UNIVERSITÀ DEGLI STUDI DI TORINO

DOCTORAL THESIS

ESSAYS ON CAREER ADVANCEMENT IN ITALIAN ACADEMIA

Author:
Massimiliano CODA ZABETTA

Supervisor:
Prof. Aldo GEUNA

Submitted to the Department of Economics and Statistics "Cognetti de Martiis" in partial fulfillment of the requirements for the degree of Doctor of Philosophy

of the

Vilfredo Pareto Doctorate in Economics — Cycle XXX
Doctoral School in Human and Social Sciences

April 2018

Acknowledgements

This thesis is the result of three years of study and research. It would be just impossible to thank all the people that, in different ways and at different stages of this work, were helpful to its realization. But I feel the need to take few lines to express my gratitude to some of them.

I am profoundly indebted with Professor Aldo Geuna. He has been an attentive supervisor. His passion, enthusiasm and support all these years taught me a lot about how academic research really should be done in this field.

I had the chance to meet and discuss both with distinguished scholars and young researchers. I am most grateful to Paula Stephan, for her useful suggestions and sharp comments, Stefano Verzillo and Orion Penner, who helped me in solving several issues with data and matching techniques. Their help is gratefully acknowledged.

All suggestions, discussions and feedbacks by Professors and PhD colleagues at the Department of Economics of the University of Turin have been crucial in the early, middle and final stage of this thesis. With some of them the months of work and long talks naturally evolved in friendship: without Enrica, Gianluca and Matteo these years would have simply been poorer.

Contents

Acknowledgements **iii**

Introduction **1**

1 Italian Doctorate Holders and Academic Career Progression **5**

1.1 Introduction 5

1.2 Data sources 7

1.2.1 Italian doctorate holders 7

Data retrieval from BNCF 9

1.2.2 Italian academics 14

1.2.3 International scientific publications 18

1.3 Matching Italian doctorate holders and academics 21

1.4 PhD holders' career: descriptive statistics and stylized facts 24

1.4.1 Seniority and gender 25

1.4.2 Inbreeding and mobility 26

1.4.3 University prestige 28

1.5 Conclusions 32

2 Early Career Decisions and Advancement in Italian Academia **35**

2.1 Introduction 35

2.2 Conceptual framework 37

2.3 Institutional setup 40

2.4 Expectations 42

2.5 Data 44

2.5.1 Data generation 44

2.5.2 Data description 46

Inbreeding 47

Mobility paths 48
Post-doc and research stays 48
Economic fields 50
2.6 Method and variables 52
2.6.1 Empirical strategy 52
2.6.2 Variables and controls 56
2.7 Results 58
2.7.1 Time to first appointment 58
2.7.2 Time to tenure 61
2.7.3 Additional networks controls 65
2.7.4 Time varying covariates 65
2.8 Conclusions 69

3 Promotion and Connections in Academic Committees 71

3.1 Introduction 71
3.2 The role of mentorship in academic careers 73
3.3 Institutional background 75
3.3.1 Selection of committees 76
3.3.2 Evaluations 76
3.4 Data 77
3.4.1 Applications 77
3.4.2 Candidates 78
3.4.3 Connections 79
3.4.4 Future performance 79
3.5 Empirical analysis 80
3.5.1 Impact of connections on success 81
Candidates' and committees' research quality 83
3.5.2 Long-term efficiency of connections 87
3.6 Conclusions 88

4 Gender Discrimination in Academic Promotions in Italy 91

4.1 Introduction 91
4.2 Institutional background 93

4.3 Theoretical expectations 96
4.4 Berlinguer's reform (1998) 98
4.4.1 Data and methods 98
Matching methods 100
4.4.2 Empirical strategy 101
4.4.3 Results 102
Propensity score matching 102
Coarsened exact matching 105
4.5 Gelmini's reform (2010) 109
4.5.1 Data and methods 109
4.5.2 Empirical strategy 109
4.5.3 Descriptive statistics 110
4.5.4 Results 112
Gender difference in success in habilitation concourse 112
Likelihood to enter the habilitation concourse 114
Networks effects 117
The competition environment 120
4.6 Conclusions 121

Concluding Remarks **123**

A Appendix to Chapter 1 **127**

B Appendix to Chapter 2 **139**

C Appendix to Chapter 3 **141**

D Appendix to Chapter 4 **143**

Bibliography **145**

List of Figures

1.1 Number of PhD holders every 10,000 inhabitants 6

1.2 Number of doctoral courses, positions, scholarships and enrollements. Source: MIUR 8

1.3 Number of doctorate holders (Source: MIUR) and thesis cataloged at the BNCF 10

1.4 Number of doctorate holders (Source: MIUR) and thesis cataloged at the BNCF by macro scientific area 15

1.5 Share of Italian professors by academic position 1990-2015 (Source: MIUR) 15

1.6 Number of professors by academic position and year of recruitment (Source: MIUR) 17

1.7 Number of professors by academic position and exit year (Source: MIUR) 17

1.8 Number of Scopus publications 20

1.9 Scientific proximity clusters for the record linkage 22

1.10 Weight formula for the record linkage 23

1.11 Hiring networks (left: STEM; right: SSH). Vertex size: in-degree; vertex color: out-degree 30

1.12 Distribution of rank changes (left: STEM; right: SSH) 32

2.1 Yearly entrances and promotions in Italian academia, 1990-2015 42

2.2 Kaplan-Meier survival estimate, whole sample (left) and by scientific field (right) 58

2.3 Kaplan-Meier survival estimate, whole sample (left) and by scientific field (right) 62

2.4 Hazard rate function, whole sample (left) and by gender (right) 62

2.5 Predicted survival probabilities by *ResearchAbroad* 69

4.1 Promoted associate and full professors and share of promoted women (MIUR) 96

4.2 Propensity score histograms by treatment and control groups 104

A.1 XML tags of a thesis stored in the BNCF OPAC 128

A.2 Example of the record linkage 130

A.3 Graphical representation of the record linkage 131

A.4 Thesis linked to an academic (%) 132

A.5 Academics linked to thesis by macro-area (86-06) 137

A.6 Academics linked to thesis by SSD (86-06) 137

List of Tables

1.1 Number of doctorate holders (Source: MIUR), thesis deposited at the BNCF and differences by year 11
1.2 Number of doctorate holders (Source: MIUR), thesis deposited at the BNCF and differences by institution 12
1.3 Number of doctorate holders, thesis deposited at the BNCF and differences by scientific field (1:BNCF, 2:MIUR) 13
1.4 Number of Italian professors by academic position 1990-2015 (Source: MIUR) 16
1.5 Contents of Italian academics MIUR data 18
1.6 Publications, academics and authors on Scopus by academic position . 20
1.7 Publications, academics and authors on Scopus by scientific field . . . 20
1.8 Examples of score values for name matching 22
1.9 Examples of score values for year matching 23
1.10 Age at first appointment and tenure and percentage tenured by scientific area and PhD 26
1.11 Gender by scientific area and PhD 26
1.12 Inbreeding by scientific area 27
1.13 Mobility and promotion by scientific area and PhD 28
1.14 Promotion and mobility by scientific area and PhD 28
1.15 Summary statistics of the hiring networks 29
1.16 STEM PhDs hired from top 10 universities 31
1.17 SSH PhDs hired from top 10 universities 31

2.1 Individual characteristics Sample vs Population 46
2.2 Percentage of inbreds by institution type 48
2.3 Promotion pattern of mobile academics 48

2.4 Mobility pattern of promoted academics 48
2.5 Postdoctoral mobility by PhD cohort 49
2.6 Performance, age and ranking measures by type of mobility for tenured academics 50
2.7 Candidates' published oeuvre by SSDs 50
2.8 Candidates' published oeuvre by SSDs if tenured or censored 51
2.9 Duration until tenure by SSDs 51
2.10 Descriptive information on the variables 55
2.11 Cox hazard model for the first appointment in Italian academia 60
2.12 Cox hazard model of the hazard of tenure in Italian academia 63
2.13 Cox hazard model of the hazard of tenure in Italian academia (with centrality controls) 66
2.14 Cox hazard model with time varying covariates 68

3.1 Descriptives statistics: examinations 78
3.2 Descriptives statistics: candidates 79
3.3 Descriptive statistics: connections 79
3.4 Impact of connection on candidates' success (ITT) 81
3.5 Impact of connection on candidates' success (IV) 84
3.6 Impact of connection on candidates' success (by candidates' productivity) 85
3.7 Impact of connection on candidates' success (by committees' productivity) 86
3.8 Future performance 87

4.1 Professors in Italian academia by gender and scientific field (MIUR) . . 92
4.2 Promotion by gender in the pre-reform waves 101
4.3 Professors with at least one Scopus publication by scientific field 103
4.4 Population propensity scores descriptives statistics 103
4.5 DD estimation - Marginal effects 104
4.6 Descriptive statistics of bibliometric indicators by disciplinary area . . 106
4.7 Variables used for coarsening 106
4.8 Treated and untreated units by CEM groups) 107

4.9 $L1$ measure for matched and original population 107
4.10 DD estimates – Marginal effects 108
4.11 Academics by rank and gender in 2015 110
4.12 Descriptives of potential candidates 111
4.13 Likelihood to be habilitated conditional on applying 113
4.14 Likelihood to be habilitated conditional on applying (by academic rank) 115
4.15 Likelihood to apply for habilitation (Marginal effects) 116
4.16 Network effects: likelihood to apply for habilitation (Marginal effects) 119
4.17 The competition environment: likelihood to apply for habilitation (Marginal effects) 120

A.1 Summay of the studies 129
A.2 Links "academic-thesis" identified per broad scientific area and year (1986-2006) 133
A.3 Thesis not linked per broad scientific area and year (1986-2006) 133
A.4 Percentage of thesis linked to an academic per broad scientific area and year (1986-2006) 133
A.5 Academics linked to thesis per disciplinary area and expected PhD year (1986-2006) 134
A.6 Academics not linked to a thesis per disciplinary area and expected PhD year (1986-2006) 135
A.7 Share of academics linked to a thesis per disciplinary area and expected PhD year (1986-2006) 136

B.1 Summary statistics (time-to-first appointment) 140
B.2 Impact of connection on candidates' success – Logit (Marginal effects) 140

C.1 Impact of connection on candidates' success – Logit (Marginal effects) 142
C.2 First-stage estimates 142

D.1 Descriptive statistics by treated and controls 144
D.2 Descriptive statistics of matched units by treated and controls 144

Introduction

Academia remains a major locus of knowledge production, generating a considerable fraction of the scientific research articles (National Science Board, 2008). Universities compete to improve their reputation in local and global academic markets, and attracting and hiring talented researchers is a key dimension in this competitive environment (OECD, 2008). There is a worldwide trend for new performance-based funding models of universities (OECD, 2010) which increasingly turned the attention of policy makers and university managers to the need of selecting the best possible researchers.

The analysis of the labor markets of academics and their careers, access to tenure, and academic promotion has a long tradition of studies centered on the USA (see Long and Fox, 1995 for a review), but, because of national diversity, analysis of academic labor markets in Europe is much scarcer.

In this context, this thesis seeks to fill a significant gap in the literature on the Italian higher education system, by looking at issues of whether access to a permanent academic position is governed by merit and universalism or more biased and particularistic factors. In particular, in order to take into account all the relevant factors, it is fundamental to start from the very beginning of scientists' academic career, namely the PhD degree. In fact, as I will show, the human and scientific relations built (in particular with the academic advisor) and the decisions made (in particular regarding mobility) at this early stage will be very influential throughout the whole researchers' career.

Doctoral programs students are generally considered to be of utmost importance in the generation of new knowledge in our economy, also playing a central role in the dissemination of scientific and technical human capita (Bozeman et al., 2001). Nonetheless, the information about these crucial players in knowledge creation and diffusion is surprisingly scarce and fragmented. For this reason, in order to build the empirical basis for my research, I have expressly assembled an *ad-hoc* database of Italian PhD holders in all disciplines from the first doctoral cycle (1986) to 2006. I matched this information with the institutional database of Italian scientists (provided by the Italian Ministry of University and Research – MIUR). In this way I built a database of academic careers of Italian university professors since the PhD over the last 30 years.

All the four chapters are written as single working papers to be read individually by the readers. Therefore it is possible to note a certain degree of overlapping that was clearly unavoidable. In the remaining part of this introduction I briefly introduce the topic of each paper.

Chapter 1 describes the methodology used to build a database on doctorate holders in all disciplines from Italian universities in the period from the first cycle of doctorate until 2006. The doctorates who pursued an academic career in Italy have been identified by matching with academics in the official statistics of the Italian Ministry of Education. I exploited an unused source of data, i.e. the repository of doctoral dissertations at the national library of Florence (*Biblioteca Nazionale Centrale di Firenze* – BNCF), established by the MIUR in 1980. The BNCF provides the identity and affiliation of dissertations' authors, advisors and in some cases committee members. Therefore, it allows to shed light on the employment outcomes of researchers who have been formed in Italian universities in a time period of 20 years.

The objective of Chapter 2 is to investigate the relationship between early career choices of researchers and permanent academic positions. My aim is to approach some of the dynamics of research careers in early stages by analyzing the relationship among tenure, mobility and scientific production. The main research question I want to address is "What does explain the timing of academic career progression?", testing the relative impact of productivity, mobility and other early career variables, In particular, I want to disentangle the effect of mentoring on career success thanks to the use of bibliometric indicators and social network variables. I use the dataset described in Chapter 1, based on institutional longitudinal data and survival analysis (event history analysis) to explore time to promotion and its covariates. In particular I seek to deepen the understanding about the role of early academic performance, mobility and social embeddedness in the period between PhD and tenure.

The aim of Chapter 3 is to analyze the impact of connections between candidates and evaluators in affecting the selection process in academia. I will use data from the centralized selection exams in Italy to highlight the role of social capital in the decisions taken by evaluating committees, distinguishing between strong and weak ties with influential actors. In particular, I will devote attention to the relationship established by researchers and the time of PhD education (namely with the advisor, PhD colleagues and advisor's coauthors), a point which is still unexplored in the literature in the case of Italian academia. I exploit the evidence provided by centralized selection evaluations in Italian academia in 2012, the year of the first round of the national habilitation, which qualifies as a large scale randomized natural experiment allowing the estimation of the effect of connections on candidates' chances of success.

Recent empirical research based on professorial appointments indicates that many mechanisms prevalent in recruitment and appointment practices of professors are disadvantageous to the career of academic women (Brink and Benschop, 2012). Since late '90s, the Italian academic recruiting system went through several reforms. i) Until 1998, career progressions were made in a nationwide public competition where the panel of commissioners was elected by the whole body of professors of the academic sub-discipline; ii) Law 210/1998 moved the recruiting system towards a decentralized mechanism: universities were entitled to manage their own concourses to fill the vacancies; iii) the most recent reform (2010) introduced a two-step procedure with a national habilitation and local concourses to recruit professors. The last Chapter aims at understanding if gender makes a difference in the path to promotion in Italian universities, in the context of the two most recent reforms of Italian academic labor market

Chapter 1

Italian Doctorate Holders and Academic Career Progression — 1987-2015

1.1 Introduction

Doctoral graduates are key players for research and innovation. They have been specifically trained to conduct research and are considered the best qualified for the creation and diffusion of scientific knowledge. The number of doctoral students provides a measure of a country's potential research capability. There were an estimated 717 thousand doctoral students in the EU-28 in 2012, compared with levels of 492 thousand in the United States and 75 thousand in Japan (European Commission, 2014).

In Italy, starting from the end of the 1990s, the number of places for PhD students largely increased: students passed from 20,657 in 2000/01 to 33,037 in 2013/14, while doctorate holders passed from 3,977 to 10,182. Nonetheless even after 30 years from the introduction of the first PhD cycle, Italy still is in a bottom end position with respect to other European countries taking into account the diffusion of the PhD title within the population. With 28 doctorates every 10,000 inhabitants Italy exhibits in fact one of the lowest ranking in Europe, the average for all countries is 65 (dotted line in Figure 1.1).

The importance of this academic title fostered the necessity of having reliable

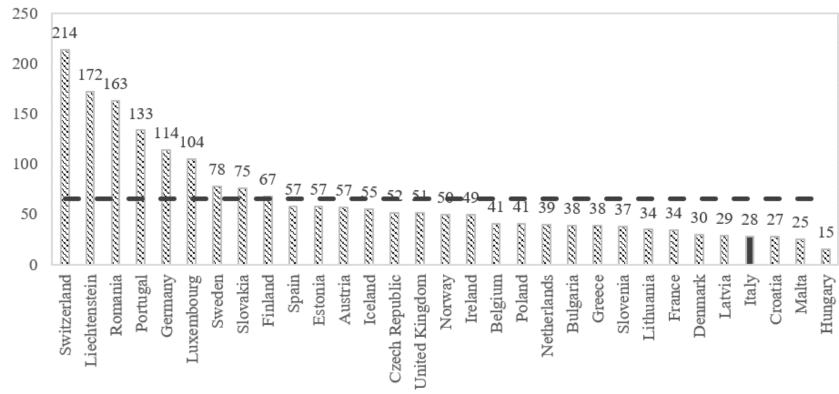

FIGURE 1.1: Number of PhD holders every 10,000 inhabitants

data through which making in depth analysis in particular about the career of doctorate holders, necessity that found an answer within the "Project on Careers of Doctorate Holders" launched by OECD, Eurostat and Unesco in 2004. This project aims at developing constant and internationally comparable data and indicators on the career and mobility phenomena of doctorates (Auriol et al., 2013).

Taking into account the Italian case, it is quite difficult to find reliable data on this phenomenon. For a long time, the only database with respect to this issue has been the one of Cineca, which was just about the offer of doctoral courses in the Italian universities. Starting from 1998 the statistical office of the Italian ministry of university and research (MIUR) registers, not without losses, data concerning students and doctorate holders from Italian universities. Beside this, there have been just sporadically some attempts to conduct surveys on the career outcome of doctorates who graduated from specific universities[1]

The database here presented is one of the first attempts of creating a database on doctorate holders who obtained their title from all Italian universities in all scientific fields. It contains information about doctorate holders in all disciplines from Italian universities in the period from the first cycle of doctorate until 2006. The doctorates who pursued an academic career in Italy have been identified by matching with academics in the official statistics of the Italian Ministry of Education (MIUR).

[1]In 2006 CNVSU and MIUR conducted an analysis for the doctorate holders from universities of Pavia, Pisa, Siena and Salerno; in 2006 university of Trento conducted an analysis of the occupational outcomes of doctorate holders from universities of Milano, Milano-Bicocca and Trento; in 2009 Scuola Superiore Sant'Anna conducted an analysis of doctorates from seven universities (Bergamo, Brescia, Milano, Milano-Bicocca, Palermo, Pisa and Scuola Superiore Sant'Anna); in 2009 and 2014 ISTAT conducted a survey on the career of Italian doctorates.

We exploited an unused source of data, i.e. the repository of doctoral dissertations at the national library of Florence (*Biblioteca Nazionale Centrale di Firenze* – BNCF), established by the MIUR in 1980. The BNCF provides the identity and affiliation of dissertations' authors, advisors and in some cases committee members. Therefore, it allows to shed light on the employment and outcomes of researchers who have been formed in Italian universities in a time period of 20 years.

The remainder of the paper is organized as follows: Section 1.2 presents the data sources we used, describes the retrieval procedure and presents some summary statistics for each of them; in Section 1.3 we discuss the record linkage performed between BNCF and MIUR data; Section 1.4 presents an exploratory analysis and some stylized facts about the academic career of Italian PhD holders; Section 1.5 provides a conclusion.

1.2 Data sources

We have collected information from three primary sources: the National Library of Florence (BNCF), the Italian Ministry of education (MIUR) and the bibliographic web version of Scopus database. Hereafter we describe the data retrieval process and provide some descriptive statistics for each of the data sources.

