1), 67–91.
- ESG (2015). Standards and Guidelines for Quality Assurance in the European Higher Education Area. Brussels, Belgium. [https://www.enqa.eu/wp-content/uploads/2015/11/ESG\\_2015.pdf](https://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf)
- Figlio, D. N., Schapiro, M. O., & Soter, K. B. (2015). Are tenure track professors better teachers? *Review of Economics and Statistics*, 97(4), 715–724.
- Francesconi, M., Aina, C., & Cappellari, L. (2011). Student performance may not improve when universities are choosier. *Evaluation & Research in Education*, 24(4), 231–242.
- Garibaldi, P., Giavazzi, F., Ichino, A., & Rettore, E. (2012). College cost and time to complete a degree: Evidence from tuition discontinuities. *Review of Economics and Statistics*, 94(3), 699–711.

- Gitto, L., Minervini, L. F., & Monaco, L. (2016). University dropouts in Italy: Are supply side characteristics part of the problem? *Economic Analysis and Policy*, 49, 108–116.
- Glocker, D. (2011). The effect of student aid on the duration of study. *Economics of Education Review*, 30(1), 177–190.
- Goldin, C., Katz, L., & Kuziemko, I. (2006). The homecoming of American college women; the reversal of the college. *Journal of Economic Perspectives*, 20(4), 133–156.
- Gunnes, T., Kirkebøen, L. J., & Rønning, M. (2013). Financial incentives and study duration in higher education. *Labour Economics*, 25, 1–11.
- Gury, N. (2011). Dropping out of higher education in France: A micro-economic approach using survival analysis. *Education Economics*, 19(1), 51–64.
- Häkkinen, I. S., & Uusitalo, R. (2003). The effect of a student aid reform on graduation: A duration analysis. Uppsala University Department of Economics Working Paper, 8.
- Hanushek, E. A., Piopiunik, M., & Wiederhold, S. (2019). The value of smarter teachers: International evidence on teacher cognitive skills and student performance. *Journal of Human Resources*, 54(4), 857–899.
- Hanushek, E. A., & Rivkin, S. C. (2006). Teacher quality. In E. A. Hanushek & F. Welch, *Handbook of the economics of education* (Vol. 2, pp 1051–1078). Elsevier.
- Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4), 267–321.
- Harvey, W. B., & Anderson, E. L. (2005). Minorities in Higher Education, 21th Annual Status Report, Washington, DC: American Council on Education.
- Harvey, L., & Newton, J. (2007). Transforming quality evaluation: moving on. In D.F. Westerheijden, B. Stensaker & M.J. Rosa (Eds.), *Quality assurance in higher education* (pp. 225–245). Springer.
- Herzog, S. (2006). Estimating student retention and degree-completion time: Decision trees and neural network vis-à-vis regression. *New Directions for Institutional Research*, 131, 17–33.
- Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41(2), 251–261.
- Horta, H., Dautel, V., & Veloso, F. M. (2012). An output perspective on the teaching-research nexus: An analysis focusing on the United States higher education system. *Studies in Higher Education*, 37(2), 171–187.
- Huisman, J., & Westerheijden, D. F. (2010). Bologna and quality assurance: Progress made or pulling the wrong cart? *Quality in Higher Education*, 16(1), 63–66.
- Jarvis, D. S. (2014). Regulating higher education: Quality assurance and neo-liberal managerialism in higher education—a critical introduction. *Policy and Society*, 33(3), 155–166.
- Jia, P., & Maloney, T. (2015). Using predictive modelling to identify students at risk of poor university outcomes. *Higher Education*, 70(1), 127–149.
- Johnes, G., & McNabb, R. (2004). Never give up on the good times: Student attrition in the UK. *Oxford Bulletin and Statistics*, 66(1), 23–48.
- Kallio, K.-M., Kallio, T. J., & Grossi, G. (2017). Performance measurement in universities: Ambiguities in the use of quality versus quantity in performance indicators. *Public Money & Management*, 37(4), 293–300.
- Lassibille, G., & Navarro Gomez, M. L. (2011). How long does it take to earn higher education degree in Spain? *Research in Higher Education*, 52(1), 63–80.
- Laureti, T., Secondi, L., & Biggeri, L. (2014). Measuring the efficiency of teaching activities in Italian universities: An information theoretic approach. *Economics of Education Review*, 42, 147–164.
- Leiber, T. (2019). A general theory of learning and teaching and a related comprehensive set of performance indicators for higher education institutions. *Quality in Higher Education*, 25(1), 76–97.
- Löfgren, C., & Ohlsson, H. (1999). What determines when undergraduates complete their theses? Evidence from two economics departments. *Economics of Education Review*, 18(1), 79–88.
- MacLeod, W. B., Riehl, E., Saavedra, J. E., & Urquiola, M. (2017). The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics*, 9(3), 223–261.

- McGuinness, S. (2003). University quality and labour market outcomes. *Applied Economics*, 35(18), 1943–1955.
- McNabb, R., Pal, S., & Sloane, P. (2002). Gender differences in educational attainment: The case of university students in England and Wales. *Economica*, 69(275), 481–503.
- Montalvo, J. G. (2018). The impact of progressive tuition fees on dropping out of higher education: a regression discontinuity design (No. 1597).
- Montmarquette, C., Mahseredjian, S., & Houle, R. (2001). The determinants of university dropouts: A bivariate probability model with sample selection. *Economics of Education Review*, 20(5), 475–484.
- Morley, L. (2003). *Quality and power in higher education*. McGraw-Hill Education, UK
- Muscio, A., Ramaciotti, L., & Rizzo, U. (2017). The complex relationship between academic engagement and research output: Evidence from Italy. *Science and Public Policy*, 44(2), 235–245.
- Neave, G., & van Vught, F. (1991). *Changing relationship between government and higher education in Western Europe*. Oxford: Pergamon Press.
- OECD (2019). *Education at a glance 2019: OECD indicators*. Paris: OECD Publishing. <https://doi.org/10.1787/f8d7880d-en>
- Oppedisano, V. (2011). The (adverse) effects of expanding higher education: Evidence from Italy. *Economics of Education Review*, 30(5), 997–1008.
- Page, L. C., Kehoe, S. S., Castleman, B. L., & Sahadewo, G. A. (2019). More than dollars for scholars: The impact of the Dell scholars program on college access, persistence, and degree attainment. *Journal of Human Resources*, 54(3), 683–725.
- Palali, A., Van Elk, R., Bolhaar, J., & Rud, I. (2018). Are good researchers also good teachers? The relationship between research quality and teaching quality. *Economics of Education Review*, 64, 40–49.
- Pascarella, E. T., & Terenzini, P. T. (1978). Student-faculty informal relationships and freshman year educational outcomes. *The Journal of Educational Research*, 71(4), 183–189.
- Power, M. (1997). *The audit society: Rituals of verification*. Oxford: OUP.
- Rebora, G., & Turri, M. (2011). Critical factors in the use of evaluation in Italian universities. *Higher Education*, 61(5), 531–544.
- Rodríguez, R., & Rubio, G. (2016). Teaching quality and academic research. *International Review of Economics Education*, 23, 10–27.
- Ryan, J. F. (2004). The relationship between institutional expenditures and degree attainment at baccalaureate colleges. *Research in Higher Education*, 45(2), 97–114.
- Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In E.A. Hanushek, S. Machin, & L. Woessmann, (Eds.), *Handbook of economics of education* (Vol. 3, pp. 249–277). Amsterdam: Elsevier Science.
- Schudde, L., & Scott-Clayton, J. (2016). Pell grants as performance-based scholarships? An examination of satisfactory academic progress requirements in the nation's largest need-based aid program. *Research in Higher Education*, 57(8), 943–967.
- Scott-Clayton, J. (2011). On money and motivation: a quasi-experimental analysis of financial incentives for college achievement. *Journal of Human Resources*, 46(3), 614–646.
- Scott-Clayton, J., & Schudde, L. (2020). The consequences of performance standards in need-based aid: Evidence from community colleges. *Journal of Human Resources*, 55(4), 1105–1136.
- Sin, C., Tavares, O., & Amaral, A. (2017). The impact of programme accreditation on Portuguese higher education provision. *Assessment & Evaluation in Higher Education*, 42(6), 860–871.
- Singell, D. L. (2004). Come and stay a while: Does financial aid effect retention conditioned on enrollment at a large public university? *Economics of Education Review*, 23(5), 459–471.
- Smith, J. P., & Naylor, A. R. (2001). Dropping out of university: A statistical analysis of the probability of withdrawal for UK university students. *Journal of Royal Statistical Society: Series A*, 164(2), 389–405.
- Stensaker, B. (2008). Outcomes of quality assurance: A discussion of knowledge, methodology and validity. *Quality in Higher Education*, 14(1), 3–13.

- Stensaker, B. (2011). Accreditation of higher education in Europe-Moving towards the US model? *Journal of Education Policy*, 26(6), 757–769.
- Stinebrickner, T., & Stinebrickner, R. (2006). What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics*, 90(8–9), 1435–1454.
- Stinebrickner, T., & Stinebrickner, R. (2008). The effect of credit constraints on the college dropout decision: A direct approach using a new panel survey. *American Economic Review*, 98(5), 2163–2184.
- Stinebrickner, T., & Stinebrickner, R. (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4), 707–748.
- Stinebrickner, T., & Stinebrickner, R. (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics*, 32(3), 601–644.
- Stratton, L. S., O’Toole, D. M., & Wetzel, J. N. (2008). A multinomial logit model of college stopout and dropout behavior. *Economics of Education Review*, 27(3), 319–331.
- Tavares, O., Sin, C., Videira, P., & Amaral, A. (2017). Academics’ perceptions of the impact of internal quality assurance on teaching and learning. *Assessment & Evaluation in Higher Education*, 42(8), 1293–1305.
- Tian, Z., Wei, Y., & Li, F. (2019). Who are better teachers? The effects of tenure-track and part-time faculty on student achievement. *China Economic Review*, 53, 140–151.
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45, 89–125.
- Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *The Journal of Higher Education*, 68(6), 599–623.
- Trivellato, P., & Triventi, M. (2009). Participation, performance and inequality in Italian higher education in the 20th century. *Higher Education*, 57(6), 681–702.
- Vidovich, L. (2002). Quality assurance in Australian higher education: Globalisation and-steering at a distance. *Higher Education*, 43(3), 391–408.

**Massimiliano Bratti** (Ph.D. University of Warwick) is Professor of Economics at the University of Milan. He specializes in the economics of education and the determinants and effects of educational choices.

**Giovanni Barbato** (Ph.D. University of Milan) is a Postdoctoral Researcher at the University of Milan. His research interests are on the measurement and evaluation, performance and strategies of public sector organizations.

**Daniele Biancardi** (Ph.D. University of Milan) is a Postdoctoral Researcher at the University of Turin. His research area are in education economics and labour economics.

**Chiara Conti** (Ph.D. University of Milan) is Assistant Professor at Sapienza University of Rome. Her main research fields are innovation, regulation, and the economic impact of public policies.

**Matteo Turri** (Ph.D. University Carlo Cattaneo) is Professor of Business Administration at the University of Milan. His research interests are public administration and higher education.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter'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.





# Teaching Efficiency of Italian Universities: A Conditional Frontier Analysis



Camilla Mastromarco, Pierluigi Toma, and Cinzia Daraio

**Abstract** The aim of this chapter is to provide a comparative analysis of the performance of teaching at Italian university by evaluating the efficiency of heterogeneous faculty courses at the national level. For this purpose, we use advanced and robust nonparametric tools recently developed in nonparametric efficiency frontier literature. This performance assessment does not rely on hypotheses about the relationship between inputs and outputs and allows us to account for the heterogeneity of the analyzed courses. The overall analysis carried out at the national level for Italy extends the traditional and limited one-dimensional indicators available through SUA-CdS data and ad hoc surveys on graduates conducted by ANVUR (National Agency for the Evaluation of Universities and Research Institutes) and MIUR (Ministry of Education, University and Research). The estimated efficiency scores are used to analyze current trends and changes in the teaching activities of Italian universities.

**Keywords** Conditional efficiency methodology · Teaching efficiency · Italian university system · Heterogeneity

---

**Supplementary Information** The online version contains supplementary material available at [https://doi.org/10.1007/978-3-031-07438-7\\_11](https://doi.org/10.1007/978-3-031-07438-7_11).

---

C. Mastromarco (✉)

Department of Economics, Statistics and Finance – Arcavacata, University of Calabria, Rende, Italy

e-mail: [camilla.mastromarco@unical.it](mailto:camilla.mastromarco@unical.it)

P. Toma

Department of Economics, Management and Quantitative Methods Ecotekne, University of Salento, Lecce, Italy

e-mail: [pierluigi.toma@unisalento.it](mailto:pierluigi.toma@unisalento.it)

C. Daraio

Department of Computer, Control and Management Engineering “A. Ruberti,” Sapienza University of Rome, Rome, Italy

e-mail: [cinzia.daraio@uniroma1.it](mailto:cinzia.daraio@uniroma1.it)

© The Author(s) 2022

D. Checchi et al. (eds.), *Teaching, Research and Academic Careers*,  
[https://doi.org/10.1007/978-3-031-07438-7\\_11](https://doi.org/10.1007/978-3-031-07438-7_11)

## 1 Introduction

Economic theory models public entities' behavior using paradigms typical of the economic sciences. However, empirical econometric or operational research is needed to obtain comparative evaluations and effective policy implications (Daraio, 2018; Johnes, 2006, 2015; Johnson & Ruggiero, 2014; Ruggiero, 1996). In the case of universities, the quantity, quality, and mix of services produced are largely due to autonomous decisions, influenced by preferences of the different categories of stakeholders (Klumpp, 2015; Nigsch & Schenker-Wicki, 2015). Universities in the current period of crisis are at a crossroads. Evaluation, rankings, and governance are at the core of recent policy agendas, because it is crucial to invest in science and education to fully implement revitalizing strategy in terms of innovation and growth (Daraio et al., 2019).

Universities can be seen as “loosely coupled systems” characterized by autonomous decision-making processes with respect to the quality, quantity, and mix of products and services provided (Bonaccorsi & Daraio, 2007a, 2007b, 2007c). Italian universities, as all other universities over the world, are involved in a series of institutional activities, namely teaching, research, and the diffusion of knowledge in society (called “Third Mission”). An important limitation of available Italian university rankings relies on an inadequate consideration of the role of teaching, one of the fundamental pillars of academic activity, causing information bias regarding the performance of individual universities and the academic system as a whole.

This chapter investigates the performance of university teaching by evaluating the efficiency of different faculty courses. The aim is to propose a new measure of university efficiency, where the unit of evaluation is not the university but the course of study. The aim is to highlight the diversity and autonomy of individual universities in relation to teaching organization, providing an empirical measure, based on comparable available data, of the results obtained in terms of the efficiency of the various courses of study.

A preliminary analysis was carried out on the University of Salento data, and the results are reported in Mastromarco et al. (2019). We extend the analysis at the national level, including all courses of all Italian universities. For this purpose, we use advanced and robust nonparametric tools recently developed in nonparametric efficiency frontier literature (Bădin et al., 2010, 2012, 2019; Daraio et al., 2018; Daraio & Simar, 2005, 2007a, 2007b).

The novelty in the performance assessment of university teaching activities consists in an extension of traditional and limited indicators, extracting information available through the SUA-CdS data sheet and ad hoc surveys on graduates conducted by ANVUR (National Agency for the Evaluation of Universities and Research Institutes) and MIUR (Ministry of Education, University and Research), in order to improve the quality of teaching monitoring and promote its dissemination through the different universities.

The chapter is organized as follows. In the next section, we summarize the relevant literature. Section 3 outlines the methodology. Section 4 describes the data and the empirical strategy. Section 5 illustrates the results, and Sect. 6 concludes the chapter. Finally, the [Appendix](#) reports additional details on the methodology.

## 2 Brief Literature Review

Although the role of universities in the knowledge society is increasingly relevant, there is a lack of systematic quantitative evidence at the micro level and, in particular, at the level of individual faculty courses. Bonaccorsi and Daraio (2007a, 2007b, 2007c) examined original data from universities in six European countries, including Italy, by applying for the first time new generations of nonparametric efficiency measures on a large scale and providing micro-based evidence on the evolution of the strategic profile of universities in terms of scientific research, contract research, education, and the Third Mission. In another study, Agasisti and Johnes (2010) evaluated the efficiency of Italian universities, demonstrating that there is a close relationship between the size and efficiency of universities; moreover, in their work, they highlighted that the growth in the size of universities reduces the overall efficiency of scientific research in the same. Even earlier, more specific analysis on the efficiency of teaching by Italian universities includes: Ferrari and Laureti (2005), Laureti (2008) and Laureti et al. (2014). A rich survey on the efficiency of universities can be found in Worthington (2001); while De Witte and López-Torres (2017) is the most recent and comprehensive survey available in the field.

The main critical points that emerge from the empirical literature relating to the evaluation and rankings of universities can be summarized as follows: (i) one-dimensionality; (ii) lack of statistical robustness; (iii) dependence on the size of the university and on its subject mix; (iv) lack of consideration of the input–output structure. See Daraio et al. (2015a, 2015b) and Daraio and Bonaccorsi (2017) for a deeper description of the literature on rankings and empirical investigations. In addition to these critical aspects, university education is subject to rapid changes, due to continuous reforms, which make the evaluation process particularly complicated. In Italy (see, e.g., Agasisti & Dal Bianco, 2009), thanks to law n.270/2004, which reformed the university system, the differentiation between three-year bachelor courses and master’s degree courses has grown considerably, gradually allowing for the adoption of different teaching practices depending on the level of university education. This aspect has not yet been adequately assessed and monitored, even though it concerns one of the main activities of universities, which involves almost two million students every year, and which represents the first source of funding for universities (both public and private) and therefore the main area of competition.

The difficulty in applying efficiency methodologies at the level of individual degree courses consists mainly in finding the data: in fact, as reported in the comprehensive analysis of the literature proposed by De Witte and López-Torres

(2017), only two papers, Cooper and Cohn (1997) and De Witte and Rogge (2011), carried out a study of the single degree courses, and neither used the methodology applied here, which allows for taking into account the temporal dimension, and therefore the adjustments dynamic, following the work of Mastromarco and Simar (2015).

In this chapter, we aim to develop new models for estimating the efficiency of Italian faculty courses exploiting the new information contents that can be processed from the SUA-CdS forms, useful for identifying the most efficient teaching practices for better monitoring the performance of teaching, comparatively, among Italian universities.

The indicators identified for the analysis of efficiency are detailed in Sect. 4, and relate to basic institutional information, geographic information, training activities conducted, personnel (including gender), and size.

### 3 Methodology

Our econometric approach is based on recent developments of Data Envelopment Analysis (DEA), which originated from the seminal papers of Farrell (1957) and Charnes, Cooper, and Rhodes (1978). DEA uses linear programming to compare and benchmark a sample of observed units—in our study, faculty courses—against the efficient production frontier, which consists of combinations of observed production possibilities. DEA relies on a minimum set of hypotheses that are: i) free disposability (that is, the possibility to destroy goods without costs) and ii) convexity.

The main advantage of the approach is its multidimensionality; i.e., multi-input multi-output performance evaluation, without any assumption on the functional relationship between inputs and outputs. DEA is nonparametric because it does not make assumptions about the distribution of inefficiencies or the functional form of the production function. On the contrary, it uses the input and output data themselves to estimate the production possibility frontier. This nonparametric approach does not require assumptions about the behavior of the analyzed units, such as cost minimization or profit maximization, which are not appropriate for the Higher Education context, and does not require the knowledge of input and output prices, which are often unknown in the Higher Education context (Daraio, 2018; Johnes, 2006). In this study we use advanced and robust nonparametric tools recently developed in nonparametric efficiency frontier literature called “conditional efficient frontier” models (Cazals et al., 2002; Daraio & Simar, 2005; Mastromarco & Simar, 2015) whose main ideas are outlined below. Additional methodological details are reported in the [Appendix](#).

The aim of this study is the analysis of the performance of teaching at Italian universities by comparing the efficiency of university faculty courses at the national level. For this purpose, we consider the university faculty courses as the relevant unit of analysis, whose efficiency in producing knowledge can be evaluated by applying

a conditional efficiency approach. In an efficiency analysis, performances are measured with respect to the best-practice frontier (or efficient frontier) constructed by comparing the outputs achieved, given the inputs possessed, of the units analyzed. In a conditional efficiency analysis, conditioning or environmental factors are included in the measurement of performance. These factors are neither inputs nor outputs of the production process but may influence the performance of the units analyzed (see, e.g., Daraio & Simar, 2007a).

Let the production process of teaching activities in university faculty courses be characterized by one *input*  $X \in \mathbb{R}_+$  (number of teachers weighted by teaching hours), and two *outputs*  $Y \in \mathbb{R}_+^2$ , (the percentage of graduated students within the legal duration of the degree course and the percentage of graduates satisfied with their faculty course).

The objective is to evaluate the efficiency with which the inputs (number of teachers compared to the number of students, weighted by the hours of teaching) of the different study programs determines the outputs (percentage of graduates within the legal duration of the degree course and percentage of students satisfied with the course of study). This model considers the quality of teaching as given and not observable, and instead analyzes the teachers/students ratio in relation to the completion of the course of study on time and student-satisfaction rate. It would also be important to take into account the quality of teaching, but data limitation does not allow us to control for quality. However, our conditional approach enables us to include other factors, which are neither inputs nor outputs, which may affect and limit the units under analysis in terms of reaching the efficient output levels. We assume that there are some external variables  $Z \in \mathbb{R}_+^d$  that may influence the performance of the faculty courses. We consider the size of the different faculty courses, which may capture the heterogeneity among them. To avoid distortion, due to time delays or time adjustments, in the evaluation of the performance of different faculty courses, we take into account temporal dynamics.

Being nonparametric, the conditional efficiency models suffer from the so-called curse of dimensionality, which means the need to use parsimonious models in terms of input and output numbers to avoid inaccurate estimates of efficiency scores.

To overcome the curse of dimensionality, then, we limit the dimension of our model and consider only one input (weighted teacher-to-students ratio), two outputs (% of graduates within regular duration and student-satisfaction rate), and two environmental variables (size and time).

## 4 Data and Empirical Strategy

Data come from SUA-CdS forms and ad hoc surveys on graduates conducted by ANVUR and MIUR and refer to the period 2013–2017.

The analysis was conducted at the level of a single university course of study, using annual data from 2013 to 2017. The entire dataset was divided into three-year degree courses and master's degree courses in order to make the analysis

more homogeneous. We have accounted for the heterogeneity among disciplines by carrying out the analysis distinguishing among subject areas, considering sciences, health sciences, social sciences, and humanities. This approach makes the empirical analysis more homogenous and allows us to consider the peculiarities of the courses in different subject areas, as with, for example, the infrastructure and the size in health science and science degree courses.

Unique-cycle degree courses (like law and veterinary) were annexed to the datasets related to master's degree courses. In this way, we obtained eight datasets and, therefore, eight analyses: the three-year degree courses were divided into the four subject areas, as were the master's degree courses. The summary statistics and the results of the analysis are therefore being presented for each subject area.

All degree courses with a temporal dimension fewer than three years were eliminated from the dataset. This threshold was chosen in accordance with the period needed to conclude one study cycle that is a three-year period. The same threshold is useful in order to improve the homogeneity of the analysis and to be able to consider the temporal trend of the efficiency of different degree courses. After cleaning the data, we obtained 1907 BA courses and 2138 master's courses.

Table 1 shows the distribution of the degree courses and master's courses by disciplinary area. We note that in Italy there is a prevalence of courses in the scientific field as scientific courses represent 41% and 45% of the overall degree courses and overall master's courses, respectively.

The purpose of this study is to analyze the performance, in terms of efficiency, of the various university courses. We chose the following outputs:

- (i) The “percentage of graduates within the legal duration of the degree course” (iC02)
- (ii) The “percentage of graduates overall satisfied with the study course” (iC25).

The chosen input is the “number of teachers over the number of students (weighted by teaching hours),” which is the inverse of the ANVUR iC27 indicator. As an external factor, we use “size”; that is, the number of enrolled students of the considered study courses.

See Table 2 for some descriptive statistics on the input, outputs, and external factor “size” by degree and master's courses and by discipline. It can be seen that the areas with the highest number of students are those related to Science and

**Table 1** Number of analyzed courses by subject area

Disciplinary field	n. of three-year degrees	n. of three-year degrees (%)	n. of master's degrees	n. of master's degrees (%)
Science	778	40.80	965	45.14
Health Science	457	23.96	237	11.09
Social Science	410	21.50	596	27.88
Humanities	262	13.74	340	15.90
<b>Total</b>	<b>1907</b>	<b>100</b>	<b>2138</b>	<b>100</b>

Note: “Bold letter” indicates female gender predominance

**Table 2** Summary statistics on input, outputs, and size

Type	Disciplinary Area	Indicator	Size	iC02	iC25	iC27	
Three-year degrees	<b>Science</b>	<b>Average</b>	13,014.88	0.44	0.90	0.05	
		Std. dev.	10,202.50	0.20	0.08	0.04	
	<b>Health Science</b>	<b>Average</b>	7053.33	0.66	0.86	0.23	
		Std. dev.	4975.51	0.18	0.13	0.25	
	<b>Social Science</b>	<b>Average</b>	11,853.33	0.50	0.89	0.03	
		Std. dev.	8018.43	0.22	0.06	0.02	
	<b>Humanities</b>	<b>Average</b>	7225.37	0.49	0.87	0.03	
		Std. dev.	5209.80	0.18	0.08	0.02	
	Master's degrees	<b>Science</b>	<b>Average</b>	13,443.39	0.55	0.90	0.13
			Std. dev.	10,305.51	0.25	0.09	0.10
<b>Health Science</b>		<b>Average</b>	6866.85	0.64	0.85	0.08	
		Std. dev.	5022.54	0.25	0.14	0.06	
<b>Social Science</b>		<b>Average</b>	12,961.06	0.64	0.90	0.06	
		Std. dev.	8241.70	0.21	0.08	0.05	
<b>Humanities</b>		<b>Average</b>	7588.54	0.52	0.89	0.09	
		Std. dev.	5209.83	0.22	0.09	0.10	

Note:: "Bold letter" indicates female gender predominance

Social Science. The average of "percentage of graduates within the legal duration of the degree course" (iC02) is quite similar between the different disciplines, while the average of "percentage of graduates overall satisfied with the course of study" (iC25) varies greatly depending on the subject area. The input "number of teachers over the number of students (weighted by teaching hours)" (the inverse of the iC27) has a lot of heterogeneity between the different groups of degrees.