1.2.1 Italian doctorate holders

In Italy doctorates have been created with the decree 382/1980 *'Riordinamento della docenza universitaria relativa fascia di formazione, nonché sperimentazione organizzativa e didattica"* and later modified with the Law 476/1984 *'Norme in materia di borse di studio e di Dottorato di ricerca nelle università"*. This has been the way through which Italy uniformed itself to the rest of the countries: identifying a clear path to train the new researchers and future academics. The main modifications to the doctorates occurred with the Law 289/1989 which assigned to the universities the power to decide on the post-doctorate scholarships; Law 210/1998 which explicitly states that the doctorate is mandatory to become researcher at universities and other research institutes and, together with Law 224/1999, gave the universities the possibility of autonomously create doctoral courses and increase the yearly number of candidates through the positions without scholarship. In the period after these two reforms there has been a strong increase in the doctoral courses created by universities (see Figure 1.2): in the period between 1999 and 2006 the number of courses increased from 1,816 to over 2,200, to then decrease progressively in the following years. This is due to the decree 124/2007, that merged the courses into the doctoral schools and introduced further requirements for the creation of courses, and to the Law 240/2010, actuated by the ministerial decree 45/2013, according to which the activation of doctoral courses should be authorized by the Italian National Agency for the Evaluation of the University and Research Systems (ANVUR). This has led, starting from 2013, to a strong reduction of the courses offered by the universities.

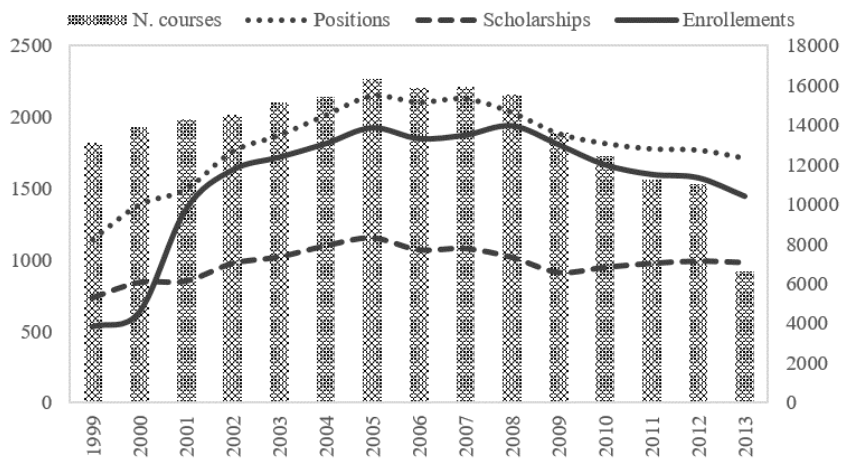

FIGURE 1.2: Number of doctoral courses, positions, scholarships and enrollements. Source: MIUR

In Italy there is no a central archive which keeps track of national PhDs and doctorate holders as in the case of university professors (see next section). The ministerial decree 45/2013 also states that the Italian Ministry of University and Research has to introduce a registry of the doctorate holders and of the doctoral thesis, in order to further investigate their career and occupational outcomes. These registries (*Anagrafe Dottorati* and *Anagrafe Dottorandi*) are available starting from the academic year 2003/2004 but have mainly administrative purpose and the university departments are not legally obliged to fulfill and update them. In particular, the registry of doctorate holders just contains information about the students enrolled in the first year of their PhD and, at the end of the doctorate, the result of their thesis defense but this outcome is not fully reported for the whole ensemble of the doctoral students, even after several years from the enrollment in the PhD. What universities are obliged to do, since the institution of doctorates in 1980, is to deposit a copy of the doctoral thesis at the BNCF. For these reasons we decided to retrieve the data about Italian doctorate holders from the open repository of this library. This is one of the first attempts of creating a database on doctorate holders who obtained their title from Italian universities in all scientific fields.

Data retrieval from BNCF

Using specific key query parameters in order to identify the doctoral thesis cataloged, we harvested all the records from the on-line public access catalog (OPAC) of the BNCF from 1986, when the PhD students of the first cycle had their dissertation, to 2015. The OPAC of the BNCF allows the download of the query results in XML format, which can be then easily codified in a spreadsheet, but just for one query result at time, so we built an automating harvesting program in order to do that for all the records (over 100 thousand). In order to avoid to have too much noise, we downloaded specific subfields of the XML structure of the records which contain this information: author, year of publication, university of affiliation, cycle, title, supervisor, scientific sector. Figure A.1 in Appendix A gives an example of the codification in a spreadsheet of the XML tags of a thesis stored in the BNCF OPAC.

We then went through a process of data cleansing where a number of duplicate and incorrect records were detected and definitively deleted from the dataset. The process of cleaning followed these steps:

- The original information from BNCF is parsed into several fields of a spreadsheet, getting rid to all the noise which come with the unnecessary strings of code.
- Parsed data on publication year and university of affiliation are cleaned by fixing spell mistakes and standardizing university names. If the corresponding field in the XML was empty the information has been retrieved from other fields, whenever possible.

- Taking into account the MIUR classification and the Dewey classification we grouped the thesis in six broad scientific areas: Medicine&Veterinary, Science (agriculture, biology, chemistry, physics, geology and mathematics), Architecture&Engineering, Humanities&Law, Social Sciences and Economics&Statistics. If those fields were empty in the XML, we retrieved the information about the title of the doctoral course and confronted it with the official Cineca repository of doctoral courses offered by Italian universities[2] which associates to each course a macro-scientific area.

- We identified and removed duplicates records, even when they presented different identifiers but the same combination of author name, publication year, university affiliation, Dewey and MIUR classifications.

We confronted the data with the information on doctorate holders from the statistical office of the MIUR, which are just available online for the period starting in 1998, and with the data from the same source reported by Avveduto and Brandi, 2004, one of the few studies covering the period 1987-1998[3].

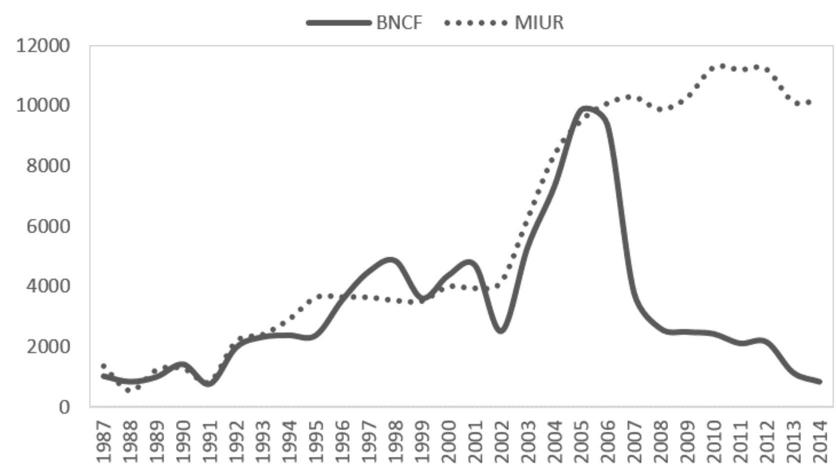

FIGURE 1.3: Number of doctorate holders (Source: MIUR) and thesis cataloged at the BNCF

As it is possible to see from Figure 1.3, there has been a substantial decline in the thesis cataloged at the national library of Florence after 2006. After the DPR

[2]See http://cercauniversita.cineca.it/php5/dottorati/cerca.php

[3]We do not take into account the so called *istituti a ordinamento speciale* (institutes with a special statute): Normale of Pisa, Sant'Anna of Pisa, SISSA of Trieste, IMT of Lucca, IUSS of Pavia and SUM of Florence. Their third level courses, in fact, are equalized to the doctoral courses of public and private Italian universities, but are not subject to the law on the legal deposit of the doctoral thesis at the BNCF.

TABLE 1.1: Number of doctorate holders (Source: MIUR), thesis deposited at the BNCF and differences by year

	MIUR	BNCF	%Δ
1998	2803	4847	-72.9
1999	3500	3605	-3
2000	3977	4362	-9.7
2001	3924	4706	-19.9
2002	4139	2508	39.4
2003	6249	5321	14.8
2004	8346	7316	12.3
2005	9477	9849	-3.9
2006	10057	9364	6.9
Tot	52472	51878	1.1

252/2006 and the MIUR Communication 1746/2007, in fact, the BNCF launched a project aimed at automatically harvesting the doctoral thesis deposited in the open access repositories of the participant universities. This procedure, however, did not completely substitute the legal deposit of the paperback thesis, which some departments continued doing together with the digital deposit in their databases. Furthermore, the BNCF continued cataloging just the paperback thesis sent by the universities and did not do the same for those digitally harvested.

However, considering the period 1987-2006, we can see that there are small differences between the data reported by MIUR and the one obtained from BNCF. According to MIUR the total number of PhD students who discussed their thesis in the period 1986-2006 is 76,681, while in the same period the thesis deposited at the BNCF are 73,950, thus with a 3.5 percentage points of imbalance with respect to the available official statistics. This slight difference is due to some limitation of the ministerial data, and to some small delay in the deposit of the thesis by the university departments or to their digital cataloging at the library. Yearly differences by year, university and scientific area are reported respectively in Table 1.1, Table 1.2 and Table 1.3.

For the previous cycles (from 1986 to 1997) there are not official data which we can use to compare the database with, but just few studies (see for example Cipollone and Avveduto 1998), that report statistics which are consistent with data contained in the database.

Even taking into account the yearly evolution of the scientific macro-areas to

TABLE 1.2: Number of doctorate holders (Source: MIUR), thesis deposited at the BNCF and differences by institution

University	MIUR	BNCF	% Δ	University (cont'd)	MIUR	BNCF	% Δ
Ancona	618	567	8.3	Napoli - "Parthenope"	167	159	4.8
Arcavacata di Rende	646	630	2.5	Napoli - "SOB"	16	21	-31.3
Bari - Politecnico	264	277	-4.9	Padova	2195	2590	-18.0
Bari	1843	1916	-4.0	Palermo	1710	1654	3.3
Benevento	72	74	-2.8	Parma	792	822	-3.8
Bergamo	60	87	-45.0	Pavia	1060	1008	4.9
Bologna	3582	3378	5.7	Perugia	946	1003	-6.0
Brescia	398	279	29.9	Pisa	2044	2066	-1.1
Cagliari	567	480	15.3	Potenza	282	274	2.8
Camerino	193	207	-7.3	Reggio Calabria	371	366	1.3
Campobasso	237	187	21.1	Roma - III	556	569	-2.3
Cassino	151	71	53.0	Roma - LUISS	89	84	5.6
Castellanza	20	19	5.0	Roma - LUMSA	15	16	-6.7
Catania	2002	1988	0.7	Roma - "La Sapienza"	2665	3577	-34.2
Catanzaro	55	57	-3.6	Roma - "Foro Italico"	6	0	100.0
Chieti e Pescara	591	605	-2.4	Roma - "Tor Vergata"	1341	1235	7.9
Ferrara	621	668	-7.6	Roma - UNINT	2	8	-300.0
Firenze	2535	2322	8.4	Salerno	672	587	12.6
Foggia	143	142	0.7	Sassari	495	384	22.4
Genova	1495	1625	-8.7	Siena	1188	1053	11.4
L'Aquila	405	450	-11.1	Siena - Stranieri	35	35	0.0
Lecce	561	526	6.2	Teramo	120	130	-8.3
Macerata	236	234	0.8	Torino - Politecnico	982	994	-1.2
Messina	989	1029	-4.0	Torino	1746	1728	1.0
Milano - IULM	47	46	2.1	Trento	540	500	7.4
Milano - Politecnico	1231	1201	2.4	Trieste	802	905	-12.8
Milano - "Sacro Cuore"	933	925	0.9	Udine	387	425	-9.8
Milano - Bocconi	217	234	-7.8	Urbino	486	397	18.3
Milano	2359	2488	-5.5	Varese	83	91	-9.6
Milano - San Raffaele	10	10	0.0	Venezia - "Cà Foscari"	474	494	-4.2
Milano - Bicocca	332	331	0.3	Venezia - IUAV	193	203	-5.2
Modena e R. Emilia	608	580	4.6	Vercelli	129	116	10.1
Napoli - Seconda	1053	486	53.8	Verona	388	389	-0.3
Napoli - "Federico II"	3760	3376	10.2	Viterbo	247	25	89.9
Napoli - "L' Orientale"	414	188	54.6	Tot	52472	51878	1.1

TABLE 1.3: Number of doctorate holders, thesis deposited at the BNCF and differences by scientific field (1:BNCF, 2:MIUR)

	1987			1988			1989			1990			1991			1992			1993			1994			1995			1996		
	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ
Science	323	399	19	303	66	-359	341	412	17	588	420	-40	264	158	-67	750	770	3	944	807	-17	993	886	-12	928	1,015	9	1,222	1,013	-21
Hum&Law	266	388	31	205	107	-92	243	271	10	296	240	-23	196	221	11	391	456	14	387	385	-1	386	561	31	436	783	44	850	824	-3
Econ&Stat	66	57	-16	35	8	-338	65	61	-7	61	37	-65	37	46	20	124	108	-15	128	96	-33	137	177	23	144	232	38	163	243	33
Med&Vet	99	182	46	144	258	44	149	197	24	236	352	33	123	270	54	264	370	29	412	652	37	329	663	50	334	771	57	550	713	23
Arch&Eng	190	254	25	89	10	-790	160	226	29	183	193	5	101	92	-10	353	406	13	346	392	12	465	544	15	459	738	38	681	749	9
SocSci	47	62	24	20	3	-567	27	51	47	31	22	-41	15	31	52	58	57	-2	56	56	0	62	67	7	67	95	29	81	82	1
Tot	991	1,342	26	796	452	-76	985	1,218	19	1,395	1,264	-10	736	818	10	1,940	2,167	10	2,273	2,388	5	2,372	2,898	18	2,368	3,634	35	3,547	3,624	2

	1997			1998			1999			2000			2001			2002			2003			2004			2005			2006		
	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ	1	2	%Δ
Science	1,533	1,045	-47	1,723	1,013	-70	1,189	1,285	7	1,438	1,574	9	1,520	1,504	-1	842	1,363	38	1,824	2,095	13	2,334	2,573	9	3,221	2,789	-15	3,029	3,055	1
Hum&Law	1,055	711	-48	1,056	800	-32	793	766	-4	998	804	-24	1,047	837	-25	547	936	42	1,181	1,451	19	1,668	2,067	19	2,483	2,376	-5	2,433	2,532	4
Econ&Stat	262	291	10	304	200	-52	269	277	3	351	273	-29	430	291	-48	187	253	26	307	379	19	454	528	14	604	595	-2	661	600	-10
Med&Vet	707	673	-5	606	804	25	485	363	-34	574	408	-41	691	410	-69	392	640	39	786	957	18	1,165	1,363	15	1,474	1,506	2	1,366	1,550	12
Arch&Eng	827	692	-20	1,054	687	-53	788	717	-10	916	803	-14	915	736	-24	510	809	37	1,051	1,178	11	1,463	1,532	5	1,735	1,823	5	1,539	1,966	22
SocSci	104	108	4	106	111	5	88	92	4	92	115	20	114	59	-93	38	102	63	174	184	5	229	242	5	331	355	7	336	325	-3
Tot	4,488	3,520	-28	4,849	3,615	-34	3,612	3,500	-3	4,369	3,977	-10	4,717	3,837	-23	2,516	4,103	39	5,323	6,244	15	7,313	8,305	12	9,848	9,444	-4	9,364	10,028	7

which the doctoral dissertations refer to, reported in Figure 1.4, it is possible to observe a similar trend between the data gathered from the MIUR and the one collected from the BNCF. In addition to what already said for the overall evolution of the thesis discussed, the small discrepancies here are probably due to the fact that the data from the MIUR database take into account aggregations of the scientific disciplinary sector (SSD) of the doctoral course in which the PhD student defending the thesis is enrolled, while the database of the BNCF registers all its entries according to the Decimal Dewey Classification, which is a general classification for books, which we have thus matched with the university SSDs and, when possible, with the official ministerial data about the offer of doctoral courses provided by Cineca.

Since data about doctoral thesis retrieved from BNCF for years after 2006 are not representative for the population of doctorates holders[4], we will just consider the precedent period, i.e. from 1986 to 2006. In this time span, the total number of thesis in the database is 73,968, among these, 1,990 do not indicate the university of affiliation of the doctorate holder while 550 do not indicate the scientific field of the dissertation, thus these records require a further online search to fill this missing information.

1.2.2 Italian academics

The Italian academic system is composed of 96 universities (30 private and 66 public) and 8 higher education institutions. Each professor working at an Italian university is categorized by a level of arrangement (full professor, associate professor and assistant professor) and by one of the scientific sectors of discipline. We obtained information about academic position, disciplinary areas and university affiliation on all the Italian academics in the period 1990-2015 from MIUR and Cineca. Table 1.4 and Figure 4.1 summarize the number and the share of Italian professors by academic position in this time period.

In 1990 there were 42,209 professors active in Italian universities, in the time period considered 45,795 academics entered in the Italian system and 33,219 exited,

[4]For the period 2007-2014 the only reliable data are about these universities (in brackets the percentage of difference with respect to MIUR data): Polytechnic university of Bari (2.5%), “KORE” university of Enna (-11.9%), “Bocconi" university of Milan (4.6%), “San Raffaele" university of Milan (19.5%), university of Pisa (7.6%), Polytechnic university of Turin (3%).

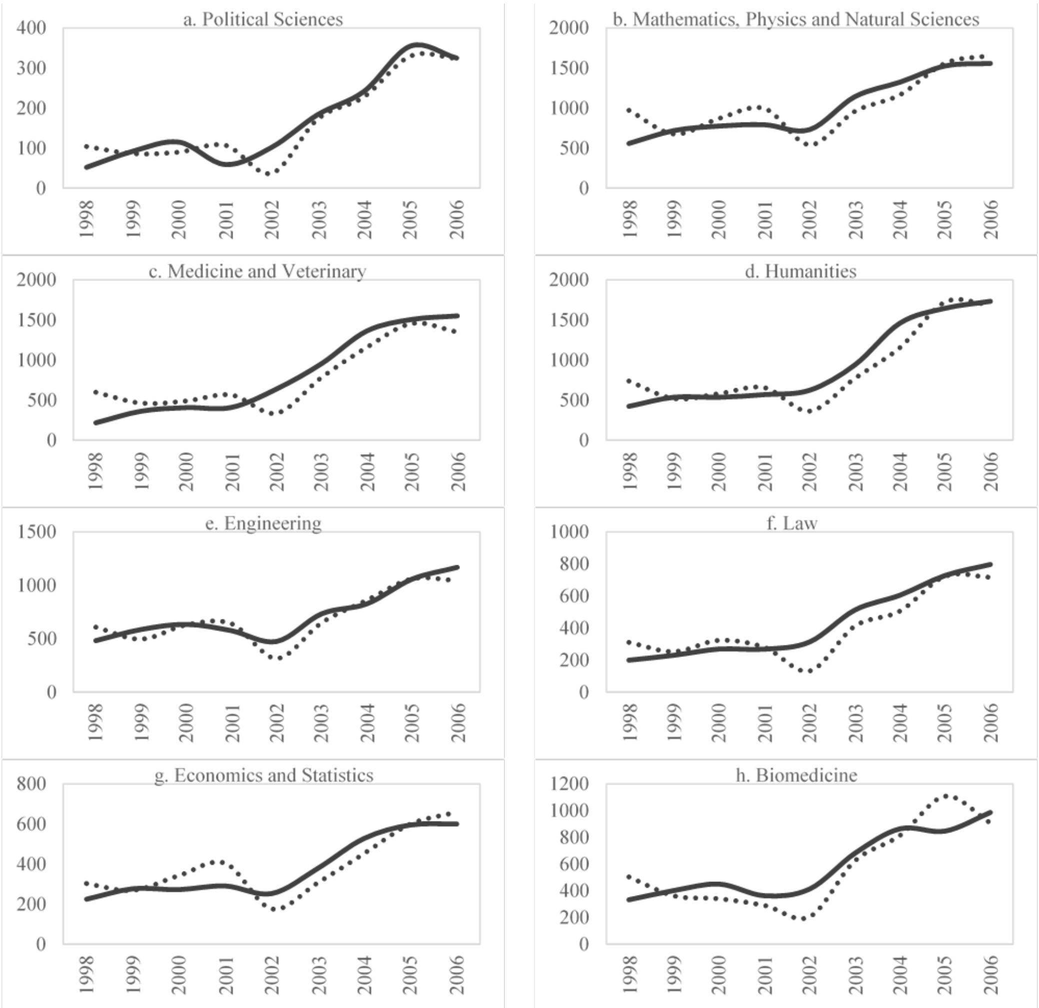

FIGURE 1.4: Number of doctorate holders (Source: MIUR) and thesis cataloged at the BNCF by macro scientific area

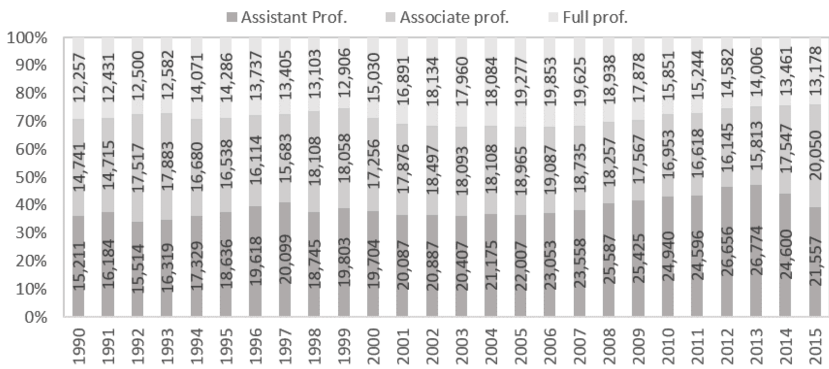

FIGURE 1.5: Share of Italian professors by academic position 1990-2015 (Source: MIUR)

TABLE 1.4: Number of Italian professors by academic position 1990-2015 (Source: MIUR)

Year	Assistant prof.	Associate prof.	Full prof.	Total
1990	15,211	14,741	12,257	42,209
1991	16,184	14,715	12,431	43,330
1992	15,514	17,517	12,500	45,531
1993	16,319	17,883	12,582	46,784
1994	17,329	16,680	14,071	48,080
1995	18,636	16,538	14,286	49,460
1996	19,618	16,114	13,737	49,469
1997	20,099	15,683	13,405	49,187
1998	18,745	18,108	13,103	49,956
1999	19,803	18,058	12,906	50,767
2000	19,704	17,256	15,030	51,990
2001	20,087	17,876	16,891	54,854
2002	20,887	18,497	18,134	57,518
2003	20,407	18,093	17,960	56,460
2004	21,175	18,108	18,084	57,367
2005	22,007	18,965	19,277	60,249
2006	23,053	19,087	19,853	61,993
2007	23,558	18,735	19,625	61,918
2008	25,587	18,257	18,938	62,782
2009	25,425	17,567	17,878	60,870
2010	24,940	16,953	15,851	57,744
2011	24,596	16,618	15,244	56,458
2012	26,656	16,145	14,582	57,383
2013	26,774	15,813	14,006	56,593
2014	24,600	17,547	13,461	55,608
2015	21,557	20,050	13,178	54,785

thus in 2015 the number of professors grew to 54,785. Figure 1.6 and Figure 1.7 show the distribution of professors by academic position and year of recruitment or exit from the system, respectively.

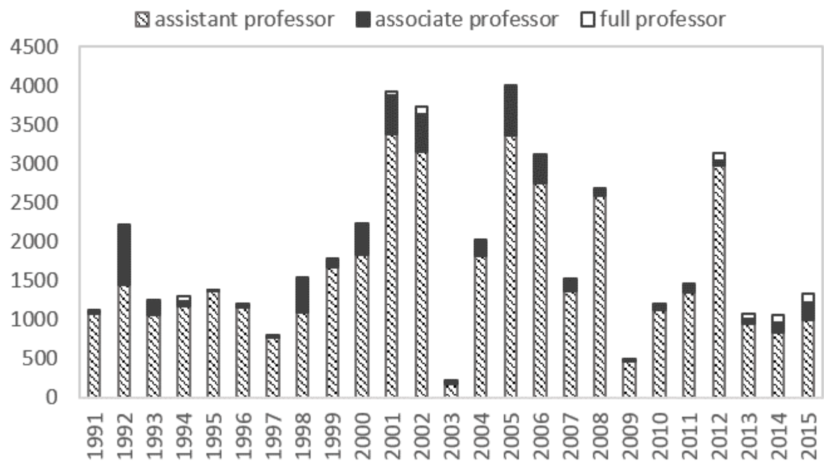

FIGURE 1.6: Number of professors by academic position and year of recruitment (Source: MIUR)

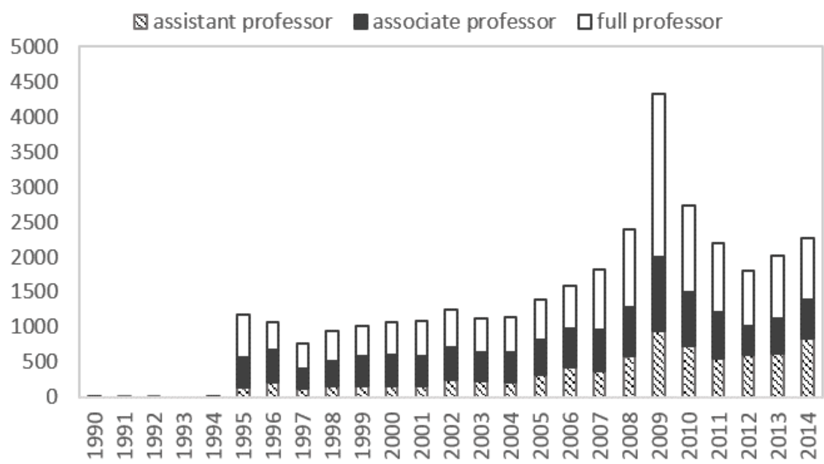

FIGURE 1.7: Number of professors by academic position and exit year (Source: MIUR)

In parallel to the creation of the doctorate database we proceeded to the collection of biographical information about Italian academics. The collection was directed to all scientific disciplines which for the purpose of this database we have classified in six macro-areas, as will be explained in next section.

TABLE 1.5: Contents of Italian academics MIUR data

Id
Surname
Name
Gender
Date of birth
Disciplinary field (per year)
Faculty and university (per year)
Rank (per year)

Information regarding academic positions, disciplinary areas, and university affiliation are available on-line from 2000 to 2015. We obtained data on academic careers before 2000 from Cineca, a MIUR agency which collects administrative data on personnel as well as on competition for professorship in Italy. These data have several known problems, often relating to the uniqueness of identifying codes of individuals, and missing data on academic disciplines over the first five years (1990-1995): we have corrected for these issues to the best of our ability.

Whatever their rank, Italian professors in public universities are tenured civil servants, and even those working in private universities are tenured and recorded for administrative purpose in the Ministry's list. However, the Ministry does not keep central records of PhD students nor of the numerous contract-based researchers and instructors who populate Italian universities. Table 1.5 summarizes the main contents of the Italian academics MIUR data. Notice that information of the professors' rank and affiliation allows for panel data or pooled cross section analysis of academic careers.

1.2.3 International scientific publications

Scopus database, published by Elsevier, indexes a greater number of journals (12,850, including 500 open access journals) within the medicine, technical and social sciences and offers systematic, quantifiable statistical information based on citation data.

The assignment of articles to academics is non-trivial. For each publication and author, Scopus provides information on the surname and on the initial (or, in some cases, initials). Problems with homonyms may arise in the case of common surnames. Further, authors may sign using their first name, their middle name, or both.

We use the following procedure in order to identify authors.

Through a Python script which performs repeated "surname+name" author search queries using Scopus API, we downloaded all available personal information for the list of academics. This information includes: affiliation, scientific research area and Scopus Author-ID, the latter is a unique identifier for each author inside Scopus database.

Then, we assigned all professors in the ministerial list and authors' record downloaded from Scopus to a broad disciplinary category. In order to attribute comparable disciplinary categories for authors and individuals, we aggregate disciplines defined by MIUR and Scopus disciplinary areas into the following categories: Agriculture; Chemistry; Biology; Physics; Mathematics and Computer Science; Architecture and Engineering; Medicine and Veterinary; Economics and Management; Humanities and Law, Sociology and Political Science. Finally, in each broad disciplinary category we matched authors with individuals in ministerial list using the information on their surnames, names and affiliation.

We dropped observations that, given the matching criteria, were assigned to more than one possible match, in this way we obtained a one-to-one correspondence between academics and Scopus Author-IDs.

After filtering, duplicates and incomplete records were deleted obtaining a consistent database of 449.664 scientific publications with at least one Italian author. We then employ a matching procedure to assign the corresponding author identifying codes to each research product (it might be possible that one paper is co-authored by two or more different individuals belonging to Italian academia). Figure 1.8 shows an increasing growth rate for the number of publications over the three decades considered.

The overall percentage of academics with at least one Scopus publication is 63.5% (see Table 1.6), and it is higher for associate professors (66%). Almost 50% of the publications are produced by academics holding an assistant professor position, one third by associate professors and the remaining portion by full professors (around 17% of the publications). Even if the overall production of papers is so diverse, on average assistant professors and full professors share the same level of productivity in the time period considered (29 papers), while academics at the associate professor

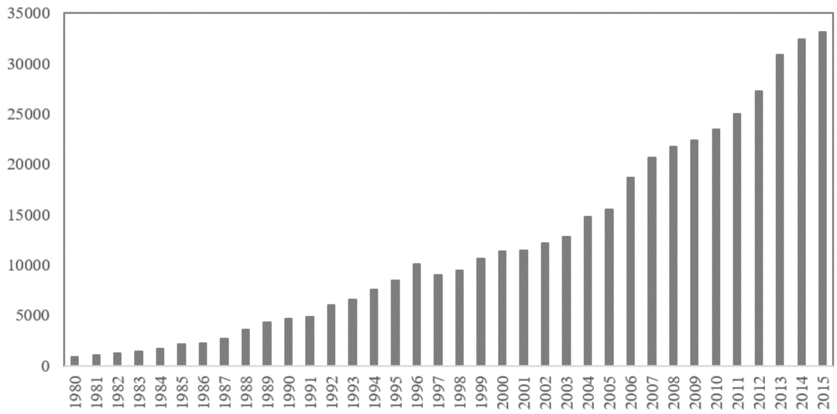

FIGURE 1.8: Number of Scopus publications

TABLE 1.6: Publications, academics and authors on Scopus by academic position

	# Publ.	# Acad.	# Auth.	% Auth.	Avg. Publ.
Assistant prof.	221,031	12,299	7,544	0.61	29.31
Associate prof.	147,811	11,209	7,418	0.66	19.92
Full prof.	77,955	4,325	2,697	0.62	28.90

level exhibit a lower scientific production (20 papers).

Table 1.7 shows the number of publication in our database per scientific area. Scientific production exhibits a high variety across fields, with Medicine and Engineering which account for 45% of the total publications. The percentage of academics who are also authors in Scopus (with at least one publication in international scientific journals) is on average 63.5%, but it varies across fields being low for Social Sciences and Humanities and higher for Medicine, Engineering and Science fields. Considering just the academics with at least one Scopus publication, on average in

TABLE 1.7: Publications, academics and authors on Scopus by scientific field

Field	# Publ.	# Acad.	# Auth.	% Auth.	Avg. Publ.
Mat	24610	1581	1346	85.14%	18.28
Fis	43310	1163	770	66.21%	56.25
Chem	64372	1753	1572	89.67%	40.95
Geo	14231	637	561	88.07%	25.37
Bio	72490	2558	2163	84.56%	33.51
Med	90096	2534	1925	75.97%	46.80
Agr&Vet	28573	1668	1437	86.15%	19.88
Arch	13938	2097	1145	54.60%	12.17
Eng	69499	3399	2975	87.53%	23.36
Hum&Law	14512	7106	1768	24.88%	8.21
Econ&Stat	12673	2400	1623	67.63%	7.81
Socsci	1359	914	374	40.92%	3.63
Tot	449663	27810	17659	63.50%	25.46

the period considered they publish 25 papers, the most prolific academics being in the field of Physics (56 papers), Medicine and Chemistry (47 and 41 paper, respectively).

1.3 Matching Italian doctorate holders and academics

The identification of academics who hold an Italian doctoral degree was pursued through the match between academics from the MIUR data with doctorate holders from Italian universities data from BNCF. The matching procedure consisted in the following steps.

- A "narrow" matching based upon academics' and doctorates' full name, gender, discipline and year of PhD.
- A "broad" matching directed at the academics and thesis that escaped the first-step matching, by first name (i.e with middle names excluded), gender, discipline and year of PhD.
- A "filtering out" procedure aimed at eliminating incongruous academic-doctorate matches.

In order to take into account typos in names in either one of the two datasets, we allowed a certain degree of flexibility in the matching. We treated names as raw strings and tested how many operations (inserting/deleting a character, switching two characters next to each other) need to be applied to the first string so that it is converted to the second. Once specified approve and disapprove levels, the score of the function is calculated using equation 1.1: *strA* is the first string, *strB* is the second string, *e(strA,strB)* is the edit distance between the two strings, *a* is the approve level and *d* is the disapprove level. Table 1.8 presents some examples.

$$score=\begin{cases} 0,\ if\ e\left(strA,strB\right)>d{*}\max\left[length\left(strA\right),lenght\left(strB\right)\right] \\ 1,\ if\ e\left(strA,strB\right)<a{*}\max\left[length\left(strA\right),lenght\left(strB\right)\right] \\ \frac{d{*}\max[length(strA),\ lenght(strB)]\ -e(strA,strB)}{(d-a){*}\max[length(strA),\ lenght(strB)]},\ otherwise \end{cases} \tag{1.1}$$

TABLE 1.8: Examples of score values for name matching

First value	Second value	e(strA,strB)	Approve level	Disapprove level	Score
Falini Andrea	Fallini Andrea	1	0.0	0.3	0.78
Maliocco Giovanni	Maliocco Giovanna	1	0.0	0.3	0.82
Piva Paolo	Pizza Paolo	2	0.0	0.3	0.44
Andrea Verdini	Andera Vedrini	2	0.0	0.3	0.83

To prevent the incongruent matching of male academic with female doctorate, and *vice versa*, with similar names (for example: Paolo and Paola, or names which in Italian could be either masculine or feminine, such as Andrea) we also took into account the gender. The MIUR data already contained the gender field for the records, while for the BNCF we downloaded a list of common Italian names-gender pairs, and associated them to the records.

We grouped academics and doctorates in six disciplines: Medicine&Veterinary, Science (agriculture, biology, chemistry, physics, geology and mathematics), Architecture&Engineering, Humanities&Law, Social Sciences and Economics&Statistics. We grouped the first three and the last three in 2 "scientific proximity clusters" and allowed the discrepancy between the discipline of the PhD and the broad scientific area of the academic giving a higher score if this discrepancy falls into one of the two clusters (Figure 1.9).

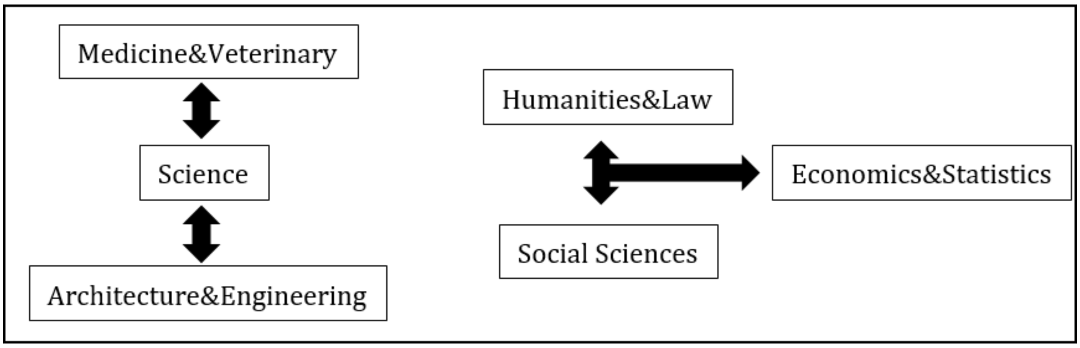

FIGURE 1.9: Scientific proximity clusters for the record linkage

Finally, we matched the year of publication of the thesis in the BNCF repository with the expected PhD year of the academics based on their birth year, allowing a time window between the 26^{th} and the 32^{nd} year for the academics. The final score is calculated using the formula 1.2. Table 1.9 presents some examples

TABLE 1.9: Examples of score values for year matching

First value	Second value	Lower bound	Upper bound	Score
1996	1999	1995 (value-1)	2001 (value+5)	0.5
1996	1997	1995 (value-1)	2001 (value+5)	0.83
1996	1994	1995 (value-1)	2001 (value+5)	0.0
1996	2001	1995 (value-1)	2001 (value+5)	0.17

$$score=\begin{cases} \frac{upperBound(v1)-v2}{upperBound(v1)-v1},\ if\ v2\in[v1;upperBound\,(v1)] \\ \frac{v2-lowerBound(v1)}{v1-lowerBound(v1)},\ if\ v2\in[lowerBound\,(v1)\,;v1] \\ 0,\ otherwise \end{cases} \quad (1.2)$$

We gave each field a weight (see Figure 1.10), two records were linked if the final score was 81 or higher. If a record from one dataset was linked to more than one record of the other, only the link with the higher score was retained.

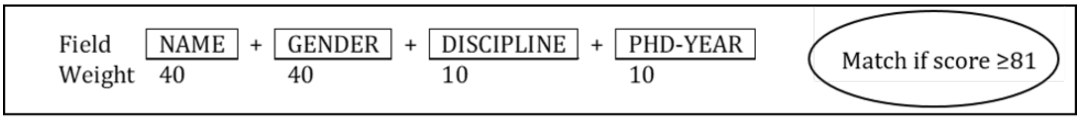

FIGURE 1.10: Weight formula for the record linkage

The record linkage was conducted between the database of doctoral dissertations, considering just those discussed in the period 1986-2006 (74,037 thesis) and taking into account all disciplines, and the one of academics. For the latter, since we did not have the year of PhD, we considered just those researchers who could have earned a doctoral degree in the period 1986-2006 based on their birth year (37,096 academics). We took into account also in this case all possible scientific fields focusing on those reported for the entry year in academia (the 4.4% of academics, in fact, changed their disciplinary scientific sector during the career).

In order to take into account for specificities of the Italian language, the following steps were undertaken to prepare the data:

- All the accents from the final vocal of the names were transformed in apostrophe, while the accents on the letters inside name were removed. For example: "FOÀ Sergio" became "FOA' Sergio" and "CALDÉS PINILLA Maria" became "CALDES PINILLA Maria".
- Whenever possible, composite names which had been recorded in the BNCF repository with only the initial, were manually corrected to the full length

name by manually searching online the person (e.g. “M. Angela” was changed to “Maria Angela”).

- All non-alphanumeric characters (e.g. the character “-” between composite names) were removed.
- The original “surname+firstname+middlenames” strings were placed in separate fields for each element, i.e. “surname”, “firstname+middlenames” and just “firstname”.

As said, we performed the matching relying on four fields: name, gender, scientific area and year of PhD. We first considered the complete names of the researchers and then we launched a second matching on the residuals records considering just the first names; we allowed a mismatch between the scientific area of the PhD and the one of the academic position; we considered a 5 year time window between the publication of the thesis and the expected year of dissertation based on researchers' birth year. At the end of the record linkage procedure we were able to identify 27,840 records: the 38% of the thesis discussed and the 75% of the academics who got their PhD in the period considered. Further information about the validation and summary statistics of the resulting database can be found in Appendix A.

1.4 PhD holders' career: descriptive statistics and stylized facts

Research in the Economics of Science field devoted major attention to the determinants of scientific productivity, while the analysis of career and mobility has not been the main focus, perhaps because both are assumed to be closely linked to productivity (Allison and Long 1990; Long et al. 1993). However, while one may expect that academic career advancement is made merely on the basis of scientific merit, there is conclusive evidence that this Is not the case. In many countries merit is not the only driver behind promotion, Long et al., 1993 show that seniority and gender are equally if not more important; similarly, since the seminal work of Crane (1965) and Crane (1970) it has been shown that hiring can rely on prestige effect, favouring graduates from top institutions; same can be said about inbreeding, when an institution hires researchers among its own graduates (see Horta et al. 2010).

Most of the existing literature on academic career paths refers to Anglo-Saxon countries, which have a very specific institutional set-up of the academic labor market. As continental Europe does not share this institutional set-up, these findings may not apply, but studies on these countries are still relatively scarce. Notable exceptions include Pezzoni et al., 2012 who study career advancement of academics in physics and chemistry in France and Italy; and Combes et al., 2008 who focus on publication profiles and network connections in the hiring procedure for economists in France.

The database presented allows to shed light on the Italian academic job market for PhDs. In what follows we provide some statistics about hiring and promotion patterns in Italian academia, focusing on the three main areas which seem to be the most studied in the literature: seniority and gender; mobility and inbreeding; university prestige.