The objective is to evaluate the efficiency with which the input (number of teachers compared to the number of students, weighted by the hours of teaching) of the different study programs determines the outputs (percentage of graduates within the legal duration of the degree course and percentage of students satisfied with the course of study). Regarding the choice of the input variable, we need to make some considerations. The Italian education system does not report costs and human resources at the level of a single university course of study. In general, the data are aggregated at the level of departments (or faculties). The input indicator we chose, then, is the only one available at the study-course level. Nevertheless, we think it may adequately proxy the concept of teaching resource at the study-course level.

The data are from a dashboard of indicators that each university uses for its own evaluation and self-assessment. These indicators are the first and only source of data available at the level of individual university courses. The originality of the work consists in the comparative efficiency analysis conducted at the level of the single course of study, carried out for the first time in the education literature considering all the courses at the national level.

As described in Sect. 3, the methodology applied in this work is conditional efficiency. It allowed us to obtain the conditional efficiency on some environmental (or contextual) factors that may influence the production process. We conditioned on size, as the number of enrolled students in the subject area of the study course. In particular, we analyzed whether and how the number of students influences the efficiency of the course. Size was measured by the number of students enrolled in bachelor's and master's degree courses of the specific subject area of the course analyzed. For example, the efficiency of the degree in economics at the University of Salento was conditioned by the number of students enrolled in social science degrees at the University of Salento.

Following the approach proposed by Mastromarco and Simar (2015), we carried out a time-dependent analysis, which allowed us to measure the time-dependent efficiency of university courses and to assess the effect of time on the performance by taking into account time delays and adjustment lags.

To illustrate the conditional efficiency scores calculated for each course of study, we considered (i) the geographic area of the universities and (ii) the gender composition of teachers.

We distinguished between universities in central-Northern and Southern Italy. This focus greatly strengthens the policy recommendations for a strategic sector, such as education, for the economic development of Italy, which has always suffered from serious geographic disparities.

Considering the importance of gender balance, especially in advanced studies in the STEM (Science, Technology, Engineering, and Mathematics) fields, we calculated an indicator of *gender prevalence*. Regarding this topic, to which policymakers have recently been paying attention, we found it difficult to find data at the level of an individual study course. To overcome this problem, our indicator was calculated for each International Standard Classification of Education (ISCED) Field of Education and Training (FOET2013) of each university as follows. If in each FOET2013, the number of female-dominated courses exceeded the number of male-dominated courses, the field was assigned the *female-oriented* discipline label. If the number of male-dominated courses exceeded the number of female-dominated courses, the FOET2013 classification of the university was assigned the *male-oriented* discipline label. If the number of male-dominated courses was equal to the number of female-dominated courses, the field was assigned the *gender-neutral* discipline label. Finally, the relevant label was assigned to each degree course based on the correspondence between the Italian degree classes and the FOET2013 nomenclature.<sup>1</sup>

---

<sup>1</sup> This correspondence is available at the Ministry of Education, University and Research website <https://www.miur.gov.it/documents/20182/1287773/DD+n.+389+ALLEGATO+2.pdf/e6ec2148-843a-4d9d-b683-26b4ff45d3a2?version=1.0&t=1551954744102>



## 5 Results

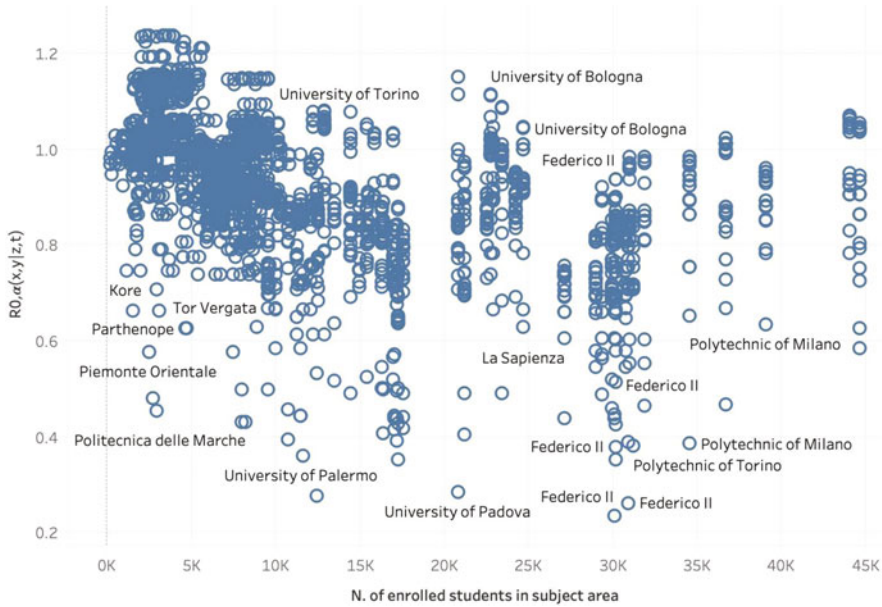
In the first step, we studied the impact of time and size on the efficient frontier of the production process of the study courses. This is done by investigating the ratios of conditional and unconditional efficiency scores for the robust full frontier calculated with  $\alpha = 0.99$  (see [Appendix](#)). Subsequently, we analyze the effect of the conditioning variables (time and size) on the distribution of efficiency—that is, on the distance of the units from the efficient frontier—by inspecting the graph related to the robust partial frontier estimated in the middle of data with  $\alpha = 0.5$  (see [Appendix](#)).

The second step involved obtaining the average efficiency results of the individual university degree courses. For the sake of clarity, only the results of the 20 best (on average) degree courses and 20 worst (on average) courses of study for each subject area are presented here.

The third and last part of the results focuses on the efficiency trends of the study courses, with particular attention given to the average variation of efficiency scores. Emphasizing the tendency of improvement or deterioration is an important aspect in evaluating a university course. The purpose of the efficiency analysis is not, in fact, to punish or reward educational institutions but to stimulate a process of change. Following the Mastromarco and Simar (2015) approach, we were able to evaluate in a robust way the dynamics of efficiency during the time of each course, which provides useful insights. Scatterplots are shown to visualize the relationship between improvements and efficiency starting levels. In the following we present the results, tables, and figures of the three-year degree courses. For space reasons, in the supplementary material, available on line, we report all the figures for the master's degree courses.

### 5.1 Effect of Size on Efficiency

We investigated the effect of size and time on the boundary of the efficient frontier, hence on the best-performing courses, in different areas. We started with science courses that represent more than 40% of our sample. Figure 1 shows the impact of size on the efficiency of three-year degree courses in science. Size (on the x-axis) is the number of enrolled students in science. On the y-axis  $R_0(x,y|z,t)$  ( $\alpha = 0.99$ , see [Appendix](#) for more details) are the ratios of the conditional to unconditional efficiency scores. An increasing (decreasing) trend of ratios identifies a positive (negative) impact of the dimension on the efficient frontier of courses in science. A flat trend shows no effect of size on the efficient frontier. Inspecting Fig. 1, we note an important effect of the size—as number of enrolled students in the subject area—on the efficient frontier. In particular, there is, first, a negative and then positive effect on the efficiency frontier (i.e., the maximum achievable output values, given the available input). This suggests a negative effect of number of students



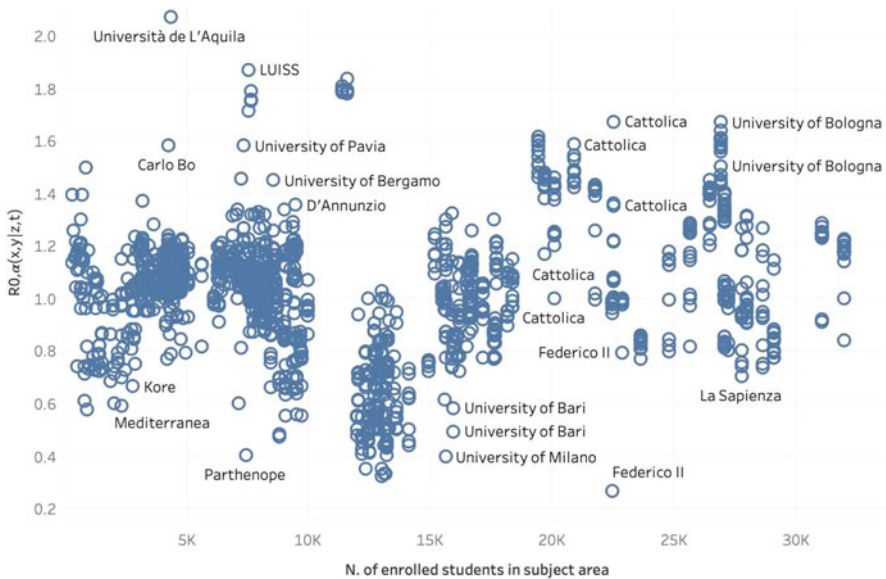
**Fig. 1** Impact of size on the efficient frontier of three-year degree courses in science. Size (on the x-axis) is the number of enrolled students in science. On the y-axis  $R_{0,\alpha}(x,y|z,t)$  ( $\alpha = 0.99$ , see Appendix for more details) are the ratios of the conditional to unconditional efficiency scores. An increasing (decreasing) trend of ratios identifies a positive (negative) impact of size on the efficient frontier of courses in science. A flat trend shows no effect of size on the efficient frontier

in small-medium universities and a positive effect of size on the most efficient courses in sciences in larger universities. This finding is explained by the fact that science courses need specialized structures as well as laboratories. Thus, only larger universities, which are able to invest in large infrastructures, may obtain good results when they increase the number of students enrolled. Indeed, the best courses in science in big universities may afford new investments to avoid congestion costs. Figure 1 also highlights some universities with the best courses in science.

The effect of the number of students enrolled in degree courses in the health sciences does not seem relevant for the most efficient courses. This result is not surprising; in fact, in order to have a degree course in health sciences, universities must collaborate with important clinical centers (e.g., hospitals), and the number of students who can access medical degrees is limited by law. In addition, the presence of some medical laboratories and other infrastructures is necessary to offer these courses. This makes the size of this area, measured as the number of students enrolled, similar for all universities with medical degree courses and, hence, not particularly influential. To save space, we do not include graphs for the other university areas, which are available upon request.

A result similar to health science is obtained for the degree courses in social sciences and humanities. In these cases, since no particular laboratories or infras-

structure is necessary, the degree courses allow for a great deal of flexibility in the number of enrolled students. Despite this great variability of size, as can also be seen from the descriptive statistics in Table 2, there is no particular influence of the size on the efficiency of the best courses in these areas. Concerning the analysis of the possible effect of the size on the distribution of the efficiency scores of the university courses, we evaluated the trend of the ratios between the conditional and the unconditional efficiency scores using a partial frontier with  $\alpha = 0.5$ , which captures the middle of the distribution of the efficiency (see Appendix). For sciences, the effect of size on the ability of degree courses below the frontier to reach full efficiency (i.e., maximum outputs value, given the available input) is similar to the previous case on the most efficient degree courses (see Fig. 1). In particular, size (i.e., the number of students enrolled in the area) has a negative impact on universities with a medium- to small-sized scientific area, while it has a positive impact on universities with a bigger scientific area. The effect of size in the health sciences does not seem relevant for the achievement of efficient output values of three-year degree courses. For social science courses (see Fig. 2), size has a slightly negative effect for medium-small universities and a positive effect for larger ones. Furthermore, for universities with a number of enrolled students higher than 20,000, a slight negative effect of size on the efficiency of university courses over time is



**Fig. 2** Impact of size on the distribution of the efficiency of three-year degree courses in social science. Size (on the x-axis) is the number of enrolled students in social science. On the y-axis  $R_{0,\alpha}(x,y|z,t)$  ( $\alpha = 0.50$ , see Appendix for more details) are the ratios of the conditional to unconditional efficiency scores. An increasing (decreasing) trend of ratios identifies a positive (negative) impact of size on the distribution of the efficiency of courses in social science. A flat trend shows no effect of size on the distribution of the efficiency

observed. This finding demonstrates how excess overcrowding can cause problems and negatively affect the efficiency of the courses.

For humanities, the number of enrolled students has a negative effect on efficiency of courses in universities with a small number of students, while it has a positive effect on efficiency of courses in medium-sized universities (i.e., 6000–11,000 enrolled students). The effect is again positive for universities with over 12,000 students.

## 5.2 *Analysis of the Efficiency Scores*

We now analyze the twenty best and worst degree courses over the five years under analysis. The top panel of Table 3 shows the ranking of the 20 most efficient degree courses in science, among which 19 universities are located in the center-North and only one university in Southern Italy. This result is not surprising. It is well known that the industrial sector is a natural job outlet for science degrees. The universities in Northern Italy have more relationships with companies and industries, as they are concentrated in the North of the country. As a result, there are many job opportunities for students who may build relationships with these companies during their university studies. In this area, with strong STEM characterization, from a gender point of view, we find a female predominance in six out of the 20 best courses (highlighted in bold) and one neutral. Among the twenty least efficient degree courses (bottom of Table 3), we find 14 degree courses in Southern Italy and six in the center-North, which is substantially reversed compared to the ranking of the best courses. From a gender point of view, we find five degree courses with female prevalence among the worst ones. In both rankings—those with the best and those with the worst courses in science—we find female gender-dominated degree courses in limited numbers. This result is common in the science area, as confirmed in the scientific literature on the subject (see Card & Payne, 2021). Despite this evidence, in Italy, in recent years, there has been a change in this trend.

In the ranking of the 20 most efficient degree courses in the health sciences, shown in the top panel of Table 4, we only find courses in the center-North, with a very high frequency of degree courses in Milan universities (11 out of 20). The regional aspect is important for health degree courses. Any university offering a course in medicine is required to establish an agreement with the regional health system. Therefore, it is clear that interregional differences in health systems are also reflected in this ranking. All of these courses are female-gender prevalent (highlighted in bold in the table). This result is not surprising, because it is well known that the presence of women is predominant in this disciplinary area. Among the 20 least efficient degree courses, displayed in the bottom panel of Table 4, considering an average of the period being analyzed, we find 13 degree courses in Southern Italy and seven in the center-North; therefore, this situation is substantially reversed compared to the ranking of the best university courses. Only one course

**Table 3** List of the 20 most efficient three-year degree courses in the scientific area (top) and the 20 least efficient three-year degree courses in the scientific area (bottom)

Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
Ingegneria gestionale	L-9	University of Bologna	07
Ingegneria informatica	L-8	Polytechnic University of Turin	06
Biologia molecolare	L-13	University of Padua	05
Ingegneria Biomedica	L-8	University of Genoa	06
Biotechnologie	L-2	University of Padua	05
Ingegneria dell'Automazione	L-8	Polytechnic University of Milan	06
<i>Design per la moda</i>	L-4	University of Campania Luigi Vanvitelli	02
Ingegneria e Scienze Informatiche	L-8	University of Bologna	06
Biotechnologie*	L-2	University of Pisa	05
<i>Biotechnologie</i>	L-2	University of Turin	05
Ingegneria Gestionale	L-8	Sapienza University of Rome	06
Biotechnologie	L-2	University of Eastern Piedmont	05
Ingegneria dei Materiali e delle Nanotecnologie	L-9	Polytechnic University of Milan	07
Fisica	L-30	University of Pisa	05
<i>Scienze dell'Alimentazione e della Nutrizione Umana</i>	L-13	Biomedical University of Rome	05
Ingegneria Meccanica	L-9	University of Modena and Reggio Emilia	07
<i>Biotechnologie</i>	L-2	Sapienza University of Rome	05
<i>Design della moda e arti multimediali</i>	L-4	Iuav University of Venice	02
<i>Design del prodotto industriale</i>	L-4	University of Bologna	02
Ingegneria Clinica	L-9	Sapienza University of Rome	07
Less Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Biotechnologie</i>	L-2	Basilicata University	05
Ingegneria Civile e Ambientale	L-7	Basilicata University	07
Fisica	L-30	University of Catania	05
<i>Corso di laurea in Ottica e Optometria</i>	L-30	University of Naples Federico II	05
Ingegneria dell'Edilizia	L-23	University of Rome II "Tor Vergata"	07
Ingegneria edile	L-23	University of Bologna	07
Ingegneria Elettronica e Informatica	L-8	University of Messina	06
Ingegneria Industriale	L-9	University of Messina	07

(continued)

**Table 3** (continued)

Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<b>Scienze e Tecnologie Informatiche</b>	L-31	Basilicata University	06
Ingegneria Civile	L-7	University of Campania Luigi Vanvitelli	07
Ingegneria per l'ambiente e il territorio	L-7	University of Cagliari	07
<b>Scienze geologiche</b>	L-34	University of Calabria	05
Ingegneria Civile e Ambientale	L-7	University of Rome II "Tor Vergata"	07
Ingegneria civile e ambientale	L-7	University of Catania	07
Ingegneria Medica	L-9	University of Rome II "Tor Vergata"	07
Ingegneria civile e dei sistemi edilizi	L-7	University of Messina	07
<b>Design e Arti</b>	L-4	Free University of Bozen-Bolzano	02
Ingegneria Civile e Ambientale	L-7	Mediterranean University of Reggio Calabria	07
Scienze Ambientali e Naturali	L-32	University of Catania	05
Ingegneria Edile	L-23	University of Naples Federico II	07

Note: "Bold letter" indicates female gender predominance and "\*" gender neutral

(the worst, in terms of efficiency) is male-gender prevalent, and three out of 20 are neutral-gender prevalent.

In the ranking of the 20 most efficient degree courses in the social sciences, shown in the top panel of Table 5, we only find courses from the central-Northern regions, with a very high frequency of degree courses in Milan universities (12 out of 20). This evidence depends on the presence, in the center-North, of universities considered to be excellent in the economic and social sciences. Many of these courses (15 out of 20) are female oriented. This result is due to pedagogy and social science courses, which mainly attract female students (Francesconi & Parey, 2018). Continuing with the evaluation of the average efficiency of courses, the bottom panel of Table 5 reports the 20 degree courses deemed least efficient. We find 12 degree courses from Southern Italy and eight from central-Northern Italy; therefore, this situation is substantially reversed with respect to the ranking of the most efficient social science courses. All of these courses, except the first one, are female oriented (highlighted in bold). This result is in line with what was previously described—namely, the fact that many courses of study in this area find the presence of many female students.

In the ranking of the 20 most efficient degree courses in the humanities, displayed in the top panel of Table 6, we find many courses in the central-Northern regions, but, unlike the previously analyzed area (social sciences), the frequency of degree

**Table 4** List of the 20 most efficient three-year degree courses in the health science area (top) and the 20 least efficient three-year degree courses in the health science area (bottom)

Health Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Tecniche audioprotesiche</i>	L/SNT3	University of Milan	09
<i>Igiene dentale</i>	L/SNT3	Vita-Salute San Raffaele University	09
<i>Logopedia</i>	L/SNT2	Catholic University of the Sacred Heart	09
<i>Infermieristica</i>	L/SNT1	University of Bologna	09
<i>Terapia occupazionale</i>	L/SNT2	Catholic University of the Sacred Heart	09
<i>Fisioterapia</i>	L/SNT2	Vita-Salute San Raffaele University	09
<i>Tecniche di radiologia medica, per immagini e radioterapia</i>	L/SNT3	Catholic University of the Sacred Heart	09
<i>Tecniche della prevenzione nell'ambiente e nei luoghi di lavoro</i>	L/SNT4	Catholic University of the Sacred Heart	09
<i>Tecnica della riabilitazione psichiatrica</i>	L/SNT2	University of Rome II "Tor Vergata"	09
<i>Terapia della neuro e psicomotricità dell'età evolutiva</i>	L/SNT2	University of Rome II "Tor Vergata"	09
<i>Dietistica</i>	L/SNT3	Catholic University of the Sacred Heart	09
<i>Scienza della Nutrizione</i>	L-29	University of Urbino "Carlo Bo"	05
<i>Fisioterapia</i>	L/SNT2	Catholic University of the Sacred Heart	09
<i>Tecniche ortopediche</i>	L/SNT3	Catholic University of the Sacred Heart	09
<i>Tecniche audioprotesiche</i>	L/SNT3	University of Padua	09
<i>Educazione professionale</i>	L/SNT2	University of Ferrara	09
<i>Infermieristica</i>	L/SNT1	Biomedical University of Rome	09
<i>Educazione professionale</i>	L/SNT2	University of Insubria	09
<i>Ostetricia</i>	L/SNT1	Catholic University of the Sacred Heart	09
<i>Tecniche di radiologia medica, per immagini e radioterapia</i>	L/SNT3	Biomedical University of Rome	09

(continued)

**Table 4** (continued)

Health Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
Less Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Tecniche di laboratorio biomedico</i>	L/SNT3	University of Messina	09
<i>Fisioterapia</i>	L/SNT2	University of Sassari	09
<i>Ostetricia</i>	L/SNT1	University of Firenze	09
Scienze e tecnologie erboristiche*	L-29	University of Milan	05
<i>Tecniche di radiologia medica, per immagini e radioterapia</i>	L/SNT3	Marche Polytechnic University	09
<i>Scienze e tecnologie erboristiche e dei prodotti per la salute</i>	L-29	University of Bari	05
<i>Tecniche di radiologia medica, per immagini e radioterapia</i>	L/SNT3	University of Siena	09
<i>Ostetricia</i>	L/SNT1	University of Catania	09
<i>Tecniche di neurofisiopatologia</i>	L/SNT3	University of Naples Federico II	09
Scienze Tossicologiche e Controllo di Qualità*	L-29	University of Cagliari	05
<i>Ostetricia</i>	L/SNT1	University of Rome II "Tor Vergata"	09
<i>Corso di Laurea interfacoltà in Scienze Erboristiche</i>	L-29	University of Naples Federico II	05
<i>Corso di laurea in Controllo di Qualità</i>	L-29	University of Naples Federico II	05
Scienze e sicurezza chimico tossicologiche dell'ambiente*	L-29	University of Milan	05
<i>Ostetricia</i>	L/SNT1	University of Messina	09
<i>Tecniche Erboristiche</i>	L-29	University of Salerno	05
<i>Informazione Scientifica del Farmaco e dei Prodotti per la Salute</i>	L-29	University of Calabria	05
<i>Scienze Farmaceutiche Applicate</i>	L-29	Sapienza University of Rome	05
<i>Scienza della Nutrizione</i>	L-29	University of Calabria	05
Scienze farmaceutiche applicate	L-29	University of Catania	05

Note: "Bold letter" indicates female gender predominance and "\*" gender neutral



**Table 5** List of the 20 most efficient three-year degree courses in the social science area (top) and the 20 least efficient three-year degree courses in the social science area (bottom)

Social Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Scienze e tecniche psicologiche</i>	L-24	University of Milano-Bicocca	03
<i>Scienze e tecniche psicologiche</i>	L-24	Vita-Salute San Raffaele University	03
Economia e finanza internazionale	L-33	Bocconi University	03
Economia e Management	L-18	LUISS University	04
<i>Linguaggi dei media</i>	L-20	Catholic University of the Sacred Heart	03
<i>Economia e finanza</i>	L-18	Bocconi University	04
<i>Scienze internazionali e istituzioni europee</i>	L-36	University of Milan	03
<i>Mediatori per l'intercultura e la coesione sociale in Europa</i>	L-39	University for Foreigners "D. Alighieri" of R.C.	09
<i>Business and Economics</i>	L-18	University of Bologna	04
Economics and Business	L-33	LUISS University	03
<i>Scienze e tecniche psicologiche</i>	L-24	University of Bari	03
<i>Relazioni pubbliche e comunicazione d'impresa</i>	L-20	IULM University of Milan	03
<i>Comunicazione e società</i>	L-20	University of Milan	03
<i>Scienze del turismo e comunità locale</i>	L-15	University of Milano-Bicocca	03
<i>Scienze e Tecniche Psicologiche</i>	L-24	European University of Rome	03
<i>Business</i>	L-18	Bocconi University	04
<i>Economia e gestione dei beni culturali e dello spettacolo</i>	L-18	Catholic University of the Sacred Heart	04
Scienze politiche e delle relazioni internazionali	L-36	Tuscia University	03
<i>Scienze e tecniche psicologiche</i>	L-24	Catholic University of the Sacred Heart	03
Scienze Politiche	L-36	LUISS University	03
Less Efficient			
Degree	Type	University	ISCED_F_1dgt
Scienze politiche e delle relazioni internazionali	L-36	University of Messina	03
<i>Consulente del lavoro</i>	L-14	University of Foggia	04
<i>Diritto dell'impresa</i>	L-14	University of Pisa	04
<i>Economia Aziendale</i>	L-18	University of Calabria	04

(continued)

**Table 5** (continued)

Social Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Scienze dell'amministrazione</i>	L-16	University of Palermo	03
<i>Scienze dell'amministrazione e dell'organizzazione</i>	L-16	University of Salerno	03
<i>Scienze del diritto italiano ed europeo</i>	L-14	University of Turin	04
<i>Economia</i>	L-33	University of Catania	03
<i>Formazione di operatori turistici</i>	L-15	University of Catania	03
<i>Scienze del servizio sociale</i>	L-39	University of Messina	09
<i>Consulente del lavoro</i>	L-14	University of Padua	04
<i>Servizi giuridici per Lavoro [...]</i>	L-14	University of Cassino	04
<i>Scienze dei Servizi Giuridici</i>	L-14	University of Cagliari	04
<i>Operatore dei servizi giuridici</i>	L-14	University of Ferrara	04
<i>Scienze dei servizi giuridici</i>	L-14	University of Rome III	04
<i>Scienze dei servizi giuridici</i>	L-14	University of Perugia	04
<i>Scienze dell'amministrazione e dell'organizzazione</i>	L-16	University of Catania	03
<i>Scienze dei servizi giuridici</i>	L-14	University of Bari	04
<i>Servizi giuridici</i>	L-14	Catholic University of the Sacred Heart	04
<i>Sociologia e servizio sociale</i>	L-39	University of Catania	09

Note: "Bold letter" indicates female gender predominance

courses in universities of Milan is only four out of 20. This result confirms that, at the level of three-year degree courses, the disparities between the North and South of Italy are great among the best courses, based on our methodology. With the exception of three courses, all humanities courses are female oriented, in line with expectations. The bottom panel of Table 6 presents the least efficient courses in this area and illustrates a substantially different situation compared to the best ones. Among the 20 least efficient degree courses, we find 18 degree courses in Southern Italy and two in the center-North. These strong territorial imbalances, even in the humanities (thus, degree courses that do not require many resources in terms of infrastructure), are worrisome. Strong disparities exist in terms of educational efficiency between the Northern and Southern universities. The Italian university system seems strongly characterized by polarization. The courses in humanities belong to a female-oriented area (highlighted in bold).