1.4.1 Seniority and gender

In academics jobs, career progress is matter of time: seniority is rewarded with promotion. The time spent by a scientist at a given academic rank is found to be one of the most important factors determining chances of promotion (Long et al. 1993; Modena et al. 1999), either directly (more senior researchers stand higher chances of being promoted, *ceteris paribus*) or indirectly, via scientific production (more senior scientists stand a chance of accumulating a longer list of publications, which may be of help in getting promoted).

From Table 1.10, taking into account first appointment and tenure in Italian universities, it is possible to notice that academics who hold an Italian doctorate degree are slightly younger at both career stages (with the exception of Medicine&Veterinary) and that among these, also the percentage of tenured researchers is higher.

In their seminal study, Long et al., 1993 also highlighted the gender gap for female scientists, who exhibits a lower change of getting promoted controlling for productivity. Gender differences in productivity and career path have attracted the

TABLE 1.10: Age at first appointment and tenure and percentage tenured by scientific area and PhD

	PhD ITA			Others		
	Age First App.	% Tenured	Age Tenure	Age First App.	% Tenured	Age Tenure
Science	34.4	49%	41.5	34.5	50%	40.8
Arch&Eng	34.1	59%	40.1	34.8	52%	40.2
Med&Vet	36.9	41%	43.5	36.6	37%	43.0
Hum&Law	35.7	53%	40.8	36.4	48%	41.1
Econ&Stat	33.2	62%	38.5	34.3	58%	38.9
SocSci	36.8	43%	42.2	37.8	37%	42.4

TABLE 1.11: Gender by scientific area and PhD

	PhD ITA		Others	
	M	F	M	F
Science	55%	45%	63%	37%
Arch&Eng	75%	25%	78%	22%
Med&Vet	51%	49%	66%	34%
Hum&Law	51%	49%	52%	48%
Econ&Stat	60%	40%	66%	34%
SocSci	58%	42%	59%	41%

attention of both sociologists and economists (see for example Kahn 1993; Levin and Stephan 1998; McDowell et al. 2001; Ginther and Kahn 2004).

Table 1.11 provides information about the gender composition of Italian universities by scientific field and PhD. Social Sciences and Humanities&Law are the fields in which the gender gap is lower (2 and 6 percentage points, respectively) while Science and Architecture&Engineering are notably more gendered. It is interesting to notice, comparing the academics with an Italian doctorate to the others, that gender gap in the fields of Science and Medicine&Veterinary is milder for the first group (17 and 30 percentage points smaller respectively).

1.4.2 Inbreeding and mobility

Academic inbreeding was very common in the United States until the late 1970s (Hargens and Farr 1973), and remains substantial in many countries in Europe, at least at the beginning of the academic career (Horta 2013; Horta et al. 2010). Godechot, 2016 has shown that in France during the 1980s, inbred PhDs were 17 times more likely to get hired than outbred PhDs. Moreover, also the relationship between academic inbreeding and scientific performance has been examined. Most such studies have shown, usually through a university of origin fixed effect, that inbred scholars are less productive scientifically (Horta 2013; Horta et al. 2010).

TABLE 1.12: Inbreeding by scientific area

	First App.	Tenure	Pure Inbreds
Science	54%	55%	96%
Arch&Eng	54%	53%	95%
Med&Vet	52%	52%	96%
Hum&Law	31%	29%	94%
Econ&Stat	30%	27%	86%
SocSci	30%	25%	94%

Note: only PhD ITA=1; percentage of pure inbreds calculated over the tenured inbreds.

Table 2.2 presents statistics about inbreeding, focusing only on academics who earned the doctorate degree from an Italian institution. Inbreeding at first appointment and tenure is considerably higher in Science, Architecture&Engineering and Medicine&Veterinary fields, where the phenomenon interests more than half of the faculty hirings from Italian universities. However, in Social Sciences, Economics&Statistics and Humanities&Law the percentage of inbreds is lower, around 30% for both career stages. Among tenured inbreds, the percentage of academics who never changed workplace through their entire career is around 95% in all disciplines, with the exception of Economics&Statistics where it is 10 percentage point lower.

On the other hand, mobility is an important characteristic because it enhances knowledge circulation and contributes to the well-functioning of the academic market. Its impact on career advancement, however, seems to depend on the general structure of the academic system. Mobility, in fact, can have a positive effect on academic career since mobile researchers have access to a bigger network of acquaintances which can give them the opportunity of increasing their scientific productivity (Jonkers, 2011). On the other hand, mobility can also have negative effects on career since mobile researchers could experience more difficulties in integrating with local environments (Melin, 2005), especially in case of young scientists in their post-doctoral period (Cruz-Castro and Sanz-Menéndez, 2010).

In Table 1.13 and Table 1.14 we report statistics about mobility and promotion patterns. In the first table we see whether changes of affiliations at the assistant and associate professor rank are related to promotion, while in the second table we see whether researchers promoted to associate and full professor rank are internal candidates or coming from a different institution. From Table 1.13 it is possible to

TABLE 1.13: Mobility and promotion by scientific area and PhD

		PhD ITA		Others	
		Prom	No prom	Prom	No prom
Assistant's move	Science	47%	53%	49%	51%
	Arch&Eng	68%	32%	76%	24%
	Med&Vet	42%	58%	46%	54%
	Hum&Law	60%	40%	61%	39%
	Econ&Stat	61%	39%	71%	29%
	SocSci	56%	44%	40%	60%
Associate's move	Science	13%	87%	13%	87%
	Arch&Eng	9%	91%	12%	88%
	Med&Vet	13%	88%	14%	86%
	Hum&Law	13%	87%	18%	82%
	Econ&Stat	11%	89%	19%	81%
	SocSci	9%	91%	27%	73%

TABLE 1.14: Promotion and mobility by scientific area and PhD

		PhD ITA		Others	
		Move	No move	Move	No move
Prom. to Ass.	Science	8%	92%	12%	88%
	Arch&Eng	9%	91%	13%	87%
	Med&Vet	5%	95%	9%	91%
	Hum&Law	14%	86%	20%	80%
	Econ&Stat	17%	83%	23%	77%
	SocSci	9%	91%	7%	93%
Prom. to Full	Science	9%	91%	10%	90%
	Arch&Eng	6%	94%	8%	92%
	Med&Vet	8%	92%	10%	90%
	Hum&Law	13%	87%	19%	81%
	Econ&Stat	11%	89%	16%	84%
	SocSci	10%	90%	14%	86%

see that mobility at the assistant professor level is mostly related to promotion, especially in Architecture&Engineering and Economics&Statistics, while mobility at the associate professor level is almost entirely non-promotion related in all scientific fields. There are no striking differences comparing academics who hold an Italian PhD degree to others. From Table 1.14 it is possible to see that, in line with what just said, promotion to associate and full professor position is hardly related with researchers' change of institution, especially considering candidates with an Italian PhD degree with respect to the others.

1.4.3 University prestige

Many empirical studies show the importance of university prestige for a successful career. Graduating and working in a prestigious institution gives visibility

TABLE 1.15: Summary statistics of the hiring networks

	STEM			SSH		
	All	F	M	All	F	M
# Nodes	95	95	95	95	95	95
# Edges	709	643	955	1087	993	1184
Density	0,159	0,147	0,218	0,243	0,222	0,265
Avg path	1,921	1,957	1,814	1,8	1,823	1,792
Diameter	4	4	4	3	4	3
Avg clust. coef.	0,541	0,601	0,731	0,518	0,663	0,667

and access both to information and to knowledge embedded in other productive scientists, which makes promotion easier, especially if associated with mobility across universities (Long et al., 1993). Several studies find that the ranking of PhD education is one of the most useful predictors of success in academia, where success means getting a job in a highly-ranked academic institution, even after controlling for productivity (Hargens and Hagstrom, 1967). However, while the reputation of the university granting the doctoral degree seems to occupy a relevant position in success in the US academic job market (Allison and Long, 1990), in the case of Europe the evidence in favor of an impact of doctoral institution prestige on careers is much weaker (Heining et al. 2007 found some limited impact studying German economists).

We performed an exploratory network analysis for the hiring patterns of PhDs in Italian universities, following the work of Clauset et al. (2015) for the US. We consider separately science engineering technology and mathematics (STEM) doctorate holders and social sciences and humanities (SSH) ones. We consider years from 1990 to 2015.

The hiring networks of STEM and SSH are two weighted and directed adjacency matrices. Each matrix M has 95 rows and columns that are the 95 Italian universities, where each entry m_{ij} represents the number of scholars with a PhD from university i and a subsequent first job in university j.

Table 1.15 shows summary statistics of the hiring network for all, male and female scholars for SET and SSH. Both macro-fields show a common pattern for females: their networks are sparser than those of males. That is, they have fewer edges and so lower density, but also higher average path length, and a lower clustering coefficient.

The networks of STEM and SSH (Figure 1.11) show common hiring corridors, where universities with high production of PhDs tend to hire more (nodes with high in-degree tend to have high out-degree) and universities with high hiring and low placement are situated in the Center and South of Italy, with the exceptions of Rome and Naples.

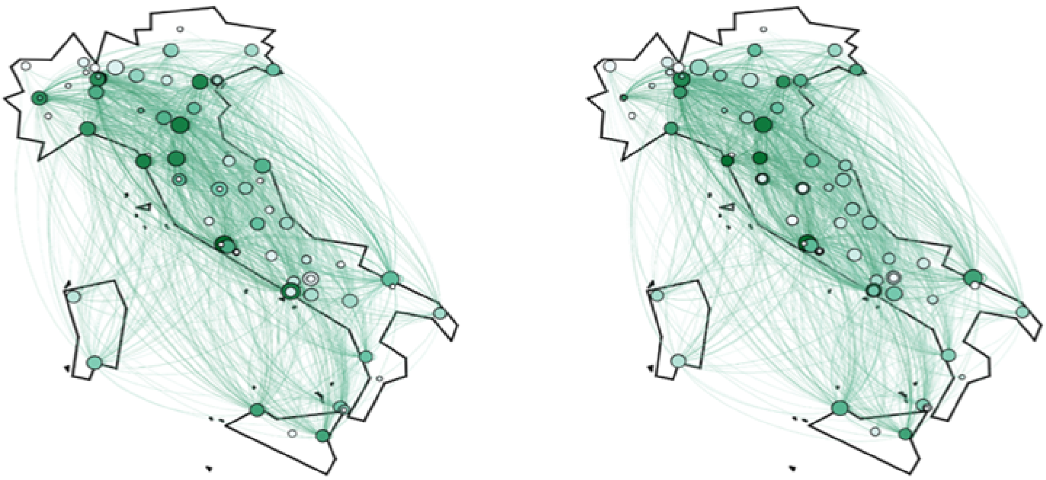

FIGURE 1.11: Hiring networks (left: STEM; right: SSH). Vertex size: in-degree; vertex color: out-degree

Table 1.16 and Table 1.17 witness the so-called "institutional stratification hypothesis" for STEM and SSH, respectively. They show the number of PhDs from the top 10 prestige universities[5] hired within the other top 10 institutions, and those hired in others. The results are striking. In STEM the top 10 prestige universities produce the 42% of all scholars in the country, while for SSH the percentage is 33. Among those, in the field of hard sciences, the 81% find a first job within these 10 institutions and 67% in SSH. This underlines the crucial role of prestige hierarchies in academia. Moreover, the lower percentage of the first job placement of the top universities in SSH with respect to SET highlights the diverse hiring processes of the two fields.

Finally, prestige rank-change measures the movements in prestige of individuals from the PhD to the first job. The measure is simply the difference between the prestige ranking of the university where a person obtained his PhD and the one of his first job. So each young scholar can move in the hierarchy in three ways:

[5]To select the top universities for both academic macro-areas, we used the number of departments of excellence selected by MIUR.

TABLE 1.16: STEM PhDs hired from top 10 universities

University	Hiring all	%	Hiring top 10	% Top 10
PADOVA	594	4,50%	489	82,32%
BOLOGNA	673	5,10%	539	80,09%
TORINO	382	2,90%	300	78,53%
ROMA La Sapienza	766	5,81%	597	77,94%
NAPOLI Federico II	828	6,28%	672	81,16%
MILANO	484	3,67%	385	79,55%
FERRARA	221	1,68%	161	72,85%
Politecnico di MILANO	789	5,98%	685	86,82%
PERUGIA	253	1,92%	190	75,10%
Politecnico di TORINO	556	4,22%	481	86,51%
Total	5546	42,05%	4499	81,12%

TABLE 1.17: SSH PhDs hired from top 10 universities

University	Hiring all	%	Hiring top 10	% Top 10
BOLOGNA	636	6,12%	452	71,07%
FIRENZE	299	2,88%	207	69,23%
PADOVA	384	3,70%	307	79,95%
TORINO	480	4,62%	298	62,08%
MILANO	362	3,48%	238	65,75%
Ca' Foscari VENEZIA	184	1,77%	113	61,41%
SIENA	162	1,56%	85	52,47%
MILANO-BICOCCA	213	2,05%	137	64,32%
TRENTO	168	1,62%	118	70,24%
ROMA La Sapienza	571	5,50%	378	66,20%
Total	3459	33,29%	2333	67,45%

- *Up*: if his prestige rank-change is positive. He gets his first job in an institution more prestigious then his PhD;
- *Down*: if his prestige rank-change is negative. He gets his first job in an institution less prestigious then his PhD;
- *Stay*: if his prestige rank-change is zero. In practice, that means he is hired by the institution granting his PhD.

Figure 1.12 shows the distributions of rank-change in relative terms for STEM and SSH where we compare female vs. male. It is possible to observe that in STEM males tend to move up the hierarchy, as the distributions peaks toward positive numbers while the opposite is true for females in SSH.

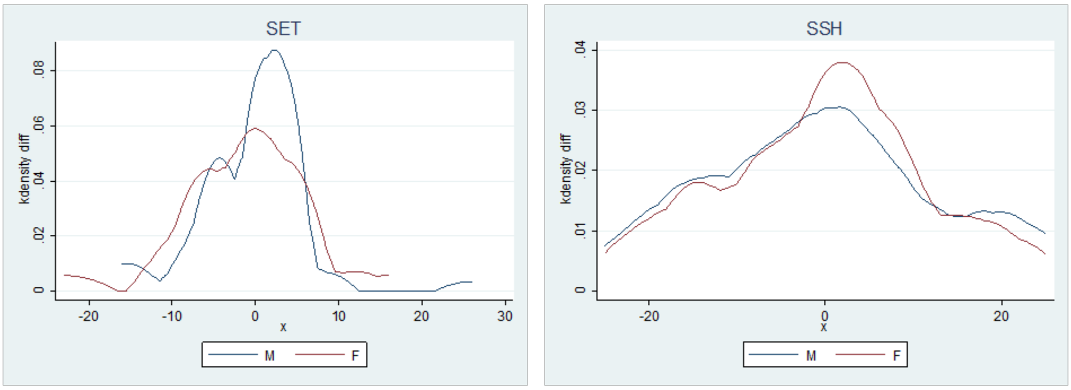

FIGURE 1.12: Distribution of rank changes (left: STEM; right: SSH)

1.5 Conclusions

Doctorate holders are central actors in the creation and dissemination of innovative scientific knowledge. The analysis of their career progression is important both to understand the academic labor market functioning and the returns for the academic system of the public resources invested in their third level education.

In this work we have reviewed the existing evidence on the phenomenon, with an eye mainly on methodological issues. In particular, we have pointed out that data used in the (scarce) literature mainly come from ad hoc surveys which cover, at times, only a limited number of institutions, years or scientific fields. Hence the need of a comprehensive database, as the one created and presented here, with the potential of shedding light on several issues.

We have then provided some examples of this potential by producing descriptive statistics and exploratory analysis based on the database constructed, a pilot dataset on Italian doctorate holders' information extracted from the BNCF repository and ministerial data on academics hired in Italian universities from MIUR. We have discussed at length the technical issues of name disambiguation and record linkage which was a crucial one for ensuring the adequate level of data quality.

The results picture on Italian academic careers presented, albeit stylized and preliminary, confirm the database potential. First, we have seen that gender gap among Italian university professors, a well-known phenomenon especially in hard science fields, is milder among academics who got the doctorate in Italy. For this group, career progression seems to take place faster at both first appointment and tenure. We also saw that the magnitude of inbred faculty is much higher in hard sciences rather then in social sciences. In fact, promotion to associate and full professor position is hardly related with researchers' mobility, especially considering candidates with an Italian PhD degree. Considering the faculty hired from the top 10 universities for STEM and SSH scientific macro-areas, the figures are striking: the 81% of the hirings in STEM (and the 67% in SSH) come from the group of top universities itself.

Chapter 2

Early Career Decisions and Advancement in Italian Academia

2.1 Introduction

The economics of the academic labor market, and higher education economics in general, has become a rapidly expanding field, see for example the symposium of the Journal of Economic Perspectives (Clotfelter, 1999). The interest in this field relies in two main reasons. First, higher education is a key industry in the production of human capital, and thus the organization of universities is of crucial importance for modern economies; and second, the academic labor market presents features that are uncommon in other markets, such as the difficulty of observing output and the institutional setting of tenure.

In this field, literature is dominated by studies which focus on the situation of the academic labor marker in the United States. The US universities not only are very numerous but are also very successful in various terms: research output, number of patents, Nobel laureates. These studies often focus on the effect of tenure (McPherson and Schapiro, 1999; Ehrenberg and Zhang, 2005) and the changing in salaries within private and public universities (Ehrenberg et al., 1998; Zoghi, 2003).

However, the university system and the related academic labor market in continental Europe is characterized by very different institutional settings to be comparable with the US one. In Europe, for example, the quality of university is less heterogeneous and education institutions are highly regulated, which is due to the fact that the majority of universities are public and thus academics are civil servants with a salary determined by law. For these reasons the insights that are coming from research on the US academic system may not be appropriate for the European one.

This work adds to the existing literature by studying the career path and career opportunities of academic economists in Italy. The objective of this study is to investigate the relationship between early career choices of researchers and tenure and permanent academic positions. The access to an associate professor position or permanent lifetime employment is probably the key reward in university career. Non-tenured researchers have a limited amount of time to show their worth so they have to take their decisions carefully especially in the beginning of their career, during their PhD and post-doc years. In particular I will argue that mobility decisions, early scientific productivity and the type of relationship with the own supervisor built during the years of the doctorate have a strong influence on career success.

Of course, the relationships among these three elements are expected to be diverse in different institutional contexts. In this sense Italy is a well fit case study: the Italian academic system, in fact, has been object of a recent reform which introduced a two-step procedure (national scientific habilitation plus local concourses) to recruit and promote professors, as opposed to the previous system where these decisions were entirely locally made by universities.

I am particularly interested in the effect of the type of relationship between advisor and doctorate in career advancement. Academic organizations, in fact, face the dilemma between the mobility and the loyalty of their best researchers by developing strategies to expand their scientific network and creating opportunities to reward their commitment. In this sense, understanding the timing of the transition to a first permanent position and tenure is essential in order to know how departments are able to commit promising candidates to the organization.

The event of interest of the analysis is career advancement from PhD to a permanent position in Italian academia, and from the latter to associate or full professor (or the direct entrance in the academic system in one of these positions). The aim is to approach some of the dynamics of research careers in early stages by analyzing the relationship among social capital, mobility and scientific production. The main question I want to address is: "What does explain the timing of academic career progression?" testing the relative impact of productivity, mobility and other early career variables.

Previous studies on academic promotion have been dominated by cross-sectional design to explain the success in the academic labor market, and when they used longitudinal information, these latter was coming from surveys questions, which are vulnerable to bias. Additionally, the findings of prior research are far from being clear in identifying the factors that actually determine success in academia. In particular, recent literature highlights the need to investigate academic career selecting from PhD cohorts (Jungbauer-Gans and Gross, 2013) and following careers from the earliest stages onwards (Lutter and Schröder, 2016).

This study contributes to prior research in several ways. In this paper I use a novel dataset based on institutional longitudinal data and survival analysis (event history analysis) to explore time to promotion and its covariates. In particular I seek to deepen the understanding about the role of early academic performance, mobility and social embeddedness in the period between PhD and tenure, a point that, as said above, is still unclear.

2.2 Conceptual framework

Most of the existing literature on academic career paths refers to Anglo–Saxon countries, which have a very specific institutional set-up of the academic labor market. As continental Europe does not share this institutional set-up, these findings may not apply, but studies on European countries are still relatively scarce. Notable exceptions include Pezzoni et al. (2012), who study career advancement of academics in physics and chemistry in France and Italy, and (Combes et al., 2008), who focus on publication profiles and network connections in the hiring procedure for economists in France. Furthermore, few studies have examined the issue of academic promotion in the case of the European academic systems focusing on researchers' early career stage. Bäker (2015), studies German-speaking researchers in economics and management finding a short-term negative effect of early job mobility on research output; Sanz-Menéndez et al. (2013) surveyed science, biomedical and engineering researchers in Spain focusing on their time to promotion finding that the academic system privileges social embedded scientists rather than mobile ones. In this section, I summarize the existing evidence on the various determinants of academic careers. From the discussion, I will derive a number of expectations to be tested by the regression analysis that follows.