**Table 6** List of the 20 most efficient three-year degree courses in the humanities area (top) and the 20 least efficient three-year degree courses in the humanities area (bottom)

Humanities			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Scienze dell'educazione e della formazione</i>	L-19	University of Macerata	01
<i>Scienze della Mediazione Interlinguistica e Interculturale</i>	L-12	University of Insubria	02
<i>Interpretariato e comunicazione</i>	L-12	IULM University of Milan	02
Lingue e comunicazione per l'impresa e il turismo	L-12	University of Aosta Valley	02
<i>Lettere</i>	L-10	University of Turin	02
<i>Educatore socio culturale</i>	L-19	University of Salento	01
<i>Scienze dell'educazione e della formazione</i>	L-19	University of Siena	01
<i>Lettere</i>	L-10	University of Milan	02
<i>Educatore servizi per infanzia</i>	L-19	University of Bologna	01
<i>Scienze dell'educazione</i>	L-19	University of Milano-Bicocca	01
Lingue e Culture Europee	L-12	University of Modena and Reggio Emilia	02
<i>Mediazione linguistica e culturale</i>	L-12	Ca' Foscari University of Venice	02
<i>Filosofia</i>	L-5	University of Bergamo	02
<i>Lingue e culture per il turismo e il commercio internazionale</i>	L-12	University of Verona	02
<i>Lingue, mercati e culture dell'asia</i>	L-11	University of Bologna	02
Filosofia	L-5	Vita-Salute San Raffaele University	02
<i>Culture e letterature del mondo moderno</i>	L-10	University of Turin	02
<i>Filosofia</i>	L-5	University of Bologna	02
<i>Mediazione linguistica e culturale</i>	L-12	Foreigners University of Siena	02
<i>Lingue e culture moderne</i>	L-11	Suor Orsola Benincasa University of Naples	02
Less Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Scienze dei beni Culturali</i>	L-1	University of Bari	02
<i>Lettere e Beni Culturali</i>	L-1	University of Molise	02
<i>Mediazione linguistica e culturale</i>	L-12	University of L'Aquila	02
<i>Lingue, culture e letterature moderne europee</i>	L-11	University of Naples Federico II	02

(continued)

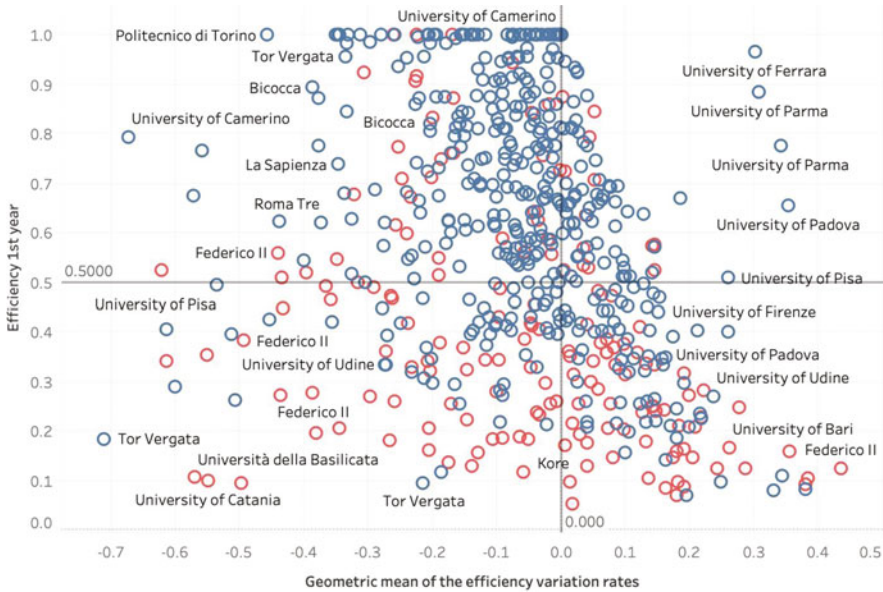
**Table 6** (continued)

Humanities			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Scienze dell'educazione e della formazione</i>	L-19	University of Catania	01
<i>Lettere</i>	L-10	University of Palermo	02
<i>Musicologia</i>	L-1	University of Pavia	02
<i>Scienze dei Beni Culturali</i>	L-1	Tuscia University	02
<i>Filosofia e teoria dei processi comunicativi</i>	L-5	University of L'Aquila	02
<i>Scienze geografiche per l'ambiente e la salute</i>	L-6	Sapienza University of Rome	03
<i>Lettere</i>	L-10	University of Messina	02
<i>Conservazione dei beni culturali</i>	L-1	University of Campania Luigi Vanvitelli	02
<i>Filosofia</i>	L-5	University of Catania	02
<i>Lettere</i>	L-10	University of L'Aquila	02
<i>Comunicazione e DAMS</i>	L-3	University of Calabria	02
<i>Lettere</i>	L-10	University of Catania	02
<i>Beni Culturali: Conoscenza, Gestione, Valorizzazione</i>	L-1	University of Palermo	02
<i>Lettere e Beni Culturali</i>	L-1	University of Foggia	02
<i>Lingue e culture europee euroamericane ed orientali</i>	L-11	University of Catania	02
<i>Beni culturali</i>	L-1	University of Catania	02

Note: "Bold letter" indicates female gender predominance

### 5.3 Trends and Prospects

We conclude the analysis of the three-year degrees with comments on the geometric mean of the efficiency variation rates to better understand the direction of the dynamism of the individual degree courses. It is important to understand which degree courses have improved the most over the period considered. Continuous improvement, evaluation, and self-evaluation are among the key principles of the Italian legislation on the evaluation of the university teaching system. Figure 3 refers to science courses and offers an overview of the relationship between the efficiency in the first year of analysis and the average rate of change during the observation period. The courses belonging to Southern universities are colored in red, while those belonging to the Northern-central universities are colored in blue. The top and bottom panels of the chart report the courses with highest and lowest levels of efficiency at the first year, the left and right panel the ones with lowest and highest levels of efficiency variation. It is possible to note how the study courses that started from lower efficiency levels are characterized by high volatility in the rate of



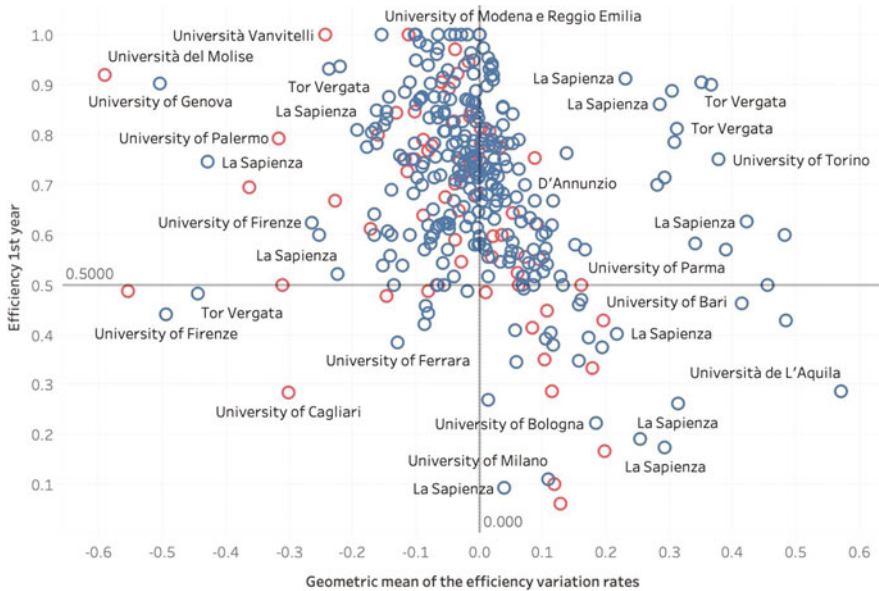
**Fig. 3** Change in efficiency of three-year degree courses in science. Efficiency in the first year of analysis (y-axis) versus the average improvement rate of the degree courses (x-axis). Blue circles are courses belonging to Northern-central universities, while red circles are courses belonging to Southern universities

change of efficiency. Some courses that started at a low level of efficiency registered significant improvement rates, while others displayed a consistent deterioration rate, especially for courses of universities in the South.

Figure 4 shows us the same relationship for degree courses in health sciences. These courses exhibit low variability. This is typical of medical degree courses. Accreditation and regulatory provisions are stringent and limit the autonomy of degree courses, which implies a reduced variability in terms of efficiency. From the same picture, we can conclude that degree courses that have shown a consistent positive rate of growth belong to universities located in the North.

Figure 5 provides useful information on the relationship between the rate of change of efficiency and the level of efficiency in the first year for social sciences courses. It is evident that the most efficient courses are located in the center and North. The courses that started at low efficiency levels are equally distributed among the geographical areas. What is important to underline in this graph is that many courses have significant improvement rates. The best courses during the first year, in contrast, have a low variability and, hence, they keep high performance during all observed periods.

Figure 6 shows the same relationship for the three-year degree courses in humanities. We can appreciate that the courses that register low efficiency in the first year of analysis have positive average rates of variation. Few courses have



**Fig. 4** Change in efficiency of three-year degree courses in health science. Efficiency in the first year of analysis (y-axis) versus the average improvement rate of the degree courses (x-axis). Blue circles are courses belonging to Northern-central universities, while red circles are courses belonging to Southern universities

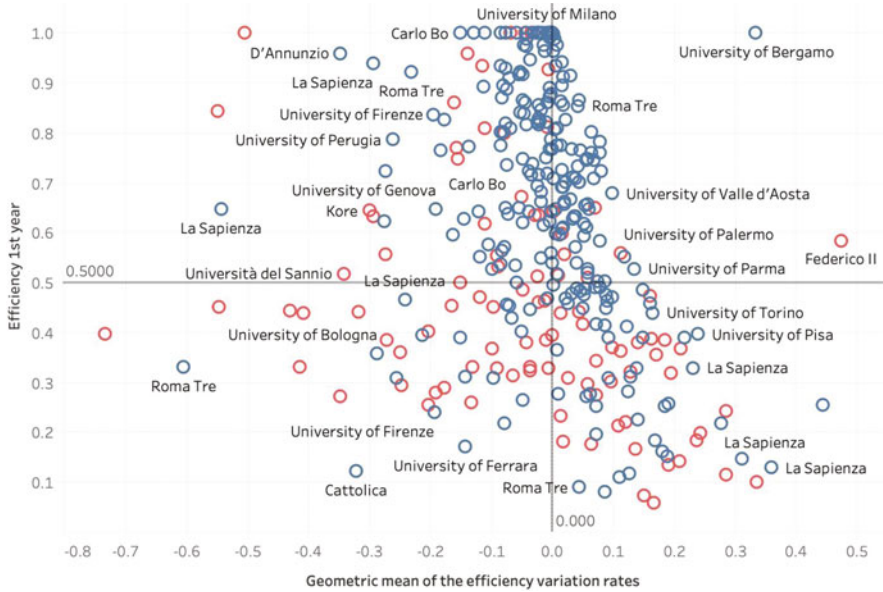
worsened their low position and find themselves in the lower-left quadrant of the chart. Yet regarding degree courses with a medium-high level of efficiency, the tendency toward a moderate deterioration is generally apparent (top-left panel of the graph).

## 6 Results on Master's Degrees

### 6.1 *Effect of Size on Efficiency*

The analysis of the efficiency of the master's degrees follows the same structure of the analysis carried out on the three-year degree courses presented in the previous section.

Figure 7 in supplementary material shows the impact of size on the efficient frontier of master's degree courses in science. In universities with small numbers, in terms of enrolled students in sciences, there is no effect of the number of enrolled students on the efficiency of the best master's courses. Therefore, the number of enrolled students does not affect the outputs (percentages of graduates within the legal duration of the course and satisfied by the course of study), given the input



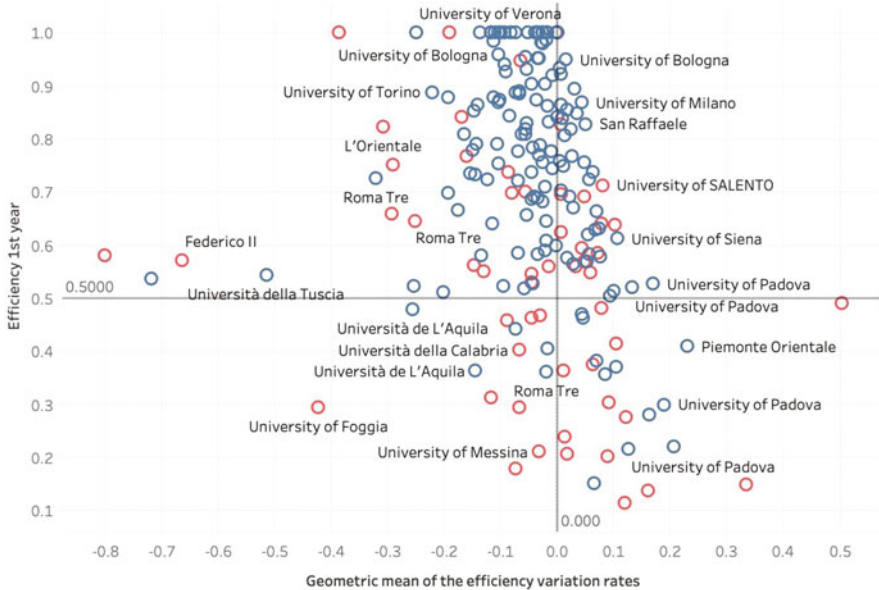
**Fig. 5** Change in efficiency of three-year degree courses in social science. Efficiency in the first year of analysis (y-axis) versus the average improvement rate of the degree courses (x-axis). Blue circles are courses belonging to Northern-central universities, while red circles are courses belonging to Southern universities

(the number of teachers per students weighted by the number of teaching hours). There is a negative effect when moving from universities with small-sized science courses to those with medium-sized ones, and then a positive effect of the size in the transition from universities with medium-sized science courses to those with large-sized ones.