Seniority In Academic jobs careers, progress is matter of time: seniority is rewarded with promotion. The time spent by a scientist at a given academic rank is found to be one of the most important factors determining chances of promotion (Long et al., 1993; Modena et al., 1999), either directly (more senior researchers stand higher chances of being promoted, *ceteris paribus*) or indirectly, via scientific production (more senior scientists stand a chance of accumulating a longer list of publications, which may be of help in getting promoted, see next paragraph).

Scientific productivity Among individual characteristics, the scientific literature has studied research productivity since the seventies both in terms of number of publications (see Clemente, 1973) and their quality, measured through citations (see Hargens and Farr, 1973), almost unanimously finding that they are positively associated with promotion. The role of early publication, instead, is more controversial. On the one hand, early work by Allison and Stewart (1974) confirmed the hypothesis, already advanced by Clemente (1973), that early publication is a powerful predictor of advancements in ranks; but this result is not confirmed by more recent works such as Long et al., 1993.

Gender In their seminal study, Long et al. (1993) also highlighted the gender gap for female scientists, who exhibits a lower change of getting promoted controlling for productivity. Gender differences in productivity and career path have attracted the attention of both sociologists and economists (e.g. Kahn, 1993; Levin and Stephan, 1998; McDowell et al., 2001; Ginther and Kahn, 2004).

University prestige Many empirical studies show the importance of university prestige for a successful career. Graduating and working in a prestigious institution gives visibility and access both to information and to knowledge, embedded in other productive scientists, which makes promotion easier, especially if associated with mobility across universities (Long et al., 1993). Several studies find that the ranking of the PhD university is one of the most useful predictors of success in academia, where success means getting a job in a highly-ranked academic institution, even after controlling for productivity (Hargens and Hagstrom, 1967). However, while the reputation of the university granting the doctoral degree seems to occupy a relevant position in success in the US academic job market (Allison and Long, 1990), in the case of Europe the evidence in favor of an impact of doctoral institutions prestige on careers is much weaker (Heining et al., 2007 found some limited impact studying German economists).

Mentorship Several studies find that the commencement and advancement of an academic career seems to correlate more with the productivity and prestige of the mentor and that of the doctoral department than with indicators of individual scientific productivity (Long et al., 1979; Reskin, 1979; Long and McGinnis, 1985). Recently, the literature has focuses on the direct role of sponsorship in promoting candidates in European state competitive exams taken upon entrance to an academic career, finding a strong positive impact on success probability when the applicant has a sponsor in the hiring committee itself (Combes et al., 2008; Zinovyeva and Bagues, 2015).

Inbreeding Academic inbreeding was very common in the United States until the late 1970s (Hargens and Farr, 1973), and remains substantial in many countries in Europe, at least at the beginning of the academic career (Horta, 2013; Horta et al., 2010). Godechot and Louvet (2008) have shown that in France during the 1980s, inbred PhDs were 17 times more likely to get hired than outbred PhDs. Moreover, also the relationship between academic inbreeding and scientific performance has been examined. Most such studies have shown, usually through a university of origin fixed effect, that inbred scholars are less productive scientifically (Horta, 2013; Horta et al., 2010).

Mobility Mobility is an important characteristic because it enhances knowledge circulation and contributes to the well-functioning of the academic market. Its impact on career advancement, however, seems to depend on the general structure of the academic system. Mobility, in fact, can have a positive effect on academic career since mobile researchers have access to a larger network of acquaintances which can give them the opportunity of increasing their scientific productivity (Jonkers, 2011). On the other hand, mobility can also have negative effects on career since mobile researchers could experience more difficulties in integrating with local environments (Melin, 2005), especially in case of young scientists in their post-doctoral period (Cruz-Castro and Sanz-Menéndez, 2010).

2.3 Institutional setup

The academic labor market in Italy is very different from other non-European countries, especially compared to the USA. The main feature is that this labor market is highly regulated. Since the employees at Italian universities are civil servants, the wage, the contract length, the tasks (teaching load and others) are determined by law and cannot be bargained at the local level. In order to understand the factors which determine success in this market it is crucial to understand these regulations.

The Italian academic system is composed of 97 universities (30 private and 67 public) and 9 higher education institutions. The latter usually dispense only masters and PhD courses, being more research oriented than most of the other universities. Three out of the 67 public universities are polytechnics. Eleven out of the 30 private institutions are distance-learning universities. The university system is divided into 372 sectors of discipline ("settori scientifico disciplinari" – SSD), grouped into fourteen research areas, as designated by the Italian National University Council (CUN). Sectors of discipline are categorized for homogeneity within each research area, and the selection of research candidates is conducted by recruitment commissions within each academic discipline in both national and local recruitment systems.

The Italian academic system has three main positions called "Ricercatore universitario" (assistant professor), "Professore associato" (associate professor) and "Professore ordinario" (full professor). Each professor working at an Italian university is then categorized by a level of arrangement (full professor, associate professor and assistant professor) and by one out of 372 sectors of discipline. Each vacancy is coded in a standardized format, and each filled position becomes tenured after a review conducted three years after hiring. Salaries in public universities are set by law and vary only by level of arrangement and seniority. Schools and departments are prevented from differentiating wages among professors, linking payment to research productivity and/or teaching loads. As a consequence, in addition to celebrity and funds attraction, the strongest incentive to scientific productivity for individuals working in academia derives from expected promotion (being hired as assistant professor, being promoted to associated or full professor). Given the public nature of the employment contracts, university professors can only be hired through public competitions that should grant publicity of the vacancy, selection of the selecting committee based on objective criteria, transparency of the selection process.

Since 1979, standardized competitions were held to hire assistants, associate and full professors and until 1998 almost all academic recruitment was substantially centralized. Despite the legislative prescription of one concourse every two years, a three to four years interval occurred. National commissions of five members were chosen by lot within a pool of elected professors (from a pool of 15[1]) belonging to the same discipline. Commissioners declared which of the candidates had the qualifications to be promoted to associate/full professorship. Eligibility was given to a number of candidates greater than the available positions (usually 20% higher) for each discipline. Universities with opening positions drew by multilateral bargaining between them from the list of eligible applicants to fulfil their vacancies. Starting in 1999[2], recruitment procedures became entirely local, and each university could hold its own selection procedure (both for assistants, associates and full professors). Local commissions were comprised of five members: one belonging to the institution itself and the four others elected by the full set of Italian professors of that discipline. After 2005, a new reform act[3] established that the commission's members had to be drawn by lot in a pool of professors of three times the size of the local commission, elected by popular vote amongst the discipline's affiliates. The commissions initially declared three qualified candidates for each concourse, but moved to two between 2007 until 2008, and only one thereafter. In the following years, universities with open vacancies could hire any candidate who had obtained a qualification.

[1]See DL 31/1979 and DPR 382/1980.
[2]"Berlinguer reform", DPR 390/1998.
[3]"Moratti reform", DL 230/2005.

The most recent reform[4] (in 2010) introduced a two-step procedure with a national habilitation and local concourses to recruit professors.

I do not have data on concourses (nor on commissions, candidates, winners or "*idonei*") in period I am considering, but I can observe entrances and promotions (from assistant professor to associate and from associate to full professor). Since concourses results do not translate immediately in new hirings, but the universities have to call the winners or "*idonei*" to fill their vacancies, I can interpret promotions and new entrances as the results of concourses. Figure 2.1 shows the number of new assistant professors and newly promoted associate and full professors of economics and statistic in Italian academia.

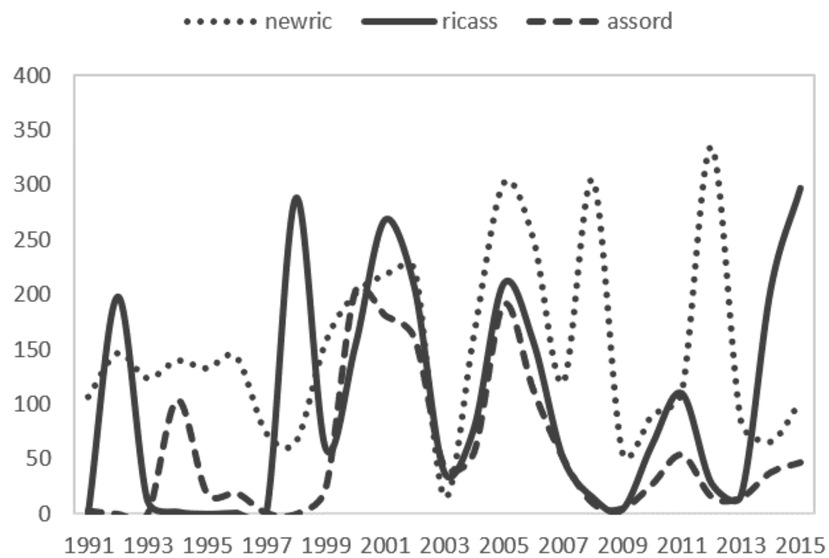

FIGURE 2.1: Yearly entrances and promotions in Italian academia, 1990-2015

2.4 Expectations

Based on the theoretical literature and the institutional setup, I develop a number of expectations to be explored in the empirical analysis. Of course, the null hypothesis in merit–based societies is that applicants displaying the highest achievement are to be rewarded with the most desirable positions, i.e. that publication output does indeed matter for getting tenure.

[4]"Gelmini reform", L 240/2010.

- *Mentorship* — Direct mentoring should have no role in candidates' chance of success, since no reference letters are admitted into the evaluation process. I expect the indirect influence of the academic supervisor to affect the odds for tenure success. I focus on two different roles that the supervisor could have embodied for his student: *collaborator*(co-authorship with own student) or *sponsor*(giving access to his network to find a job). The potential effect of co-authoring with the own supervisor could be interpreted in two opposite ways. On the one hand, the research collaboration with the PhD supervisor could increase the visibility of the work of the researcher, send signals to colleagues that he is integrated in the scientific community and thus play a positive role in the academic career. On the other side, this type of co-authorship could also have an opposite effect on researchers' career advancement. In fact, one could argue that this relationship is unbalanced towards the supervisor who, thanks to his major experience, is probably giving the most significant contribute to the research, while the younger scientist is just doing the "dirty work". Similar reasoning could be applied to the *sponsor* role. A supervisor is supposed to know better the "unobservable habits of action" (Baruffaldi et al., 2016) of his PhD students and thus can help them in their early career stage by giving them access to his network of acquaintances and previous affiliations; the downward side is that the researcher, rather than develop his own scientific network, can remain anchored to his doctorate supervisor's.

- *Seniority* — Length of career, in principle, should not affect rank advancement chances, since seniority does not enter the examination criteria of concourses.

- *PhD years* — The egalitarian norms typical of the Italian legislation forbid the consideration of university prestige as a factor to be evaluated by the examining commissions: all PhD titles have to be considered equal, since no official ranking of universities exists, and no unofficial information can be deemed as relevant by examiners. Nonetheless, I argue that early publications, during or immediately after the PhD, are the most important for a researcher who wants to pursue academic career, because they allow him to prove his worth already in his junior years.

- *Gender, Age/Cohort* — At higher age, researcher find it more costly to get involved in the scientific community (see Fitzenberger and Schulze, 2014 for a similar argument). Also, older cohorts are less likely to be affected by the university reforms. Furthermore, the higher the age, the more likely the PostDoc may not have been considered for a professorship, resulting in older PostDocs being more negatively selected. As for gender, no apparent reason exists to think of peculiarities for France and Italy with respect to the US-based evidence at hand. However, a quick look at available statistics show women are underrepresented at the higher levels of the academic ladder.

- *Academic fields* — In Italy disciplines compete for resources, which allow for cross-disciplinary differences in the availability of new jobs and promotion opportunities. That's why it is important to take in to account the average probability of being promoted in a given discipline.

- *Inbreeding & Mobility* — Inbreeding, in principle, should play no role in academic recruitment, but universities face the dilemma between the mobility or the loyalty of their best researchers by developing strategies to expand their scientific network and creating opportunities to reward their commitment.

2.5 Data

2.5.1 Data generation

I test the above mentioned expectations using event history analysis on a novel panel dataset that covers full career profiles of a sample of economists employed at Italian universities.

I have collected information from three primary sources: the national library of Florence (BNCF), the Italian Ministry of Education and Research (MIUR) and the bibliographic web version of Scopus databases.

- *BNCF* — Since the DPR 382/1980 all the doctoral dissertations discussed in Italian universities by law have to be deposited at the BNCF repository. I observe 73,950 such dissertations for the period 1986-2006, retrieving all the information available from the BNCF on-line public access catalog, which provides title of the thesis, author, supervisor, university, field and year of degree. Although the registration is compulsory, BNCF repository does not include information on approximately 3.5% on all dissertations discussed in Italy[5].

- *MIUR* — Information regarding academic positions, disciplinary areas, university affiliation and personal information, such as birth year and gender, are available on-line from 2000 to 2015. I obtained data on academic careers before 2000 from Cineca, a MIUR agency which collects administrative data on personnel as well as on competition for professorship in Italy.
- *Scopus* — Scopus, published by Elsevier, indexes a greater number of journals (12.850, including 500 open access journals) within the medicine, technical and social sciences and offers systematic, quantifiable statistical information based on citation data.

From MIUR I collected information about the career path of Italian assistant professors, associate professors and full professors in economics from 1990 to 2015. There are 3,710 academics left after excluding academics who:

- by year of birth were unlikely to have obtained a PhD in the period 1986-2006, and thus would not appear in the BNCF repository;
- already covered the position of full professor at the beginning of the observational period (i.e. in 1990);
- had homonyms colleagues (taking into account surname and initial of name) in the same university, in order to avoid mistakes in the association of the publications.

From BNCF I downloaded all the record concerning the doctoral thesis discussed in Italian universities from the first cycle of doctorate until 2006. I matched these data with those from MIUR and was able to identify 2,515 academics who pursued their PhD in Italy. Missing information may be due to the fact that:

- individuals defended their dissertation after 2006;

[5]See Chapter 1 for further information on the database and the retrieval process.

TABLE 2.1: Individual characteristics Sample vs Population

	Sample	Official data
Male (female)	58.2% (41.7%)	59.2% (40.7%)
Avg age at first app.	33.3	32.4
Avg age at tenure	39.4	38.6

– there are spelling mistakes either in the MIUR or in the BNCF data which do not allow the matching between the records;

– the dissertation was not included in the BNCF repository for unknown reasons.

I focus in this study on those individuals for whom information on the PhD advisor is available. This restriction is important since I need this information in order generate some of the control variables. Restricting to those individuals for whom I have all the relevant information leaves with 1,916 individuals, thus I lose 24% of observations. This is due to the fact that the older dissertations deposited at the BNCF (before 1993) do not report information on the supervision. Note that missing information is a standard problem in using publicly available data when doing research (see Dietz et al., 2000). In order to preserve representativeness I then further restricted the sample to thesis discussed in the period 1993-2006, which are 1,672 (67% of the linked records).

Using aggregate data on the characteristics of tenured professors, which is provided by MIUR, it is possible to be confident that the sample is representative of the population. Table 2.1 shows the proportion of men (women), the average age when getting the first job in the Italian academia and the average age at tenure from the sample and from the official data.

The comparison shows that the average professor in the sample is slightly older in terms of age at tenure and age at first job. Note that this bias might well be due to aggregation. In sum the values of the key characteristics in the sample and in the (aggregate) population data are reasonably similar.

2.5.2 Data description

Of the 1,672 researchers for whom I have full information, 698 are women (42% of the sample). Doctoral education provenience is quite spatially concentrated, the 43%

of the sample received the PhD degree from a university in the north of the country. The region with the higher number of doctorates is Lombardia. In Milan alone, in fact, there are six universities with economics departments which offer third level education courses. However, the university with the higher number of doctorates is located in the south: the 8% of the sample got the degree from the university of Naples Federico II.

As said, there are three main position in the Italian academic career: assistant professor, associate professor and full professor, I consider these two latter position together, as tenured position. Taking in to account the 59% of the sample who got tenured, the average professor finished her PhD at the age of 30 and got her first position as assistant professor in Italian academia at the age of 33, she was then granted tenure (associate or full professorship) at 39 years old. Note that not all professor in the sample follow this career path: 82 of them entered in the Italian academia directly as associate or full professors at the average age of 37 years.

Inbreeding

Academic inbreeding is the practice of universities hiring their own graduates. Previous studies on this phenomenon in the US academic system, showed that elite institutions are those in which this phenomenon is prevalent. I then identified the top 10% economics institutions in Italy according to the IDEAS ranking[6] and calculated the percentage of inbreds among the sample at the first job and at tenure. As Table 2.2 shows, inbreeding is not a prevalent phenomenon in Italy, however it tends to persist in top universities for tenured professors. Furthermore, if among inbred academics I consider only those who never moved from their alma mater since the start of their career, meaning that they have not gained any experience outside their PhD institution, I found that this percentage of "pure inbreds" Horta, 2013 is higher for non-elite universities.

[6] Further information at ideas.repec.org/top/top.italy.html. The top universities identified are: Bocconi, Bologna, Firenze, Brescia, Roma Tre, Ca' Foscari, Politecnica delle Marche, Cattolica del Sacro Cuore, Padova, Modena e Reggio Emilia, Roma Tor Vergata, Milano, Verona, Torino, Luiss Guido Carli and Milano-Bicocca.

TABLE 2.2: Percentage of inbreds by institution type

	First job	Tenure	Pure inbreds
Top Uni	39%	34%	75%
Other Uni	36%	26%	78%

TABLE 2.3: Promotion pattern of mobile academics

	Promoted	Not promoted
Assistant's move	132/227 (58%)	95/227 (42%)
Associate's move	18/71 (25%)	53/71 (75%)

Mobility paths

Taking into consideration the period after the first appointment in Italian academia (i.e. at least as assistant professor) I measure the job mobility of the academics in the sample. The group of immobile academics constitute 85% of the sample, indicating a very low degree of mobility among university professors in Italy, in line with what found by previous studies (see for example Pezzoni et al., 2012).

Table 2.3 shows that of those mobile at the assistant professor level, 58% moved to be promoted to a higher rank; while among those mobile at the associate professor level, just 25% moved to gain the position of full professor. Thus in Italy mobility is linked to promotion opportunity just at the lowest academic rank.

Table 2.4 reports the mobility pattern of researchers who gain a tenured position in Italian academia. It shows that overall 13% of new associate professors moved from another Italian university to gain promotion and 8% came from outside Italian academia; at the full professor level, 9% moved to get promotion and the 4% came from outside the national universities. This indicates that there is a general tendency to fill professorial positions with internal candidates rather that hiring external academics.

Post-doc and research stays

MIUR data on academic career shows researchers yearly career starting from their first appointment in Italian universities. In addition to that, I gathered publication

TABLE 2.4: Mobility pattern of promoted academics

	Internal	Mobile	Abroad
Promoted to AP	762/967 (79%)	130/967 (13%)	75/967 (8%)
Promoted to FP	181/209 (87%)	20/209 (9%)	8/209 (4%)

TABLE 2.5: Postdoctoral mobility by PhD cohort

Cohort	Post-doc	Alma mater	Abroad	Hired at post-doc inst.	Multiple aff.
93-98	13%	34%	48%	31%	13%
98-02	17%	49%	24%	51%	13%
02-06	26%	54%	25%	43%	12%

data from Scopus that allows to identify other types of mobility:

- from publications published between PhD and first appointment in Italian academia, I retrieved *postdoctoral positions* and *research stays*;
- from publication published after first appointment in Italian academia with a different affiliation with respect to the one in the MIUR data, I retrieved *multiple affiliations*.

20% of the sample published a paper before the first appointment in Italian academia, 49% of them spend a post-doc period in the own *alma mater* and 29% spend a research period outside Italian academia. In 43% of the cases researchers obtained a position in the institution where they carried out their postdoctoral research.

12% of the whole sample of researchers exhibit multiple affiliations, meaning that they publish being affiliated to an institution where they worked in their career which is not the current one, in 11% of the cases this second affiliation is not Italian.