The impact of size on the efficiency of the health science master's courses is illustrated in Fig. 8 in supplementary material. There is no influence of the number of students enrolled in degree courses in the health sciences on the most efficient courses of study. The same conclusion can be drawn for the social sciences (see Fig. 9). It is useful to emphasize how the distribution of size in this discipline is quite homogeneous.

Figure 10 in supplementary material shows the impact of size on the efficiency of master's courses in the humanities. The ratios show a negative influence of the size on the efficiency of degree courses in the humanities. As the size of the disciplinary area increases, the effect becomes ever greater and remains constantly negative, indicating that master's courses in humanities are more efficient when the number of students in this area is small.

The relationship between conditional and unconditional efficiency, calculated with respect to the partial frontier with  $\alpha = 0.5$ , shows the effect of the conditional factor (in our case size) on the distribution of efficiency and therefore on the



**Fig. 6** Change in efficiency of three-year degree courses in humanities. Efficiency in the first year of analysis (y-axis) versus the average improvement rate of the degree courses (x-axis). Blue circles are courses belonging to Northern-central universities, while red circles are courses belonging to Southern universities

convergence process of the least efficient courses compared to those at the efficiency frontier (the best-performing ones).

Figure 11 in supplementary material shows the impact of size on the efficiency distribution of master’s courses in science. The effect of students enrolled in this area on the master’s degrees courses efficiency distribution is negative in universities with small-sized science departments (0–25,000) and positive in universities with medium-large science departments (30,000–45,000). Hence, the increasing number of enrolled students in this area improves the efficiency of master’s courses. Furthermore, it is worthwhile to note that the size has a positive effect on efficiency only starting from some values. Therefore, it seems there is a threshold effect, and this area needs to reach a certain number of enrolled students to have a positive effect on efficiency from an increasing number of enrolled students.

In health and social science (not reported for space reasons), size does not seem to affect the process of convergence of less efficient master’s courses toward the efficient frontier, so size does not improve efficiency for health and social science master’s degrees. In the case of master’s degrees in the humanities, the effect of size on the distribution of efficiency is quite heterogeneous (see Fig. 12 in supplementary material). Fig. 12, however, shows that the size of the humanities (number of enrolled students) certainly has a negative effect on universities with a small area in this field. Concerning universities with medium-sized humanities



areas, the effect seems to be positive, which indicates that growth of enrolled students favors improvement of the efficiency of master's degree courses. This trend is quite similar to science master's degrees.

## 6.2 *Analysis of the Efficiency Scores*

Now we move on to assess the efficiency of master's degree courses, focusing on the 20 that performed best and worst on average during the five years considered in the study.

In the ranking of the 20 most efficient master's degree courses in the sciences, shown in the top panel of Table 7, we mainly find courses from the center-North (16 out of 20). The majority of these courses (13 out of 20) belong to female-oriented disciplines and are highlighted in bold in the table.

The bottom panel of Table 7, concerning the less efficient master's degree courses, illustrates a substantially different situation compared to the more efficient master's degree courses. It is interesting to note that this is the first case in which all three of the worst courses in terms of efficiency belong to universities in the center-North. Among the 20 least efficient degree courses, on average, in the five years considered, we find five degree courses from Southern Italy and 15 from central-Northern Italy. Overall, master's degree courses in the sciences present a more balanced situation in the distribution of inefficient courses between the North and South of the country.

All of the courses at the top of the ranking of the least efficient master's degree courses in the sciences are male oriented. The situation is similar to the three-year degrees in the sciences. This finding corroborates our hypothesis that the more developed industrial context in central-Northern Italy plays a central role in improving the efficiency in sciences master's courses. Moreover, the two outputs we have selected—regularity in studies and satisfaction in the courses—are aspects related to the labor market.

In the ranking of the 20 most efficient degree courses in the health sciences, shown at the top of Table 8, we mainly find courses in the center-North (16/20), with a high frequency of degree courses in universities in the Lombardy region (5/20). The ranking shows that a large part (16/20) of health science degrees is female prevalent (highlighted in bold in the table). The bottom part of Table 8, regarding the less efficient master's courses in health science, illustrates a substantially similar situation compared to the more efficient courses in the discipline. Among the 20 least efficient degree courses, in the period studied, we find eight degree courses in universities in Southern Italy and 12 in central-Northern Italy. All these courses, except for “Medicina e Chirurgia” at the University of Pavia, belong to the female-dominated field.

In the ranking of the 20 most efficient degree courses in the social sciences, presented in the top panel of Table 9, we find courses all in universities in the center-North (except one that belongs to Federico II University), with a very high frequency

**Table 7** List of the 20 most efficient three-year degree courses in the science area (top) and the 20 least efficient three-year degree courses in the science area (bottom)

Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Scienze Biologiche</i>	LM-6	University of Naples Federico II	05
<i>Scienze dell'educazione motoria e delle attività adattate</i>	LM-67	University of Turin	09
<i>Scienze dell'Alimentazione e della Nutrizione Umana</i>	LM-61	Biomedical University of Rome	05
<i>Molecular biology and genetics</i>	LM-6	University of Pavia	05
<i>Scienze e tecniche dello sport (...)</i>	LM-67	University of Perugia	09
<i>Scienze e Tecniche dello Sport</i>	LM-68	University of Bari	09
<i>Biotecnologie Mediche</i>	LM-9	University of Turin	05
<i>Stochastics and data science</i>	LM-40	University of Turin	05
Controllo e sicurezza degli alimenti	LM-70	University of Modena and Reggio Emilia	08
Management delle Attività Motorie e Sportive	LM-47	University of Bologna	04
Ingegneria civile	LM-23	Kore University of Enna	07
Ingegneria Gestionale	LM-31	Sapienza University of Rome	07
Management of Built Environment	LM-24	Polytechnic University of Milan	07
<i>Scienze e Tecniche dello Sport</i>	LM-68	University of Pavia	09
Automation and Control Engineering	LM-25	Polytechnic University of Milan	07
<i>Biotechnologia e Biomedicina</i>	LM-9	Vita-Salute San Raffaele University	05
<i>Scienze e tecnologie alimentari</i>	LM-70	University of Parma	08
<i>Scienze della Nutrizione Umana</i>	LM-61	University of Naples Federico II	05
<i>Geologia degli idrocarburi</i>	LM-74	University of Perugia	05
Natural resources management for tropical rural development	LM-69	University of Firenze	08
Less Efficient			
Degree	Type	University	ISCED_F_1dgt
Ingegneria Elettronica	LM-29	Sapienza University of Rome	07
Ingegneria elettrica	LM-28	University of Naples Federico II	07
Ingegneria Chimica	LM-22	University of Calabria	07
Ingegneria civile	LM-23	University of Padua	07
Architettura	LM-4 C.U.	University of Catania	07

(continued)

**Table 7** (continued)

Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
Ingegneria Robotica e dell'Automazione	LM-25	University of Pisa	07
Ingegneria delle infrastrutture civili e dell'ambiente	LM-23	University of Pisa	07
Ingegneria edile	LM-4 C.U.	University of L'Aquila	07
Ingegneria per l'Ambiente e il Territorio	LM-35	University of Calabria	07
Ingegneria Civile Strutturale e Geotecnica	LM-23	University of Catania	07
Ingegneria edile	LM-4 C.U.	University of Pisa	07
Ingegneria edile	LM-4 C.U.	Marche Polytechnic University	07
Architettura (Progettazione Architettonica)	LM-4	University of Naples Federico II	07
Ingegneria edile	LM-4 C.U.	University of Bologna	07
Ingegneria Civile	LM-23	Sapienza University of Rome	07
Ingegneria per l'Ambiente e il Territorio	LM-35	University of Salerno	07
Ingegneria edile	LM-4 C.U.	University of Trento	07
Ingegneria edile	LM-4 C.U.	University of Padua	07
Ingegneria Biomedica	LM-21	University of Pisa	07
Ingegneria edile	LM-4 C.U.	Sapienza University of Rome	07

*Note:* "Bold letter" indicates female gender predominance

of degree courses in universities in Milan (10/20). The majority of these courses (13/20) belong to the female-dominated field and are highlighted in bold in the table. The bottom part of Table 9, regarding the less efficient courses in social sciences, illustrates a substantially different situation compared to the most efficient ones. It is interesting to note that most of these courses are in law disciplines and belong to universities in Southern Italy. Among the 20 least efficient degree courses, on average, for the period under analysis, we find 11 degree courses in Southern Italy and nine in the center-North. It is, therefore, a vastly different situation compared to the ranking of the best courses.

In the ranking of the 20 most efficient degree courses in humanities, exhibited at the top of Table 10, we find nearly all courses in the center-North (18/20) with a high frequency—although a lower one than the three-year degrees—of university degree courses in Milan (6/20). Except for the cognitive science program at the University of Trento, all the best humanities master's degrees belong to female-dominated fields and are highlighted in bold in the table. The bottom part of Table

**Table 8** List of the 20 most efficient master's degree courses in the health science area (top) and the 20 least efficient master's degree courses in the health science area (bottom)

Health Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Chimica e tecnologia farmaceutiche</i>	LM-13	University of Salerno	09
<i>Odontoiatria e protesi dentaria</i>	LM-46	University of Modena and Reggio Emilia	09
Medicina veterinaria	LM-42	University of Parma	08
<i>Odontoiatria e protesi dentaria</i>	LM-46	University of Perugia	09
<i>Scienze delle professioni sanitarie tecniche diagnostiche</i>	LM/SNT3	University of Turin	09
Farmacia	LM-13	University of Camerino	09
<i>Scienze riabilitative delle professioni sanitarie</i>	LM/SNT2	University of Turin	09
<i>Odontoiatria e protesi dentaria</i>	LM-46	University of Campania Luigi Vanvitelli	09
<i>Odontoiatria e protesi dentaria</i>	LM-46	University of Pavia	09
<i>Odontoiatria e protesi dentaria</i>	LM-46	Catholic University of the Sacred Heart	09
<i>Scienze delle professioni sanitarie tecniche assistenziali</i>	LM/SNT3	University of L'Aquila	09
<i>Chimica e tecnologia farmaceutiche</i>	LM-13	University of Perugia	09
<i>Farmacia</i>	LM-13	University of Urbino "Carlo Bo"	09
<i>Scienze infermieristiche e ostetriche</i>	LM/SNT1	University of Milano-Bicocca	09
<i>Scienze riabilitative delle professioni sanitarie</i>	LM/SNT2	University of Padua	09
<i>Odontoiatria e protesi dentaria</i>	LM-46	University of Genoa	09
<i>Scienze delle professioni sanitarie tecniche diagnostiche</i>	LM/SNT3	University of Ferrara	09
Medicina veterinaria	LM-42	University of Milan	08
<i>Odontoiatria e protesi dentaria</i>	LM-46	University of Milano-Bicocca	09
Medicina e chirurgia	LM-41	Magna Græcia University of Catanzaro	09
Less Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Scienze infermieristiche e ostetriche</i>	LM/SNT1	University of Campania Luigi Vanvitelli	09
<i>Chimica e tecnologia farmaceutiche</i>	LM-13	University of Ferrara	09
<i>Medicina e chirurgia</i>	LM-41	University of Verona	09
Medicina e chirurgia	LM-41	University of Pavia	09

(continued)

**Table 8** (continued)

Health Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<b>Farmacia</b>	LM-13	University of Bologna	09
<b>Farmacia</b>	LM-13	University of Bari	09
<b>Scienze riabilitative delle professioni sanitarie</b>	LM/SNT2	University of Genoa	09
<b>Scienze infermieristiche e ostetriche</b>	LM/SNT1	University of Pavia	09
<b>Scienze delle professioni sanitarie della prevenzione</b>	LM/SNT4	University of Milan	09
<b>Medicina e chirurgia</b>	LM-41	Vita-Salute San Raffaele University	09
<b>Scienze infermieristiche e ostetriche</b>	LM/SNT1	University of Palermo	09
<b>Scienze delle professioni sanitarie della prevenzione</b>	LM/SNT4	University of Turin	09
<b>Farmacia</b>	LM-13	University of Calabria	09
<b>Medicina e chirurgia</b>	LM-41	University of Perugia	09
<b>Medicina e chirurgia 'F'</b>	LM-41	Sapienza University of Rome	09
<b>Farmacia</b>	LM-13	University of Cagliari	09
<b>Medicina e chirurgia</b>	LM-41	University of Bologna	09
<b>Scienze delle professioni sanitarie tecniche diagnostiche</b>	LM/SNT3	University of Naples Federico II	09
<b>Scienze riabilitative delle professioni sanitarie</b>	LM/SNT2	University of L'Aquila	09
<b>Odontoiatria e protesi dentaria</b>	LM-46	University of L'Aquila	09

Note: "Bold letter" indicates female gender predominance

10 reports degree courses rated as less efficient. Among the twenty less efficient degree courses, we find 13 degree courses in Southern Italy and seven in the center-North; therefore, the situation is substantially reversed compared to the ranking of the best master's courses in humanities.