As Table 2.5 shows, early career mobility has different characteristics across cohorts. Postdoctoral mobility became more common in recent years, also thanks to the formal introduction of research fellowship status (*assegnista di ricerca*) with law 449/1997. After that law, not only diminished the percentage of researchers who spent their post-doc period outside Italian academia, but also became prevalent to remain in the own PhD granting institution. On the other hand, remaining affiliated to a previous institution does not show substantial differences across cohorts.

I collected from Scopus the number of publications in international journals in the fields of economics, business, statistics and decision science for each researcher; in addition I also collected the number of citations received by each publication as a quality measure. Each professor published on average three articles during the observation period (from PhD to tenure, or until the spell is censored) and publications received on average 27 citation.

TABLE 2.6: Performance, age and ranking measures by type of mobility for tenured academics

	Mobile		Inbred		Abroad	
	1	0	1	0	1	0
Years from PhD	7.45	8.45	8.74	7.89	8.39	8.09
#Publications	2.40	2.89	3.24	2.53	4.55	2.43
Avg Citations	36.51	33.28	40.16	33.08	52.65	28.51
Rank PhD Uni	43.22	41.62	45.61	40.63	47.91	41.99
Rank Current Uni	33.52	39.25	45.61	34.37	42.05	37.03

TABLE 2.7: Candidates' published oeuvre by SSDs

	Econ	Busi	Hist	Stat
N	469	728	65	399
Tenured	267	443	28	239
% Tenured	0.57	0.61	0.43	0.60
Articles	8.54	3.16	1.51	6.24
Book Chapters	0.79	0.73	0.45	0.39
Books	0.12	0.05	0.12	0.02

Table 2.6 show various performance, age and ranking measures by type of mobility for academics who got tenure. Academics who have been job mobile since entering their first appointment receive tenure earlier than academics who do not move. Inbred academics, on the other hand (those who received tenure from their PhD granting institution) are more likely to hold an appointment at a university with a higher rank with respect to the other categories. This indicates that mobility in Italy is usually downward. Academics engaging in research visits outside Italian academia publish more and receive more citations, which may point toward a positive network effect.

Economic fields

A total of 977 researchers who earned PhD in the period 1993-2006, received tenure in economics and related fields until 2015. In Table 2.7 and Table 2.8 I report candidates' published oeuvre (refereed articles, books and books chapters) at the time of tenure or at the end of observation period ("censored"). I grouped individuals in four categories according to their SSD.

In Table 2.7 I report publications per capita and I can see that researchers in the field of economics publish a higher number of articles with respect to the other subdisciplines. Economists also publish more books, together with economic historians, and they have more contributions to collective volumes with respect to the others.

TABLE 2.8: Candidates' published oeuvre by SSDs if tenured or censored

	Econ	Busi	Hist	Stat
		Tenured		
Articles	0.51	0.15	0.08	0.29
Book Chapters	0.04	0.03	0.02	0.01
Books	0.00	0.00	0.00	0.00
Productivity	0.30	0.05	0.03	0.13
		Censored		
Articles	0.28	0.10	0.06	0.17
Book Chapters	0.03	0.02	0.02	0.01
Books	0.00	0.00	0.00	0.00
Productivity	0.11	0.03	0.02	0.07

TABLE 2.9: Duration until tenure by SSDs

	Econ		Busi		Stor		Stat		Tot	
Years	abs	cum	abs	cum	abs	cum	abs	cum	abs	cum
1	1%	1%	3%	3%	0%	0%	1%	1%	2%	2%
2	3%	4%	4%	7%	0%	0%	2%	3%	3%	5%
3	2%	6%	7%	14%	0%	0%	4%	7%	5%	10%
4	7%	13%	9%	23%	0%	0%	8%	15%	8%	18%
5	6%	19%	7%	30%	11%	11%	7%	22%	7%	25%
6	11%	30%	11%	41%	4%	15%	9%	31%	10%	35%
7	9%	39%	10%	51%	7%	22%	13%	44%	11%	46%
8	11%	50%	9%	60%	11%	33%	8%	52%	9%	55%
9	12%	62%	11%	71%	11%	44%	12%	64%	12%	67%
10	12%	74%	12%	83%	18%	62%	7%	71%	11%	78%
>10	26%	100%	17%	100%	38%	100%	29%	100%	22%	100%

In Table 2.8 I report the average publications per capita by year until the tenure is awarded or until the end of the observation period. Here the difference between the publication record of economists and the other researchers is even more striking. At the time of the tenure decision, economists have published in journals between 1.7 (with respect to statisticians) and 6.4 (with respect to historians) as much as newly appointed professors of the other sub-disciplines. This is a very substantial difference. In part this may be due to a stronger selection in part may be the result of a longer duration between PhD and tenure in economics.

In Table 2.9 I report the duration until tenure professorship is earned for those who were eventually tenured. After eight years 54% of the researchers have been awarded with tenure, and for less than the 25% it will take more than 10 years. I can observe, however, that there is a distinct difference in the time to tenure for business candidates. The percentage of tenured for researchers in business discipline grows faster than the others. It takes seven years to business candidates to reach the 50%

of tenured researchers, while corresponding figure is nine years for the economists and ten years for historians.

2.6 Method and variables

2.6.1 Empirical strategy

After the descriptive results on the data set that I have generated, I turn to the empirical analysis of the data. As already said, I will focus in this paper on the analysis of the duration from the PhD degree until achieving a position in Italian academia (either the first appointment or tenure). Thus, I analyze the probability of becoming professor given the time an individual is already in the academic labor market. This conditional probability, the hazard rate, is the measure of success in the academic labor market. Furthermore, I analyze the impact of explanatory variables (the covariates) on the hazard rate.

Let the duration T of the state of being a post-doc researcher until being awarded a position be a random variable. Denote t a realization of the random variable. Finally, denote $f(t)$ the density (which has a positive support since the realizations of T can by construction only be positive) and $F(t)$ the cumulative distribution function. The hazard is in this case defined as (see for details Cameron and Trivedi, 2005):

$$\lambda(t) = \lim_{\Delta \to 0} \frac{Prob(t \leq T < t + \Delta | T \geq 1)}{\Delta} = \frac{f(t)}{1 - F(T)} \tag{2.1}$$

The hazard rate is in this case the rate at which the "failure" event occurs given that the individual has "survived" until t. This approach has two major drawbacks. First of all, data are characterized by right-censorship and by the fact that more than one failure occurs at the same time, which is impossible with a continuous function. Second, and more importantly, with this approach I would have to assume an exact functional form without knowing the real underlying process.

To avoid this problems, I first estimate a non-parametric duration model using the Kaplan-Meier estimator (Kaplan and Meier, 1958), which gives the hazard rate for the event without taking into account the explanatory variables. The estimator for the hazard rate at some time t_i is the following. Let n_i denote the number of those individuals that have not yet failed or censored just until time t_i (the number of individuals at risk), let d_i denote the number of those individuals who experience a failure event at t_i. The estimated hazard rate is hence:

$$\lambda_i = \frac{d_i}{n_i} \tag{2.2}$$

The Kaplan-Meier (KM) estimate shows the evolution of the success in the academic labor market as functions of the time already spent in the career path towards a position in Italian academia. The crucial assumption in this analysis, however, is the heterogeneity among individuals. The KM does not differentiate with respect to personal characteristics or different times of the tenure decision. But it exists heterogeneity between individuals, and gender or publication record are likely to determine the success in the academic labor. In order to take into account these and related issues concerning the determinants of the hazard rate of becoming tenured (i.e. success in the academic labor market) it is necessary to apply some sort of regression analysis. The literature offers two types of regression analysis for failure time data. One can choose a parametric approach, which again implies assumptions on the exact functional form of the cumulative distribution function. The alternative approach is the semi-parametric Cox proportional hazard model.

In this study I favor the Cox model (Cox, 1972), because it is more flexible than the parametric one and poses less restrictions on the data. Hence, it has become the standard method for the analysis of failure time data (see Van den Berg, 2001 for an overview of the many application of the Cox model). The idea of the Cox proportional model is that the hazard rate (conditional on the vector of covariates of an individual $\mathbf{X}$) can be separated into the product of the baseline hazard function (which is only a function of duration time) and a function of the covariates. Assuming the covariates to have an exponential impact on the hazard rate, I get:

$$\lambda(t|\mathbf{X}) = \lambda_0(t) exp(\mathbf{X(t)fi}) \tag{2.3}$$

where $\lambda_0(t)$ denotes the baseline hazard function. Due to the semi-parametric nature of the approach, I do not need to specify the functional form of this baseline hazard. So the basic idea of the Cox approach is to estimate the impact of covariates on the hazard rate relative to the unspecified baseline hazard rate. Changes in the covariates hence shift for all individuals the common baseline hazard rate. In many applications it is assumed that the covariates do not change over the duration. Thus, their impact is constant over time. Hence, the name proportional hazard models.Kalbfleisch and Prentice (2002), however, note that considering time constant covariates is a unnecessary restriction of the model. Hence, I allow for time varying coefficients. The parameter vector **fi** is estimated by maximization of the partial likelihood function. The partial likelihood function is the product of the failure probability of an individual conditional on the failure probabilities of all individuals being at risk at some duration t_j.

Another crucial assumption of the Cox model is that the baseline hazard rate is assumed to be the same for all individuals. The hazard rate of getting tenure is without doubt also driven by some unobserved ability or talent level. Hence, one might worry whether this homogeneity assumption results in a misspecification of the model. In order to account for unobserved heterogeneity I estimated a random effect Cox model (see Cameron and Trivedi, 2005):

$$\lambda(t|\mathbf{X}, v) = \lambda_0(t) v exp(\mathbf{X(t)fi}) \tag{2.4}$$

Where v is random variable with $E(v) = 1$ and $Var(v) = \sigma_v$ which reflects the unobserved heterogeneity. Estimation of this model, assuming v being Gamma distributed, show that the estimated variance σ_v is not significantly different from zero. Hence the results reported in section 2.7 obtained from the model without unobserved heterogeneity are valid.

The next challenge was to conceptualize the dependent variable, which is the hazard of entering into an academic position. I divide the analysis in i) time to first appointment and ii) time to tenure, developing *ad hoc* explanatory variables for the two models. The goal is in fact to better disentangle the role played by mentorship and other relevant variables on the different stages of academic careers. I built three classes of explanatory variables: individual characteristics, human and social capital and mobility. Table D.1 presents descriptive information on all variables used.

TABLE 2.10: Descriptive information on the variables

Variable	Explanation
	Individual characteristics
Female	1: individual is a woman
AgePhd	Age at which the PhD is earned
Precocity	1: publication in international journals during the PhD
#Publications	Number of publication until the failure event (or until censored)
#Citations	Number of citations until the failure event or until censored
Productivity	Average productivity until the failure event or until censored
	Human and Social Capital
SharePubAdvisor	Share of articles coauthored with advisor until failure event or until censored
ExternalAdvisor	1: advisor is not affiliated with the PhD granting university
NetworkAdvisor	Number of affiliations in which the advisor has worked or has coauthors
TopPhdUni	1: PhD university is among top 10% institutions in Italy according to IDEAS rank
	Mobility
Inbred	1: inbreeding for the first affiliation
Inherit	1: first affiliation is "inherited" from advisor's network of affiliations
External	1: first affiliation is different from PhD university and external to advisor's network
ResearchAbroad	1: researcher has published a paper being affiliated to a foreign institution
#ChangeAff	Number of affiliation changes until tenure (only time-to-tenure analysis)
	Other Controls
SSD dummies	Economics; Business; History of economics; Statistics
Time dummies	3 PhD cohorts: 1993-1997; 1998-2002; 2003-2006

2.6.2 Variables and controls

Among the individual characteristics, research productivity measured by publication is of course a central element in academic performance and thus, in a merit-based system like the Italian one, it should play a central role in career advancement. In particular, I argue that the most important phase in which a researcher should try to publish is the beginning of his career. I extracted from Scopus all the scientific articles published in international economics, statistics and business journals authored by at least one of the academics in the sample. Each Scopus record lists, in separate fields, the author's or co-authors' names, and affiliations with one-to-one correspondence. Since I expect that scientific production and especially publication in international journals has a large impact on the hazard of being tenured, I use three bibliometric measures, in order to fully take in to account quality and quantity of the publications: $\#Publications$ is the number of articles published in the time period between PhD and the failure event (either first appointment or tenure); $\#Citations$ is the number of citations received in the same period; $Productivity$ is a measure amended from Pezzoni et al. (2012), which consists of the sum of articles, adjusted for journal quality, through the SCImago Journal Rank, and the number of authors until the observation fails or is censored. The formula for the $Productivity$ variable is the following: $\sum_{a=1}^{Ni,t} \left(\frac{SJR}{aut}\right)_a$. I also built the dummy $Precocity$ if the researcher has published, during or within one year after finishing his PhD. Other variables of interest at the individual level are scientists' age at the year in which they have obtained their doctoral degree ($AgePhd$), and gender ($Female$ dummy).

The explanatory variables of main interest are those related to the creation of social capital among the researchers. A specific type of social capital is acquired in the early stage of academic life, in the context of the relationship between the PhD student and his supervisor. I measure human and social capital using variables derived from the doctoral dissertation database and Scopus records. First I take into account the role of *coauthor* that an advisor could have embodied: $SharePubAdvisor$ is the share of articles coauthored with the advisor until first appointment or tenure (or the end of observation period). The idea behind this variable is that it gives an indication of how important and persistent is the advisor–advisee relationship. Second, I explore to what extent advisor's network of acquaintances is responsible for researchers' career. *ExternalAdvisor* is a dummy which takes value 1 if the advisor is not affiliated with the PhD granting university, giving access to her protégé to a scientific network external from her own. To fully explore the width of advisor's network, the variable *NetworkAdvisor* counts the number of affiliations in which the advisor has worked or draws her coauthors.

Although in Italy all universities and thus all PhD titles should be considered equal, I also paid attention to the reputation of the PhD granting institution, since this may affect researchers' career success. I included a dummy variable, *TopPhdUni* which takes value 1 if the Italian university where the researcher earned the doctorate is among the top 5% international institutions in economics according to the IDEAS ranking.

Finally, I operationalize mobility by several variables. First, I focus on mobility decisions during the early career stage, taking into account the first affiliation after the PhD. I consider as first affiliation either the one of the first publication after the PhD or the first appointment in Italian academia. I then build three mutually exclusive dummy variables: *Inbred* when the first affiliation is the PhD granting university; *Inherited* when the first affiliation belongs to the advisor's network (is one of the advisor's previous affiliations or one of his coauthors' affiliation); *External* when the first affiliation is different from the PhD granting university and external to advisor's network.

Furthermore, I created the dummy *ResearchAbroad*, for those who spent a period away from Italian universities before returning to the country. I have been able to identify the researchers who went abroad thanks to the affiliation reported in the publication data from Scopus. For the time-to-tenure analysis I added the variable *#ChangeAff* which is the number of affiliation changes until tenure (or the end of observation period). Theoretically, on the one hand, staying in the same university for the whole career could have the effect of reinforcing the relations with the local environment; on the other hand, hiring a mobile researcher could allow the new institution to inherit his scientific network endowment.

In order to control for cross-disciplines characteristics I inserted in the specification a dummy variable for each macro scientific sector[7]. To control for changes in the institutional framework of the academic labor market I have included time dummies into the regression which indicate in which five-year window the PhD was awarded. Table B.1 and Table B.2 in Appendix B reports summary statistics for the variables for the following empirical analysis.

2.7 Results

2.7.1 Time to first appointment

To provide a first impression of the survival process I depict the Kaplan-Meier survival estimate for the whole sample and separately for candidates from each subscientific field in economics (Figure 2.2).

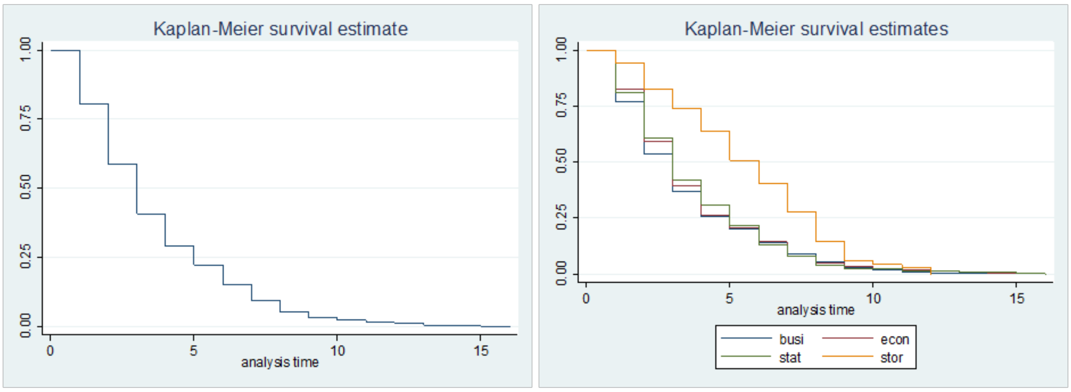

FIGURE 2.2: Kaplan-Meier survival estimate, whole sample (left) and by scientific field (right)

The survival function shows the probability of surviving (not entering Italian academia) for realizations of the random variable "duration until first appointment". The median of the survival function is three years, i.e. with a probability of 50% an individual has to wait more than three years before obtaining a position in Italian

[7]The macro scientific sectors considered are: 13/A–Economia (SECS-P/01 - Economia politica, SECS-P/02 - Politica economica, SECS-P/03 - Scienza delle finanze, SECS-P/06 - Economia applicata, SECS-P/05 -Econometria), 13/B – Economia aziendale (SECS-P/07 -Economia aziendale, SECS-P/08 - Economia e gestione delle imprese, SECS-P/10 - Organizzazione aziendale, SECS-P/11 - Economia degli intermediari finanziari, SECS-P/09 - Finanza aziendale, SECS-P/13 - Scienze merceologiche), 13/C – Storia economica (SECS-P/12 - Storia economica, SECS-P/04 - Storia del pensiero economico) and 13/D – Statistica (SECS-S/01 - Statistica, SECS-S/02 - Statistica per la ricerca sperimentale e tecnologica, SECS-S/03 - Statistica economica, SECS-S/04 - Demografia, SECS-S/05 - Statistica sociale, SECS-S/06 - Metodi matematici dell'economia e delle scienze attuariali e finanziarie).

academia. Furthermore, note that, after this point, this probability is declining more slowly. Moreover, the survival function shows that between the duration of one to three years, the probabilities of reaching a position within the next year are larger (as indicated by the size of steps of the survival function). From the figure detailed by scientific field, it is possible to notice that economic historians are the candidates which take more years to obtain a position in Italian universities, with respect to the other fields.

The estimation results of the Cox model are given in Table 2.11. Let us first of all comment on the results of the most obvious specification, the one in which I only allow for direct effects of explanatory variables on the hazard rate. I incorporated the several groups of variables step by step into the model. I begin with personal determinants (Column 1), then I add human and social capital (Column 2), mobility (Column 3) and finally I add the controls for scientific fields (Column 4) and PhD cohort (Column 5). I display the hazard ratios to facilitate the interpretation of results. This implies that an effect is positive if the hazard ratio is above 1, and negative if below 1.

First of all, there seems to be no clear gender pattern in any statistically significant way, nor youth seems to play a role at this career stage. On the other hand, precocity gives a highly significant and strong advantage, persistent among the different specifications: having published during or immediately after the PhD, increases the likelihood to obtain a position by more than 30%. It transpires that variables about publications and citations are statistically significant and positively correlated with the hazard rate of getting the first appointment, while the productivity index is negative, which may indicate that spending effort in trying to publish on better journals could be detrimental in terms of how fast the first appointment is obtained, but it is not significant.