### 6.3 Trends and Prospects

This section is devoted to the analysis of changes in efficiency of master's degrees in order to better understand the direction of the dynamism of the degree courses. University legislation is geared toward improving quality. Therefore, it is useful to understand which master's course has improved the most or which one has deteriorated the most, starting from the initial level of efficiency of the degree

**Table 9** List of the 20 most efficient master's degree courses in the social science area (top) and the 20 least efficient master's degree courses in the social science area (bottom)

Social Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Environmental and Food Economics</i>	LM-76	University of Milan	04
Marketing	LM-77	LUISS University	04
<i>Studi e ricerche interdisciplinari sull'Europa Orientale</i>	LM-52	University of Bologna	03
<i>Food marketing e strategie commerciali</i>	LM-77	Catholic University of the Sacred Heart	04
<i>Management</i>	LM-77	Catholic University of the Sacred Heart	04
<i>Psicologia</i>	LM-51	Vita-Salute San Raffaele University	03
Management Internazionale*	LM-77	Bocconi University	04
<i>Psicologia clinica</i>	LM-51	University of Bologna	03
<i>Psicologia per il benessere</i>	LM-51	Catholic University of the Sacred Heart	03
Marketing Management*	LM-77	Bocconi University	04
<i>Psicologia del lavoro, delle organizzazioni e delle risorse umane</i>	LM-51	University of Bologna	03
<i>Economia e Finanza</i>	LM-56	University of Naples Federico II	03
Business Administration—Gestione d'impresa	LM-77	University of Rome II "Tor Vergata"	04
<i>Resource economics and sustainable development</i>	LM-56	University of Bologna	03
Economia e Management nelle Arti, Cultura, Media e Intrattenimento*	LM-77	Bocconi University	04
<i>Valutazione del funzionamento individuale in psicologia clinica e della salute</i>	LM-51	University of Perugia	03
<i>Management</i>	LM-77	LUISS University	04
Amministrazione, finanza aziendale e controllo*	LM-77	Bocconi University	04
<i>Amministrazione, finanza e controllo</i>	LM-77	LUISS University	04
Economia e Management dell'Innovazione e della Tecnologia	LM-56	Bocconi University	03

(continued)

**Table 9** (continued)

Social Science			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
Less Efficient			
Degree	Type	University	ISCED_F_1dgt
Giurisprudenza	LMG/01	Parthenope University of Naples	04
<i>Statistica ed Informatica per le Decisioni e le Analisi di Mercato</i>	LM-82	University of Calabria	05
<i>Comunicazione della cultura e dello spettacolo</i>	LM-65	University of Catania	02
Economia e Diritto*	LM-77	University of Padua	04
<i>Management della Pubblica Amministrazione</i>	LM-63	University of Catania	03
<i>Innovazione e servizio sociale</i>	LM-87	University of Padua	09
<i>Scienze Economiche</i>	LM-56	University of Rome III	03
<i>Scienze dello spettacolo e produzione multimediale</i>	LM-65	University of Padua	02
Finanza e assicurazioni	LM-16	Sapienza University of Rome	04
<i>Giurisprudenza</i>	LMG/01	University of Ferrara	04
<i>Scienze Pedagogiche e Progettazione Educativa</i>	LM-85	University of Catania	01
<i>Giurisprudenza</i>	LMG/01	University of Teramo	04
<i>Giurisprudenza</i>	LMG/01	University of Salerno	04
Giurisprudenza	LMG/01	University of Sannio	04
<i>Scienze statistiche</i>	LM-82	Sapienza University of Rome	05
Scienze Statistiche	LM-82	University of Padua	05
Giurisprudenza*	LMG/01	University of Molise	04
<i>Giurisprudenza</i>	LMG/01	University of Messina	04
<i>Giurisprudenza</i>	LMG/01	University of Catania	04
Giurisprudenza*	LMG/01	University of Padua	04

Note: "Bold letter" indicates female gender predominance and "\*" gender neutral

course. This will help in understanding the dynamism of the master's courses of the Italian university system.

As done in the previous section, we present the results by relating the level of efficiency in the first year of analysis and the average improvement rate of each master's degree course. In the following graphs, the units in red are the master's degree courses offered in universities located in Southern Italy, and the circles in blue are the master's degree courses offered in universities in central-Northern Italy.

The results for the science master's degree programs are shown in Fig. 13 in supplementary material. For most of them, the rate of change is low. Master's

**Table 10** List of the 20 most efficient master's degree courses in the humanities area (top) and the 20 least efficient master's degree courses in the humanities area (bottom)

Humanities			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Traduzione specializzata</i>	LM-94	University of Bologna	02
<i>Gestione di contenuti digitali per i media, le imprese e i patrimoni culturali</i>	LM-19	Catholic University of the Sacred Heart	03
<i>Scienze della formazione primaria</i>	LM-85	University of L'Aquila	01
<i>Editoria e giornalismo</i>	LM-19	University of Verona	03
Cognitive Science Scienze Cognitive	LM-55	University of Trento	03
<i>Scienze della formazione primaria</i>	LM-85	University of Macerata	01
<i>Lingua e cultura italiane per stranieri</i>	LM-39	University of Bologna	02
<i>Filologia, letterature e storia dell'antichità</i>	LM-15	University of Turin	02
<i>Diagnostica dei Beni Culturali</i>	LM-11	University of Salento	05
<i>Comunicazione per l'impresa, i media e le organizzazioni complesse</i>	LM-19	Catholic University of the Sacred Heart	03
<i>Italianistica, culture letterarie europee, scienze linguistiche</i>	LM-14	University of Bologna	02
<i>Geografia e scienze territoriali</i>	LM-80	University of Turin	03
<i>Interpretariato e traduzione editoriale, settoriale</i>	LM-94	Ca' Foscari University of Venice	02
<i>Scienze linguistiche</i>	LM-38	Catholic University of the Sacred Heart	02
<i>Scienze della formazione primaria</i>	LM-85	University of Bologna	01
<i>Scienze della formazione primaria</i>	LM-85	University of Udine	01
<i>Lingue e culture per la comunicazione e la cooperazione internazionale</i>	LM-38	University of Milan	02
<i>Lingue, Culture, Turismo</i>	LM-37	University of Eastern Piedmont	02
<i>Lingue, letterature e culture straniere</i>	LM-37	Catholic University of the Sacred Heart	02
<i>Traduzione specialistica e interpretariato di conferenza</i>	LM-94	IULM University of Milan	02

(continued)



**Table 10** (continued)

Humanities			
Most Efficient			
Degree	Type	University	ISCED_F_1dgt
Less Efficient			
Degree	Type	University	ISCED_F_1dgt
<i>Scienze filosofiche</i>	LM-78	University of Catania	02
<i>Studi letterari e culturali</i>	LM-14	University of L'Aquila	02
<i>Storia dell'arte</i>	LM-89	University of Firenze	02
<i>Storia dell'arte</i>	LM-89	University of Turin	02
<i>Conservazione e restauro dei beni culturali</i>	LMR/02	University of Palermo	05
<i>Archeologia</i>	LM-2	University of Palermo	02
<i>Beni archeologici e storico-artistici</i>	LM-2	D'Annunzio University of Chieti-Pescara	02
<i>Filologia moderna</i>	LM-14	University of Catania	02
<i>Scienze archeologiche</i>	LM-2	University of Padua	02
<i>Archeologia</i>	LM-2	University of Pisa	02
<i>Archeologia e Storia dell'arte</i>	LM-2	University of Naples Federico II	02
<i>Storia e cultura dei paesi mediterranei</i>	LM-84	University of Catania	02
<i>Filologia classica</i>	LM-15	University of Catania	02
<i>Archeologia e storia dell'arte</i>	LM-2	University of Campania Luigi Vanvitelli	02
<i>Lingue per la cooperazione internazionale</i>	LM-38	University of Catania	02
<i>Archeologia</i>	LM-2	University of Siena	02
<i>Antropologia e linguaggi dell'immagine</i>	LM-1	University of Siena	03
<i>Lingue e culture europee ed extraeuropee</i>	LM-37	University of Catania	02
<i>Storia dell'arte e beni culturali</i>	LM-89	University of Catania	02
<i>Lingue e Civiltà Orientali</i>	LM-36	Sapienza University of Rome	02

Note: "Bold letter" indicates female gender predominance

degrees that have a low to medium level of efficiency in the first year and improve significantly can be seen in the lower-right quadrant of the figure. Regarding the courses in the health sciences, as presented in Fig. 14 in supplementary material, we note a greater variability compared to the three-year degrees in the same area. Furthermore, it is evident from the graph that master's courses that started from a situation of high efficiency show positive variation rates (upper-right quadrant of the figure). Therefore, these courses further increase the efficiency. The courses that started from a low level of efficiency further worsened their performance

(lower-right quadrant of the figure), except few of them, and especially the courses in Federico II University, which start from a very low level of efficiency but improve their performance and register the highest rate of positive efficiency change. As for the social sciences, displayed in Fig. 15 in supplementary material as a pyramid structure, the courses that started with particularly good efficiency have undergone little variation. The degree courses that started from lower efficiency, in contrast, have registered a high level of variability in the average rate of change. As for the humanities master's degree courses, as revealed in Fig. 16 in supplementary material, the courses that already started from a good position have mostly improved. Unfortunately, many post-graduate courses that had lower levels of efficiency have deteriorated, especially in the North (blue circles in the lower-left quadrant in the figure).

## 7 Conclusion

This chapter contributes to the literature on the evaluation of universities by proposing a new method of evaluating the ranking and performance of Italian universities from a perspective of the efficiency of the teaching offered at the level of study course, overcoming simple monodimensional indicators used in this context.

For the first time, an evaluation of the efficiency of study courses is proposed and assessed by considering as outputs the number of graduates within the legal duration of the courses and their satisfaction with the course of study followed, and as input the number of teachers per student weighted by the number of teaching hours. The efficiency is calculated taking in account the effect of time and size (number of students enrolled in the disciplinary fields). Furthermore, for the first time for this type of analysis, the results are analyzed in a gender-balance perspective, identifying to which gender-oriented area the best and worst courses belong.

The results highlighted a greater efficiency of the university courses of the universities of the center-North. In addition, the analysis has highlighted how the universities in the islands, especially in Sicily, suffer from serious problems of inefficiency for their courses of study, with a tendency to worsen, as evidenced by the negative temporal variation rates of efficiency.

High efficiency values are obtained in particular from the scientific and health science areas, which involve disciplines with a high technological content strictly connected with the industrial sector. The values of the efficiency of the courses of study in these areas seem to reflect the economic reality of the territories in which they operate. In particular, high values of efficiency in the scientific and health science area study courses in the center-North stand out, where there is an industrial sector capable of absorbing the human capital formed by universities, enhancing their skills, and interacting in the formation of the same through partnerships and projects with the university world aimed mainly at the professional integration of specific figures required by the world of work. This happens mainly for scientific areas, such as engineering, which in Northern Italy has employment rates that

exceed 80%. With regard to gender policies, in this STEM area, an interesting aspect is highlighted: the number of the best courses that are female oriented is growing, while the worst courses are highly male oriented. As far as the health science sector is concerned, the gap between North and South is even wider in terms of degree course efficiency. The management of health care on a regional basis and greater efficiency in the allocation of economic resources in these territories indicates that the universities, which work alongside the hospitals of excellence, widely present in the center-North, also incentivize efficiency, as well as the satisfaction of recent graduates, as seen from the results obtained on the study courses of this disciplinary area.

The humanities area, on the other hand, presents a trend reversal in relation to the efficiency of its study courses, which are also efficient in the South—certainly an encouraging result for this disciplinary area.

A possible extension of this work could consist in the inclusion of variables of the socioeconomic context of the territories in which Italian universities operate—territories with different economic resources and employment scenarios, which follow their own dynamics, often divergent from those of the efficiency of university centers. One possibility would be to include postgraduate employment data. The only source in this regard is the AlmaLaurea database, which, however, does not provide information on the type of employment contract.

Another interesting extension of this work could be the application of recently developed efficiency methodologies (see Daraio et al., 2021) able to estimate quality as a latent heterogeneity factor in the efficiency of the university faculty courses providing quality-adjusted rankings of Italian university faculty courses. All these extensions are left for future research.

The main contribution of this work is to have shown how, by using university faculty courses as a relevant unit of analysis, it is possible to apply the new conditional efficiency analysis methodologies to provide multidimensional comparisons and rankings by discipline at the national level. Obviously, the proposed analyses suffer from limitations due to existing data, highlighting the need to invest in the creation of more complete databases based on information coming from different sources.

**Acknowledgements** We thank ANVUR for the data used in this chapter and in the projects EDUCO, “III Concorso idee di Ricerca” and Giampiero D’Alessandro who helped with the data selection and participants at the ANVUR Meeting (Rome, November 2019) who gave feedback to various versions of this chapter. These individuals and organizations are not responsible for the views expressed here.

## Appendix

Let the production process of teaching activities in the university faculty courses be characterized by a vector of *inputs*  $X \in \mathbb{R}_+^p$  that produces a vector of *output*  $Y \in \mathbb{R}_+^q$

and a vector of environmental variables  $Z \in \mathbb{R}_+^d$  may affect the performance of this production process. For each time period  $t$ , the attainable production set  $\Psi_t^z \subset \mathbb{R}_+^{p+q}$  can be defined as the support of the conditional probability of being dominated (Cazals et al., 2002; Daraio & Simar, 2005; Mastromarco & Simar, 2015), given by

$$H_{X,Y|Z}^t(x, y|z) = \text{Prob}(X \leq x, Y \geq y|Z = z, T = t). \tag{A1}$$

As Mastromarco and Simar (2015) suggest, for each period  $t$ , the *conditional output-oriented efficiency* of a production plan  $(x, y)$  facing conditions  $z$ , is defined as

$$\lambda_t(x, y|z) = \sup \{ \lambda | (x, \lambda y) \in \Psi_t^z \} = \sup \{ \lambda | S_{Y|X,Z}^t(\lambda y|x, z) > 0 \} \tag{A2}$$

where  $S_{Y|X,Z}^t(y|x, z) = \text{Prob}(Y \geq y | X \leq x, Z = z, T = t)$  is a nonstandard conditional survival function.