With respect to human and social capital, two variables turn out to be significant. Publishing with the advisor is significant through all the specifications and negative for the chance of getting a job in academia, probably because a prestigious author might take greater credit for the performed work. On the other hand, the greater advisor's network of acquaintances slightly enhances the probability of getting a position, even if the effect diminishes in significance and eventually disappears by

TABLE 2.11: Cox hazard model for the first appointment in Italian academia

	(1)	(2)	(3)	(4)	(5)
Female	0.997	1.002	1.008	0.992	0.994
	(0.0509)	(0.0513)	(0.0517)	(0.0509)	(0.0513)
AgePhd	1.009	1.008	1.007	1.013	1.013
	(0.0114)	(0.0114)	(0.0114)	(0.0117)	(0.0116)
Precocity	1.293**	1.314**	1.321**	1.322**	1.320**
	(0.124)	(0.127)	(0.128)	(0.128)	(0.128)
#Publications	1.034	1.060*	1.078**	1.080**	1.080**
	(0.0253)	(0.0264)	(0.0272)	(0.0275)	(0.0276)
#Citations	1.139	1.197	1.249	1.331*	1.331*
	(0.157)	(0.160)	(0.169)	(0.183)	(0.183)
Productivity	0.997	0.993	0.991	0.990	0.989
	(0.00708)	(0.00710)	(0.00704)	(0.00710)	(0.00710)
SharePubAdvisor		0.780+	0.711*	0.703**	0.694**
		(0.102)	(0.0954)	(0.0943)	(0.0932)
ExternAladvisor		1.056	1.020	1.069	1.080
		(0.0605)	(0.0675)	(0.0714)	(0.0725)
NetworkAdvisor		1.009*	1.009*	1.007+	1.006
		(0.00403)	(0.00407)	(0.00418)	(0.00423)
TopPhdUni		1.008	1.007	1.030	1.031
		(0.0575)	(0.0576)	(0.0591)	(0.0593)
ResearchAbroad			0.567**	0.582**	0.582**
			(0.0798)	(0.0822)	(0.0824)
Inbred			1.149*	1.126+	1.115
			(0.0769)	(0.0756)	(0.0750)
Inherited			1.237**	1.212*	1.203*
			(0.0996)	(0.0979)	(0.0972)
Busi				1.192**	1.193**
				(0.0777)	(0.0781)
Hist				0.534**	0.535**
				(0.0741)	(0.0743)
Stat				1.164*	1.164*
				(0.0835)	(0.0840)
Cohort9397					0.844*
					(0.0687)
Cohort9802					0.943
					(0.0603)
Individuals	1,590	1,590	1,590	1,590	1,590
Observations	5,942	5,942	5,942	5,942	5,942
BIC	20799.64	20806.1	20808.13	20786.76	20803.95
Log Likelihood	-10374	-10360	-10348	-10324	-10324
Chi-square	17.24**	45.53	69.57	117	117.2

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

adding the cohort controls. The effect of the variables about having an external advisor and having obtained the doctoral degree from a top university follow the expectations, but are not statistically different from zero.

Variables about mobility shows that the effect of having published an article being affiliated to a foreign university significantly, and persistently across specifications, reduces the hazard rate of getting a position in Italian academia. This highlights that, in the Italian academic system, spending a research period abroad is detrimental in the short period for an academic career. Taking as reference category researcher who, after the PhD, move to a university different from their *alma mater* and external to their advisor's network, it is possible to see that both inbreds and those who inherited their first affiliation exhibit a higher hazard rate of getting an academic position. In particular, the second category shows a relevant and significant effect even after introducing scientific field and cohort controls in the specification.

Scientific field controls coefficients have the expected sign and cohort controls indicate that those earning the PhD in more recent years (reference cohort is 2002-2006) are more likely obtain an academic position than those who finished during the 1990s.

2.7.2 Time to tenure

Turning to time-to-tenure analysis, I can see from the Kaplan-Meier curve (Figure 2.3) that the median of the survival function is eleven years for the whole sample, but it is much higher for the candidates in the economic history scientific field. The hazard rate function (Figure 2.4) is hump shaped, implying that the probability of being tenured is increasing at the beginning of the career, peaks and then decreases. This is not too surprising since this is a standard career pattern. The interesting point, however, is the exact duration of the peak. The hazard rate is increasing until twelve years after the PhD degree. Thus, until this time it pays (in terms of increasing the probability of being tenured) to stay in the academic labor market. From Figure 2.4 I can also see that the probability of getting tenure for female candidates is lower at any point with respect to male candidates.

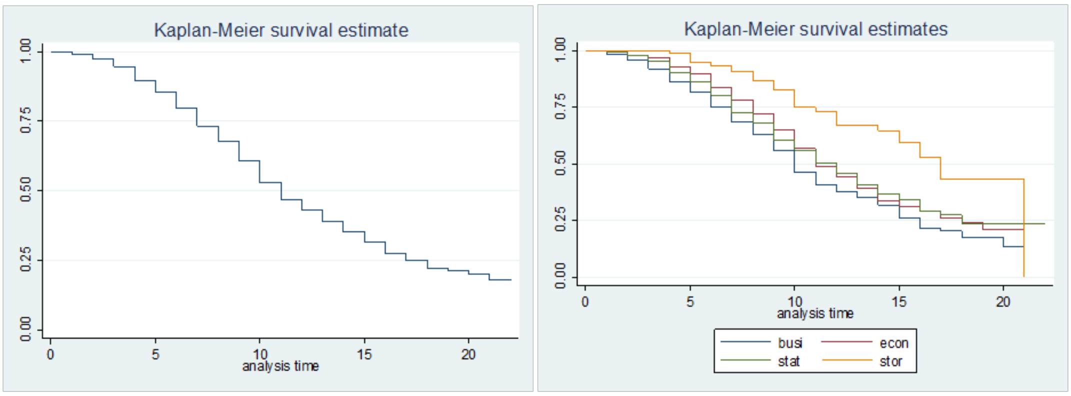

FIGURE 2.3: Kaplan-Meier survival estimate, whole sample (left) and by scientific field (right)

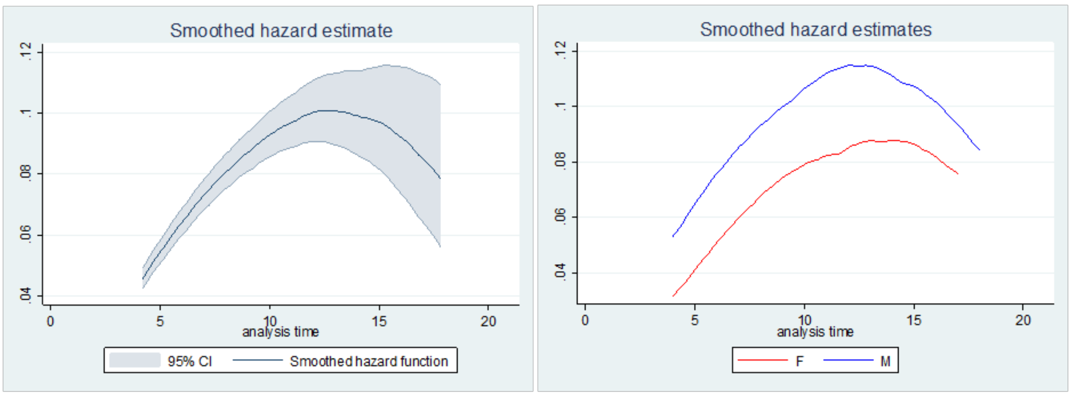

FIGURE 2.4: Hazard rate function, whole sample (left) and by gender (right)

TABLE 2.12: Cox hazard model of the hazard of tenure in Italian academia

	(1)	(2)	(3)	(4)	(5)
Female	0.671**	0.668**	0.666**	0.650**	0.691**
	(0.0454)	(0.0452)	(0.0451)	(0.0440)	(0.0473)
AgePhd	0.953**	0.956**	0.955**	0.963**	0.967*
	(0.0129)	(0.0130)	(0.0130)	(0.0132)	(0.0138)
Precocity	1.528**	1.525**	1.512**	1.540**	1.678**
	(0.148)	(0.148)	(0.149)	(0.153)	(0.168)
#Publications	1.074**	1.080**	1.080**	1.088**	1.109**
	(0.00933)	(0.00938)	(0.00938)	(0.00930)	(0.00922)
#Citations	1.056	1.060	1.071	1.081	1.242**
	(0.0845)	(0.0842)	(0.0846)	(0.0843)	(0.0937)
Productivity	1.313+	1.329+	1.264	1.390*	1.337+
	(0.207)	(0.213)	(0.209)	(0.226)	(0.207)
SharePubAdvisor		1.034	1.241*	1.237+	1.252*
		(0.105)	(0.127)	(0.136)	(0.139)
ExternAladvisor		1.318**	1.265**	1.317**	1.239**
		(0.0870)	(0.100)	(0.106)	(0.101)
NetworkAdvisor		0.983**	0.984**	0.990+	0.995
		(0.00559)	(0.00562)	(0.00571)	(0.00540)
TopPhdUni		1.078	1.091	1.109	1.106
		(0.0709)	(0.0720)	(0.0736)	(0.0737)
Inbred			1.181+	1.138	1.140
			(0.113)	(0.109)	(0.110)
Inherited			1.147	1.070	1.063
			(0.106)	(0.100)	(0.100)
ResearchAbroad			0.725**	0.727**	0.852
			(0.0816)	(0.0839)	(0.0992)
#ChangeAff			1.103*	1.112*	1.084+
			(0.0524)	(0.0527)	(0.0517)
Busi				1.616**	1.865**
				(0.136)	(0.158)
Hist				0.733	0.807
				(0.150)	(0.165)
Stat				1.237*	1.176+
				(0.114)	(0.109)
Cohort9397					2.448**
					(0.240)
Cohort9802					1.821**
					(0.154)
Individuals	1,672	1,672	1,672	1,672	1,672
Observations	16,454	16,454	16,454	16,454	16,454
BIC	13351.05	13363.08	13391.04	13375.01	13251.07
Log Likelihood	-6646	-6633	-6628	-6605	-6533
Chi-square	145.6**	172.4**	183.3**	228.5**	371.8**

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 2.12 presents the results of a series of nested Cox regressions, which estimate the factors that increase or decrease the odds that the event of interest occurs, that is that a researcher gets a tenured position.

Column 1 in Table 2.12 is a baseline model which includes the variables which represents individual characteristics. It is possible to see that they are highly significant, and stay so throughout the other specifications of the model. First I notice that women have a 31% lower chance to get tenure with respect to men with same characteristics. A smaller but also negative effect has the age at which one gets the PhD: an additional year lowers the hazard rate of 3%. Having published an article in international journals during the PhD years has a strong and positive effect toward the hazard of getting tenure, while the results for the other productivity variables are mixed. Interestingly, the coefficient for the number of citations is positive but the variable affects the "hazard" for tenure only in Column 5, while the number of publications has a positive and significant effect. This may mean that quantity matters more than quality for getting a professorial position in Italy, in fact the productivity measure, which takes both characteristics into account, is positive and significant.

Column 2 introduces human and social capital variables into the model. It shows that the prestige of the university where someone received the doctorate does not affect the odds of receiving tenure. The coefficient of the variable for the width of the advisor's network is surprisingly negative, but is not significant when scientific fields and cohorts controls are added to the model. On the other hand, having an advisor external to the own *alma mater* gives a substantial advantage in the hazard of getting tenure. Also the share of publications coauthored with the advisor has a positive effect, and significant at the .05-level.

Mobility variables, regarding the first job affiliation and having published a paper being affiliated to a foreign university, are no longer significant in determining the hazard of getting tenure when scientific fields and cohorts controls are introduced into the specification. However the number of affiliation changes remains slightly significant at the 10 percent level and positive. Thus developing research experience within several Italian university is a valuable characteristics to get tenure in the national academic system.

Finally, taking into account Column 5, scientific field controls are in line with the descriptive analysis above: business researchers exhibit higher odds of getting tenure, while economic historians are those who have to wait for the longer period of time. Cohort controls show that researchers who received PhD until 2002 are more likely to get tenure with respect to the more recent cohort of doctorate holders (until 2006).

2.7.3 Additional networks controls

Additionally to the previous results, I added as regressors two measures of centrality, betweenness and closeness, calculated for the PhD supervisor inside the network of all Italian economists for each of the four sub disciplines: economics, business, economic history and statistics. To build the networks I downloaded all supervisors' publications from Scopus and calculated the supervisor's centrality measures at the year in which each individual earned the doctorate.

Table 2.13 shows the results of the Cox proportional hazard model. It is possible to see that for time to tenure analysis in Column 2 the closeness centrality variable is positive and significant at the 10% level, while in the time to first appointment analysis, betweenness centrality is also significant at the 10% level but influences negatively the hazard rate. Results show a paradoxical network effect: if advisor centrality seems to play a positive role in the road to tenure, it is also detrimental at the beginning of the career, making the post-doc duration last longer.

2.7.4 Time varying covariates

Specifications in Table 2.11 and Table 2.12 do not consider time varying covariates (TVC). By the construction of the Cox Model, this implies that the effect of the covariates is a proportional shift of the baseline hazard rate. As such, the shift of the baseline hazard rate is independent of the duration. If, however, the covariate effect is changing over time, the model would be misspecified.

To test the proportionality assumption, I apply a test which is based on Schoenfeld residuals (see Grambsch and Therneau, 1994). The residuals are basically the

TABLE 2.13: Cox hazard model of the hazard of tenure in Italian academia (with centrality controls)

	Time to first app. (1)	Time to tenure (2)
Female	1.019	0.689**
	(0.0544)	(0.0471)
AgePhd	1.014	0.966*
	(0.0108)	(0.0139)
Precocity	1.480**	1.736**
	(0.135)	(0.173)
#Publications	1.059*	1.108**
	(0.0270)	(0.00926)
#Citations	1.295+	1.240**
	(0.180)	(0.0936)
Productivity	0.992	1.354+
	(0.00717)	(0.246)
SharePubAdvisor	0.717*	1.256*
	(0.0932)	(0.140)
ExternalAdvisor	1.114+	1.230*
	(0.0707)	(0.100)
NetworkAdvisor	1.006	0.996
	(0.00421)	(0.00557)
TopPhdUni10	1.045	1.106
	(0.0565)	(0.0738)
Betweenness	0.764+	0.804
	(0.122)	(0.155)
Closeness	1.061	1.241+
	(0.116)	(0.161)
Inbred	1.119+	1.142
	(0.0709)	(0.111)
Inherited	1.214**	1.063
	(0.0912)	(0.100)
ResearchAbroad	0.672**	0.847
	(0.0902)	(0.0992)
#ChangeAff		1.092+
		(0.0525)
Busi	0.989	1.877**
	(0.0656)	(0.160)
Hist	0.506**	0.810
	(0.0711)	(0.166)
Stat	1.013	1.174+
	(0.0762)	(0.109)
Cohort9397	0.837*	2.393**
	(0.0636)	(0.235)
Cohort9802	0.924	1.797**
	(0.0553)	(0.152)
Individuals	1,590	1,672
Observations	5,942	16,454
BIC	20829.64	13269.87
Log Likelihood	-9593	-6533
Chi-square	123.7	372.4**

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

difference between observed and expected values of the covariates for an individual. If the effect of a specific covariate is constant over the duration, this residual should not change (see for example Box-Steffensmeier and Jones, 2004).

I conducted proportionality tests for every covariate in the full specifications in Table 2.11 and Table 2.12. As such, I identified those covariates for which the test rejects the Null Hypothesis at the 90% level and hence I suspect the impact of these variables to change with the duration. A way to control for the non-proportionality is to interact the covariate with a function of the failure time. With this, it is possible to disentangle the impact of the covariate into a fixed and a time-varying effect (see Kalbfleisch and Prentice, 2002). The covariates which I have identified to have a substantial (in the sense of rejecting the proportionality assumption at the 90% level) non-proportional impact are *SharePubAdvisor* and *ResearchAbroad* in the time to first appointment analysis; *Precocity* and *ResearchAbroad* for the time to tenure analysis.

Hence, in Table 2.14 I include interaction between these covariates and the log of duration $ln(t)$, which is the most common function of time used in time-dependent covariates.

Considering the duration interactions modifies the impact of the covariates on the hazard rate. In the time to first appointment analysis, the effect of *SharePubAdvisor* is significantly negative, but this effect is declining in duration. Similarly, in the time to tenure analysis, *Precocity* is highly positive and significant on the hazard rate of getting a professorial position, with a declining effect over time. Of particular interest are the results regarding the international mobility of researchers. The effect of the *ResearchAbroad* dummy is significantly negative and quite large for both analyses. However, the effect is declining in duration and, looking at Figure 2.5, it is possible to notice that after 8 years of duration it is even true that having spent a period abroad in the beginning of the career, increases the hazard rate of getting tenure. The results concerning the effect of individual characteristics or social capital are not affected by the consideration of time variation. Hence, the results turn out to be robust.

TABLE 2.14: Cox hazard model with time varying covariates

	Time to first app.		Time to tenure	
	Main	TVC	Main	TVC
Female	0.988		0.693**	
	(0.0509)		(0.0474)	
AgePhd	1.036**		0.965*	
	(0.0103)		(0.0138)	
Precocity	1.308**		2.699**	0.709*
	(0.112)		(0.886)	(0.119)
#Publications	1.044		1.109**	
	(0.0289)		(0.00950)	
#Citations	1.331*		1.227**	
	(0.184)		(0.0936)	
Productivity	0.992		1.329	
	(0.00731)		(0.242)	
SharePubAdvisor	0.274**	2.001**	1.261*	
	(0.0985)	(0.448)	(0.145)	
ExternalAdvisor	1.092		1.219*	
	(0.0692)		(0.0993)	
NetworkAdvisor	1.004		0.995	
	(0.00406)		(0.00542)	
TopPhdUni	1.040		1.100	
	(0.0562)		(0.0733)	
Inbred	1.117+		1.108	
	(0.0710)		(0.105)	
Inherited	1.226**		1.047	
	(0.0923)		(0.0993)	
ResearchAbroad	0.421**	1.089*	0.0927**	2.867**
	(0.104)	(0.0388)	(0.0452)	(0.628)
#ChangAaff			1.083+	
			(0.0509)	
Busi	1.189**		1.605**	
	(0.0779)		(0.138)	
Hist	0.542**		0.666*	
	(0.0753)		(0.137)	
Stat	1.167*		1.055	
	(0.0843)		(0.0979)	
Cohort9397	0.837*		2.322**	
	(0.0637)		(0.227)	
Cohort9802	0.919		1.763**	
	(0.0552)		(0.149)	
Individuals	1,642		1,667	
Observations	5,942		16,454	
BIC	20802.42		13258.06	
Log Likelihood	-10314		-6513	
Chi-square	136.1		411.9	

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

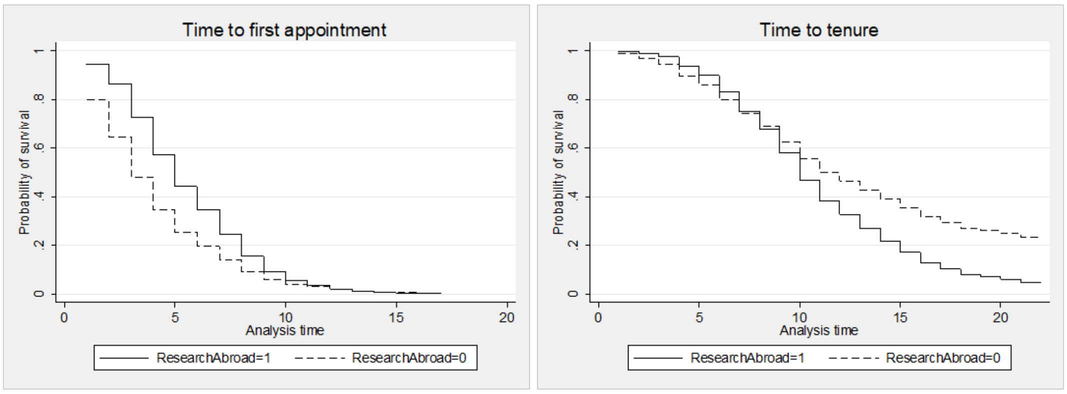

FIGURE 2.5: Predicted survival probabilities by *ResearchAbroad*

2.8 Conclusions

Using a novel dataset on career path of economists in the academic labor market in Italy, I have analyzed which variables determine the success in this market. The notion of success used is the hazard rate, i.e. the conditional probability of getting an assistant professor position in Italian academia and getting tenure. Thanks to the dataset presented in Chapter 1, it was possible to follow the individuals since the PhD through their career path and to get information on the advisor–advisee research relationship. Most of other studies suffer from the survivor bias, i.e. they include only successful persons, thus they are not able to reveal the determinants of career success.

First I applied the non-parametric Kaplan-Meyer estimator to get information on the evolution of the hazard rate for first appointment and tenure. For the first appointment case, the survival function shows that between the duration of one to three years the probabilities of reaching a position within the next year are larger. In the case of tenure, the hazard rate is hump shaped and its rate is increasing until twelve years after the PhD degree.