The *unconditional output-oriented efficiency* of the production plan  $(x, y)$  is given by

$$\lambda(x, y) = \sup \{ \lambda | (x, \lambda y) \in \Psi \} = \sup \{ \lambda | S_{Y|X}(\lambda y|x) > 0 \} \tag{A3}$$

where  $S_{Y|X}(y|x) = \text{Prob}(Y \geq y | X \leq x)$  is a nonstandard survival function conditioned only to the inputs but non-conditioned to time and  $Z$ .

The estimation of conditional distributions  $S_{Y|X,Z}^t(y|x, z)$  where we condition on  $X \leq x$  and a particular value of  $Z = z$  and  $T = t$ , is given by

$$\hat{S}_{Y|X,Z}^t(y|x, z) = \frac{\sum_{j=(i,v)} I(x_j \leq x, y_j \geq y) K_{h_z}(z_j - z) K_{h_t}(v - t)}{\sum_{j=(i,v)} I(x_j \leq x) K_{h_z}(z_j - z) K_{h_t}(v - t)} \tag{A4}$$

where  $K_{(\cdot)}$  are kernels with compact support and  $h_{(\cdot)}$  are the bandwidths or smoothing parameters (see Bădin et al., 2010 and 2019, for technical details).<sup>1</sup> Optimal bandwidths are selected by least squares cross-validation (LSCV), which is asymptotically equivalent to the maximum likelihood (see, for example, Li & Racine, 2007). Daraio and Simar (2005, 2007a) and Bădin et al. (2010) discuss in detail how to choose the appropriate bandwidths. They are determined by the estimation of conditional distributions  $S_{Y|X,Z}^t(y|x, z)$  on  $X \leq x$  and a particular value of  $Z = z$  and  $T = t$ , following the approach suggested by Hall et al. (2004) and Li and Racine (2007).

For our analysis, we follow Daouia and Simar (2007) and Mastromarco and Simar (2015) and apply order- $\alpha$  partial frontiers, to provide efficiency scores more robust to outliers and extreme observations. *Unconditional* and *conditional*

---

<sup>1</sup> Only the variable  $Z$  requires smoothing and appropriate bandwidths.

output-oriented robust (or partial) efficiency scores are defined for any  $\alpha \in (0, 1)$ , respectively, as follows:

$$\lambda_{\alpha}(x, y) = \sup \{ \lambda | S_{Y|X}(\lambda y | x) > 1 - \alpha \} \quad (\text{A5})$$

$$\lambda_{t,\alpha}(x, y | z) = \sup \{ \lambda | S'_{Y|XZ}(\lambda y | x, z) > 1 - \alpha \} \quad (\text{A6})$$

The partial frontiers do not depend on full support of  $Y$  under the conditioning, but on a less extreme quantile (unless  $\alpha = 1$ ) and for this reason are also called robust frontiers. The partial frontiers estimated with values of  $\alpha$  close to one provide the same information as the full frontier estimates, but do not envelop all the data points and for this reason are more robust to extremes and outliers. By choosing a central quantile, such as the median (i.e.,  $\alpha = 0.5$ ), it is possible to investigate the effect of  $Z$  on the distribution of inefficiencies. On the contrary, values of  $\alpha$  close to one (i.e.,  $\alpha = 0.99$ ) allow us to analyze the effect of  $Z$  on the efficient frontier.

The ratios to be analyzed are

$$R_O(x, y | z, t) = \frac{\lambda_t(x, y | z, t)}{\lambda(x, y)} \quad (\text{A7})$$

whose numerator has been defined in Eq. (A2) and the denominator has been defined in Eq. (A3). To be more robust to extremes and outliers, in this study we applied the robust ratios calculated using the partial robust frontiers of order  $-\alpha$  given by

$$R_{O,\alpha}(x, y | z, t) = \frac{\lambda_{t,\alpha}(x, y | z)}{\lambda_{\alpha}(x, y)} \quad (\text{A8})$$

whose numerator and denominator have been defined, respectively, in Eqs. (A6 and A5).

We apply these efficiency ratios to explore the influence of the external variables  $Z$  on the efficient frontier (using  $\alpha = 0.99$ ) and on the distribution of the efficiency (using  $\alpha = 0.50$ ).

To detect the impact of the external (environmental) variables  $Z$  on the efficient frontier, Daraio and Simar (2005, 2007a, 2007b) and Bădin et al. (2012) propose to plot the ratios of the conditional to unconditional efficiency scores as a function of the  $Z$  variable. In our *output-oriented* framework, we consider given the input and look at the maximum feasible expansion of the outputs. In this framework, an *increasing trend* of the ratios denotes a *positive* impact of  $Z$  on the efficient frontier. On the contrary, a *decreasing trend* of the ratios points to a *negative* impact of the  $Z$  variable on the efficient frontier. A *flat trend* of the ratios identifies *no impact* of the  $Z$  on the efficient frontier. In our case, the ratios  $R_{O,\alpha}(x, y | z, t)$  are calculated using  $\alpha = 0.99$  to ensure a robust estimation of the full efficient frontier.

To investigate the impact of  $Z$  on the distribution of efficiency scores, it is necessary to inspect the plot of the ratios  $R_O(x, y|z, t)$  calculated using a frontier that captures the center of the distribution, specifying a  $\alpha = 0.50$ . Again, an increasing (decreasing) trend of the ratios identifies a positive impact (negative) of the  $Z$  on the distribution of the efficiency scores. A flat trend shows no effect of  $Z$  on the distribution of efficiency scores.

## References

- Agasisti, T., & Dal Bianco, A. (2009). Reforming the university sector: Effects on teaching efficiency—Evidence from Italy. *Higher Education*, 57(4), 477.
- Agasisti, T., & Johnes, G. (2010). Heterogeneity and the evaluation of efficiency: The case of Italian universities. *Applied Economics*, 42(11), 1365–1375.
- Bădin, L., Daraio, C., & Simar, L. (2010). Optimal bandwidth selection for conditional efficiency measures: A data-driven approach. *European Journal of Operational Research*, 201(2), 633–640.
- Bădin, L., Daraio, C., & Simar, L. (2012). How to measure the impact of environmental factors in a nonparametric production model. *European Journal of Operational Research*, 223(3), 818–833.
- Bădin, L., Daraio, C., & Simar, L. (2019). A bootstrap approach for bandwidth selection in estimating conditional efficiency measures. *European Journal of Operational Research*, 277, 784–797.
- Bonaccorsi, A., & Daraio, C. (2007a). Theoretical perspectives on university strategy. In A. Bonaccorsi & C. Daraio (Eds.), *a cura di* (pp. 3–30) EID: 2-s2.0-68049100773.
- Bonaccorsi, A., & Daraio, C. (2007b). Universities as strategic knowledge creators. Some preliminary evidence. In A. Bonaccorsi & C. Daraio (Eds.), *a cura di* (pp. 31–84) EID: 2-s2.0-84881859632.
- Bonaccorsi, A., & Daraio, C. (Eds.). (2007c). *Universities and strategic knowledge creation. Specialization and Performance in Europe*. Edward Elgar Publisher.
- Card, D., & Payne, A. A. (2021). High school choices and the gender gap in STEM. *Economic Inquiry*, 59(1), 9–28.
- Cazals, C., Florens, J. P., & Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of Econometrics*, 106(1), 1–25.
- Cooper, S. T., & Cohn, E. (1997). Estimation of a frontier production function for the South Carolina educational process. *Economics of Education Review*, 16(3), 313–327.
- Daouia, A., & Simar, L. (2007). Nonparametric efficiency analysis: A multivariate conditional quantile approach. *Journal of Econometrics*, 140(2), 375–400.
- Daraio, C. (2018). Nonparametric methods and higher education. In P. Teixeira & J. Shin (Eds.), *Encyclopedia of international higher education systems and institutions*. Springer. [https://doi.org/10.1007/978-94-017-9553-1\\_82-1](https://doi.org/10.1007/978-94-017-9553-1_82-1)
- Daraio, C., & Bonaccorsi, A. (2017). Beyond university rankings? Generating new indicators on universities by linking data in open platforms. *Journal of the Association for Information Science and Technology*, 68(2), 508–529.
- Daraio, C., Bonaccorsi, A., & Simar, L. (2015a). Rankings and university performance: A conditional multidimensional approach. *European Journal of Operational Research*, 244, 918–930.
- Daraio, C., Bonaccorsi, A., & Simar, L. (2015b). Efficiency and economies of scale and specialization in European universities. A directional distance approach. *Journal of Informetrics*, 9, 430–448.

- Daraio, C., Heitor, M., Meoli, M., & Paleari, S. (2019). Policy turnaround: Towards a new Deal for research and higher education. Evaluation, rankings and governance for a new era of data science. *Higher Education Quarterly*, 73(1), 3–9.
- Daraio, C., & Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: A probabilistic approach. *Journal of Productivity Analysis*, 24(1), 93–121.
- Daraio, C., & Simar, L. (2007a). Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications. Springer Science & Business Media.
- Daraio, C., & Simar, L. (2007b). Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach. *Journal of Productivity Analysis*, 28(1–2), 13–32.
- Daraio, C., Simar, L., & Wilson, P. W. (2018). Central limit theorems for conditional efficiency measures and tests of the “Separability” condition in nonparametric, two-stage models of production. *The Econometrics Journal*, 21, 170–191.
- Daraio, C., Simar, L., & Wilson, P. W. (2021). Quality as a latent heterogeneity factor in the efficiency of universities. *Economic Modelling*, 99(2021), 105485.
- De Witte, K., & Rogge, N. (2011). Accounting for exogenous influences in performance evaluations of teachers. *Economics of Education Review*, 30(4), 641–653.
- Ferrari, G., & Laureti, T. (2005). Evaluating technical efficiency of human capital formation in the Italian university: Evidence from Florence. *Statistical Methods and Applications*, 14(2), 243–270.
- Francesconi, M., & Parey, M. (2018). Early gender gaps among university graduates. *European Economic Review*, 109, 63–82.
- Hall, P., Racine, J., & Li, Q. (2004). Cross-validation and the estimation of conditional probability densities. *Journal of the American Statistical Association*, 99(468), 1015–1026.
- Johnes, J. (2006). Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review*, 25(3), 273–288.
- Johnes, J. (2015). Operational research in education. *European Journal of Operational Research*, 243(3), 683–696.
- Johnson, A. L., & Ruggiero, J. (2014). Nonparametric measurement of productivity and efficiency in education. *Annals of Operations Research*, 221(1), 197–210.
- Klumpp, M. (2015). Efficiency in higher education: Requirements, theory, methods and decision areas. In J. Huisman & M. Tight (Eds.), *Theory and Method in Higher Education Research* (pp. 93–118). Emerald Group Publishing.
- Laureti, T. (2008). Modelling exogenous variables in human capital formation through a heteroscedastic stochastic frontier. *International Advances in Economic Research*, 14(1), 76–89.
- Laureti, T., Secondi, L., & Biggeri, L. (2014). Measuring the efficiency of teaching activities in Italian universities: An information theoretic approach. *Economics of Education Review*, 42, 147–164.
- Li, Q., & Racine, J. S. (2007). *Nonparametric econometrics: Theory and practice*. Princeton University Press.
- Mastromarco, C., & Simar, L. (2015). Effect of FDI and time on catching up: New insights from a conditional nonparametric frontier analysis. *Journal of Applied Econometrics*, 30(5), 826–847.
- Mastromarco, C., Toma, P., & Daraio, C. (2019). Exploring the teaching activities of the Italian universities through conditional efficiency analysis. In G. Catalano, C. Daraio, M. Gregori, H. F. Moed, & G. Ruocco (Eds.), *Proceedings of the 17th international conference on Scientometrics & Informetrics, 2-5 September 2019* (pp. 2656–2657). Sapienza University of Rome, Italy, Edizioni Efestò, ISBN: 978-88-3381-118-5, August 2019.
- Nigisch, S., & Schenker-Wicki, A. (2015). Frontier efficiency analysis in higher education. In *Incentives and performance* (pp. 155–170). Springer.
- Ruggiero, J. (1996). On the measurement of technical efficiency in the public sector. *European Journal of Operational Research*, 90, 553–565.

- Witte, K. D., & López-Torres, L. (2017). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society*, 68(4), 339–363.
- Worthington, A. C. (2001). An empirical survey of frontier efficiency measurement techniques in education. *Education Economics*, 9(3), 245–268.

**Camilla Mastromarco** (Ph.D., University of Glasgow) is a Professor of Econometrics at the University of Calabria. Her research interests are in the areas of parametric and nonparametric econometrics.

**Pierluigi Toma** (Ph.D., University of Salento) is an Assistant Professor of Econometrics at the University of Salento. His research interests are related to nonparametric methods applied to environmental and innovation economics.

**Cinzia Daraio** (Ph.D., Sant'Anna School of Advanced Studies) is a Professor of Management Engineering at the University of Rome La Sapienza. Her research interests include methodological and empirical studies in efficiency analysis, science and technology indicators, and higher education microdata.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.