Additionally to the pure description of the evolution of the hazard rates, I analyzed which covariates influence this hazard. To minimize the functional assumption posed on the regression model, I applied a semi-parametric Cox model. Estimates for the first appointment case show that publish with the advisor has a negative effect on the chance of getting a job, probably because a prestigious author might take greater credit for the performed work. The advisor role at this career stage, however, seems to be very relevant in a more indirect way. Researchers who inherited their first affiliation from the advisor's network show a relevant and significant effect of having a higher hazard rate of getting an academic position.

Turning to the time-to-tenure analysis, I find, as expected, that gender and precocity have a strong effect on the hazard rate. Women have 31 percentage points less chance of getting tenure with respect to male researchers with similar characteristics. The age at which the doctoral degree is granted and having published in the early stage of academic career (during or immediately after the PhD) also are significant. If the prestige of the own *alma mater* does not seem to play a role in this context, having worked in several universities is a valuable characteristic in the Italian academic labor market. For this analysis I further explore the influence of the advisor's centrality within the network of Italian economists for each of the four sub-disciplines considered, finding that closeness centrality has a positive and significant effect in the full specification of the model.

The results additionally show that the selection process taking place during academic careers differ among disciplines, in economic history researchers have a lower probability of getting tenure and have to wait more years, while researchers in business reach professorial position with a higher probability and significantly faster. Furthermore the analysis with time varying covariates show that having spent a research period abroad in the beginning of the career is detrimental for duration length, but this negative effect is declining with time. In terms of hazard rate of getting tenure, after 8 years of duration researchers may eventually benefit from international mobility experience.

Chapter 3

Promotion and Connections in Academic Committees

3.1 Introduction

The role played by social networks and personal contacts in getting a job is one of economic sociology's most known propositions (Granovetter, 1973; Granovetter, 1996). Indeed, labor surveys have shown repeatedly that an important fraction of the population in developed countries cites contacts as a reason they were hired in their current jobs (Marsden and Gorman, 2001; Ioannides and Loury, 2004).

Yet despite the widespread view that personal contacts, and particularly strong ties, often facilitate job finding, the empirical evidence for a clear link between social networks and employment outcomes is limited. A broad survey of the literature on the causal effects of social capital (Mouw, 2006) argues that there is actually little empirical evidence demonstrating a link between contacts and job outcomes. It points to unobserved heterogeneity and reverse causality, two classic sources of bias that are more likely to occur with network variables, as potentially leading to substantial overestimation of the impact of contacts. The article forcefully advocates for methods, such as natural experiments and randomized experiment techniques, which can overcome the statistical limitations.

In this Chapter I study the role of connections in academic promotions using the outstanding evidence provided by a large scale randomized natural experiment: centralized selection exams in Italian academia in 2012.

All candidates for full professor and associate professor positions first had to qualify in a centralized selection exam. Successful candidates could then apply for a position at the university level. Evaluators were selected from a pool of eligible evaluators in each scientific field using a random draw: this feature allows for the consistent estimation of the effect of connections on candidates' chances of success.

For the analysis I consider information on around 9,000 candidacies from three broad academic disciplines: Science, Architecture&Engineering and Social Sciences. These candidacies were evaluated by 75 committees, one for each "*settore concorsuale*". I consider five possible links between candidates and evaluators: supervision of the PhD thesis (*advisor*), co-authorship of a scientific article (*coauthor*), affiliation to the same university (*colleague*), affiliation to the same university during PhD (*PhD colleague*) and co-authorship of a scientific article with the PhD supervisor (*advisor's coauthor*).

I use the information from a novel dataset on Italian PhD holders which provides the identity and affiliation of dissertations' authors and advisors. Only a few other studies focused on the Italian academic labor market and none of them takes into account the crucial relationship between advisor and advisee. Checchi (1999) studies the national academic competition for associate professorship in Economics which took place in Italy during the academic year 1997-98 attempting to measure the relative contribution to the appointment probability made by the candidates' academic credentials. Bagues et al. (2015) use the data on the first round of the Italian national habilitation to examine how the presence of connections in scientific committees affects researchers' decision to apply for a promotion and their chances of success, while Bagues et al. (2017) use the same data, together with the applications to associate and full professorships in Italy and Spain to study how the presence of women in academic committees affects the chances of success of male and female candidates. However, in both studies they do not take into account PhD mentorship as link between candidates and evaluators. This analysis follows the identification strategy of this recent literature and contributes to this stream of research providing for Italy a first empirical evidence regarding the importance of that particular type of strong tie represented by the advisor–advisee relationship in academic selection processes.

The Italian *Abilitazione Scientifica Nazionale* is not at all different from hiring procedure used in others European countries. For instance, the French *concours* and the *concurso de habilitacion nacional* in Spain are national selection processes for academics which are quite similar to the Italian case. But it is indeed arguable that the network variables explored in this study are likely to play a role in hiring decisions in countries, such as Netherlands and UK, where the recruitment system is not based on nationwide exams but on direct negotiations between universities and candidates.

The remainder of the Chapter is organized as follows: Section 3.2 establishes links with previous studies of the academic labour market; Section 3.3 contextualizes the institutional setup of the national habilitation concourse in Italy; Section 3.4 presents the data; Section 3.5 discusses method and results; Section 3.6 concludes.

3.2 The role of mentorship in academic careers

In contrast to some labor markets where network influence in hiring is seen as having a neutral or even positive effect in terms of efficiency, the fact that contacts and networks play a role in academic labor markets is not generally viewed as valuable (Merton, 1976). Science relies on characteristics such as meritocracy and credit granted by peers in a non-market reward system (Stephan, 1996). If the system fails there could be dangerous consequences which could damage the quality of the scientific research itself. The fact that evaluators in charge of hiring or promotion decisions would favor connected candidates, is indeed a potential threat.

One common finding of quantitative studies on academic careers is that productivity is at best a very partial predictor of academic careers (Long et al., 1979; Long and McGinnis, 1981; Leahey, 2007). The commencement and advancement of an academic career seems to correlate more with the productivity and prestige of the mentor and that of the doctoral department than with indicators of individual scientific productivity (Long et al., 1979; Reskin, 1979; Long and McGinnis, 1981). Most studies insist on the overwhelming importance of a sponsor or a mentor, and in particular the PhD advisor (Reskin, 1979; Cameron and Blackburn, 1981; Long and McGinnis, 1985). Future productivity is therefore more a consequence of contextual effects than of initial talent (Long and McGinnis, 1981).

Studies on academic careers in the United States generally focus on long-term outcomes such as career advancement or wages among a set of scholars who have generally succeeded in getting at least their first job in the academic system after the PhD (Hargens and Hagstrom, 1967; Long et al., 1979; Long and McGinnis, 1981; Leahey, 2007). But these studies usually fail to investigate properly the role played by social capital at the entrance to the academia. Analyzing the European state habilitation exams taken upon entrance to an academic career can help to enrich previous studies by focusing on two elements that are often overlooked: the possibility of comparing researchers who succeed to those who fail, and the opportunity to delve more deeply into the social capital mechanisms (direct support or indirect prestige) by which a sponsor may help a PhD to get a job. In the Italian economic field, PhDs benefit from the social capital of their advisor (see Chapter 2). The number of contacts and the importance of the structural holes of the advisor within the network of Italian economists are a predictor of the velocity with which PhDs will enter the academic career.

It is likely, however, that sponsorship becomes effective not only through indirect efforts at promoting the candidate, but also when the applicant has a sponsor on the hiring committee itself. Evaluators may favor acquaintances because of nepotism or clientelism, but it could also be the case that evaluators are better informed about unobservable characteristics of acquainted candidates.

The empirical evidence regarding the effect of having connections among the evaluators in selection processes is relatively scarce. In their study of the *Agrégation du supérieur*, Combes et al. (2008) find that the presence of a person's PhD advisor on the hiring committee has a strong positive impact on the likelihood of that person getting hired, one equivalent to the candidate having written five additional articles. They also find that the presence of colleagues from the applicant's own department has a moderate impact. However, the authors find no significant impact if the hiring committee includes either other faculty from the applicant's doctoral university or coauthors of the applicant's PhD advisor. Zinovyeva and Bagues (2015) find very similar results in their study of the first step in academic recruitment of university professors (*catedrático de universidad*) and associate professors (*profesor titular de universidad*) for all disciplines in Spain from 2002 to 2006: the strongest effect, tripling the odds of recruitment, comes from the presence of the PhD advisor on the selection committee. This effect is followed by the presence of an applicant's coauthor, a colleague from the same university, or another member of the PhD committee.

These studies indicate that in academic labor markets, acquaintances count, with in particular the advisor–advisee contact holding significance. In the case of the Italian academic labor market, previous work by Bagues et al. (2015) and Bagues et al. (2017) provide evidence regarding the existence of a substantial premium emerging from connections in the evaluating committees of the national habilitation concourse. They find robust results pointing towards the fact that connections with coauthors and colleagues are helpful for promotion. However, the previous literature regarding Italian academia, did not explore the role played by that particular type of strong tie represented by the mentor–mentee relationship. This contribution aims at filling this gap.

3.3 Institutional background

For Italian public universities the recruitment of full and associate professors is regulated by national laws. Until 2010, recruitment procedures were managed locally by each university. In 2010, a reform introduced a two-step procedure with a national habilitation (*Abilitazione Scientifica Nazionale* – ASN) and local concourses to recruit professors, similarly to the systems already in place in other European countries, such as Spain and France. Successful candidates at the national level receive the scientific habilitation, which lasts four years, while those who fail to obtain the qualification cannot participate in the ASN for two years. Local universities can choose the professors to hire among the researchers who are granted the national habilitation. The first round of the national scientific qualification was launched in 2012. After the deadline for candidates' application, the evaluation committees were randomly selected. The evaluation procedure have been completed in the first half of 2014. Hereafter I summarize the main characteristics of the Italian selection process[1].

[1]For a more detailed explanation of the process see Bagues et al. (2015)

3.3.1 Selection of committees

For each of the 184 scientific fields designed by the Ministry of Education, University and Research (MIUR), a committee of five members (four full professors from Italian Universities and one foreign member from OECD countries) was randomly selected among the full professors in each field who volunteered for the task and reached some scientific productivity standards.

Evaluation committees could be composed by academics based in Italy or affiliated to a foreign university, with the constraint that no university could have more than one evaluator within the committee. The conducted randomization leaved little room for manipulation. Eligible evaluators in each sub-field were alphabetically ordered and assigned a number. A sequence of numbers was then randomly selected, and used to select committee members in several different fields. Evaluators are in charge for two rounds of the ASN. If a selected evaluator decides to resign, a substitute evaluator is randomly drawn from the corresponding group of eligible evaluators.

3.3.2 Evaluations

Differently from Spanish and French systems, where candidates are evaluated both on the basis of their CVs and oral presentations, in the Italian ASN, evaluation committees assess candidates exclusively on the basis of their publications and CVs. Committees had full autonomy on the criteria to be used in the evaluation but some criteria were suggested by MIUR in relation to the research productivity of candidates in the previous ten years, as measured by some bibliometric indicators. In bibliometric fields, candidates deserving qualification should have a score above the median of the professors of the targeted position in at least two of these three criteria: a) the number of articles published in scientific journals; b) the total number of citations; c) the h-index. In Social Sciences and in Humanities (nonbibliometric fields), successful candidates should pass the median in at least one of the following indicators: a) the number of articles published in scientific journals; b) the number of articles published in high quality journals; c) the number of books. There were no limits to the number of qualifications awarded in each field.

Only family relationships between evaluators and candidates were officially against the conflict of interest rule. When these cases arose, the evaluator in violation of the rule could not participate in the deliberation and the final voting decision. Relationships such as those with advisors, coauthors and colleagues are not affected by the these rules.

3.4 Data

I have collected data from several sources. First, I use information on exams for associate professor (AP) and full professor (FP) positions that were held in Italy in the first round of national scientific habilitation. The data include the identity of candidates and eligible evaluators, the outcome of the lotteries that determined committee composition, and the list of promoted candidates. Second, I have gathered information on the research output of candidates and eligible evaluators published in international academic outlets (Scopus). I also use this database to identify coauthorship. Third, I have collected information for the subsample of candidates' who hold a PhD from an Italian university and evaluators' activity as their PhD advisors. This a novel dataset on Italian doctorate holders which collects the PhD thesis discussed in the period 1986-2006 in Italian universities[2]. Due to the cover period of the latter information source, I consider the habilitation exams of candidate who are active in Italian academia at the time of the evaluation (as assistant or associate professors) excluding the younger cohorts of researchers, which are unlikely to have obtained a PhD within 2006.

3.4.1 Applications

The dataset includes information on 9165 applications for the national scientific habilitation in 75 disciplines, of which 6130 are applications for positions as associate professor and 3035 are applications for positions as full professor. Table 3.1 provides descriptive information on the characteristics of these applications. For both academic position the percentage of habilitate candidates is slightly higher in the Science scientific fields.

[2]More detailed information about the database can be found in Chapter 1.

TABLE 3.1: Descriptives statistics: examinations

	All	Science	ArchEng	SocSci
Associate Professors				
Applications	6130	2131	1922	2077
Habilitated	0,656	0,699	0,683	0,585
Full Professors				
Applications	3035	928	1027	1080
Habilitated	0,570	0,616	0,565	0,534
TOTAL	9165	3059	2949	3157

3.4.2 Candidates

Column 1 in Table 4.12 provides descriptive information about the candidates' characteristics. On average, candidates applied 1.3 times, either because they applied simultaneously in several related "*settori concorsuali*" or for both associate and full professor academics position. One-third of candidates are women and the average applicant is 41 years old. The average candidate had around 22 publications and received around 42 citations. I use the H-index to measure the quality of international publications. The H-index and the number of citations received is larger among candidates to full professor positions than among candidates to associate professor positions, and it is also larger in Science relatively to the other scientific fields.

Columns 2 and 3 provide information disaggregated by type of exam (FP and AP). Candidates for AP positions were three years younger than candidates for FP positions and the share of women among AP candidates was quite larger (36 percent versus 24 percent). Candidates for FP positions had published on average six articles more than candidates for AP positions. The quality of their publications, measured by citations and the H-index, was significantly higher.

Columns 4–6 report candidates' characteristics in the three broadly defined groups of disciplines: Science, Architecture&Engineering and Social Sciences[3]. Social Sciences is the most feminized disciplinary group, with almost 40 percent of the candidates being women. The least feminized field is Architecture&Engineering, where there is only one female candidate for every four male candidates. The most prolific authors are in Science, with a total of 47 publications in international scientific journals. The last row in the table shows that on average 75% of candidates had obtained

[3]Macro-scientific areas considered are: 01-*Scienze Matematiche e Informatiche*; 02-*Scienze Fisiche*; 03-*Scienze Chimiche*; 08-*Ingegneria Civile e Architettura*; 09-*Ingegneria Indistriale e dell'Informazione*; 13-*Scienze Economiche e Statistiche*; 14-*Scienze Politiche e Sociali*.

TABLE 3.2: Descriptives statistics: candidates

	All	FP	AP	Science	ArchEng	SocSci	5y aft
Female	0,32	0,24	0,36	0,34	0,24	0,38	0,320
Age	41,29	43,17	40,36	41,59	41,8	40,7	46,29
# Publications	22,27	26,36	20,24	47,18	16,4	1,72	10,28
# Citations	42,61	51,15	38,37	79,17	26,45	22,1	72,41
H-index	7,28	8,63	6,61	12,25	5,99	3,2	-
PhD-Ita	0,74	0,74	0,74	0,74	0,79	0,68	0,74
App. per capita	1.31	1.24	1.34	1.30	1.25	1.50	1.31
Individuals	6992	2438	4554	2350	2537	2105	6992

TABLE 3.3: Descriptive statistics: connections

	All	AP Exam	FP Exam
Advisor	3,52%	3,71%	3,12%
Coauthor	5,82%	5,33%	6,79%
Colleague	15,58%	15,40%	15,95%
PhD Colleague	10,54%	10,29%	11,04%
Advisor's Coauth.	4,02%	4,32%	3,38%
All ties	27,43%	29,95%	22,34%
Observations	9165	6130	3035

their PhD degree from an Italian university.

3.4.3 Connections

I consider five types of connections between candidates and evaluators in Table 3.3: those found between advisors and students, between coauthors, between colleagues at the time of the evaluation, between colleagues at the time of the PhD and between candidate and the advisor's coauthors. Approximately 4 percent of applicants were evaluated by a committee that included their thesis advisor, 6 percent by coauthors, and 16 percent by colleagues; the other two (more indirect) ties apply in the 10 percent and 4 percent of cases, respectively (Table 3.3, Column 1).

Overall, more than one fourth of applicants had a link in the committee. The presence of coauthors and colleagues in committees is slightly larger in exams for FP positions, while the opposite is true for advisors and their coauthors (Table 3.3, Columns 2 and 3).

3.4.4 Future performance

I measure the future performance of candidates in two different ways. Firstly, I have gathered information on productivity indicators during the five years following the examination (Table 4.12, Column 7). During this period the average candidate had published 10 publications which received around 70 citations. Additionally, for the sample of candidates who were habilitated for AP or FP positions, I have collected information on their promotion at the university level up to December 2016. About 45 percent of these candidates managed to be promoted.

3.5 Empirical analysis

If connected evaluators are better informed and/or biased, the presence of a connection in the committee is expected to increase candidates' chances of success. Following the empirical strategy adopted byZinovyeva and Bagues (2015), I test this hypothesis using the following equation:

$$Y_{i,e} = \beta_0 + \mathbf{C}_{i,e}\beta_1 + \mathbf{E}_{i,e}\beta_2 + \epsilon_{i,e} \tag{3.1}$$

The dependent variable is a dummy that indicates whether individual i qualified in exam e. The vector of coefficients β_1 identifies the causal effect of connections, the vector $\mathbf{E}_{i,e}$ includes the dummies for the expected number of each connection in the committee before the random selection of the committees takes place. To increase the accuracy of the estimates I include a set of area fixed effects, accounting for possible differences across scientific areas and, in some specifications, a vector $\mathbf{X}_{i,e}$ which contains the characteristics of the candidate i in exam e. I consider committees composition using the outcome of the initial random draw, which might be slightly different from the committee composition that actually evaluated the candidates due to the resignation of few evaluators. Hence in this specification β_1 provides the intention to treat effect (ITT). I compute robust standard errors clustered at the exam level, reflecting the fact that evaluations within each field are done by the same committee, and in order to account for common shocks within the same exam and for the existence of heteroskedasticity.

TABLE 3.4: Impact of connection on candidates' success (ITT)

	All (1)	AP (2)	FP (3)	PhD ITA (4)	All (5)	Experimental (6)
Advisor	0.115**	0.0958**	0.154*	0.102**	0.113**	0.119**
	(0.0316)	(0.0321)	(0.0590)	(0.0313)	(0.0315)	(0.0323)
Coauthor	0.117**	0.115**	0.132**	0.0927**	0.113**	0.112**
	(0.0237)	(0.0254)	(0.0429)	(0.0263)	(0.0230)	(0.0232)
Colleague	0.0585**	0.0421+	0.0885**	0.0317	0.0594**	0.0600**
	(0.0177)	(0.0230)	(0.0309)	(0.0192)	(0.0180)	(0.0186)
PhD Colleague	-0.00705	-0.00575	0.000708	0.00547	-0.00433	-0.0135
	(0.0212)	(0.0227)	(0.0346)	(0.0213)	(0.0216)	(0.0226)
Advisor's Coauth.	0.00284	-0.00288	-0.0104	0.0148	-0.0133	-0.0251
	(0.0321)	(0.0366)	(0.0562)	(0.0334)	(0.0323)	(0.0295)
Female					-0.00883	-0.00442
					(0.0144)	(0.0172)
Age					0.0399	0.0280
					(0.0251)	(0.0296)
Age^2					-0.000662*	-0.000515
					(0.000306)	(0.000365)
Experience					0.00386+	0.00286
					(0.00205)	(0.00252)
#Publications					0.00150+	0.00115
					(0.000763)	(0.000762)
#Citations					4.35e-05*	4.31e-05*
					(1.81e-05)	(1.79e-05)
Exp. ties	Yes	Yes	Yes	Yes	Yes	Yes
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,165	6,130	3,035	6,772	9,165	6,680
R-squared	0.057	0.061	0.061	0.056	0.075	0.078

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

3.5.1 Impact of connections on success

I estimate equation 3.1 using OLS[4], results are reported in Table 3.4, Column 1. In line with the findings of Zinovyeva and Bagues (2015) and Bagues et al. (2015) and Bagues et al. (2017), the presence of an acquainted evaluator in the committee has a significant positive effect on candidates' chances of success.

This effect varies according to the type of the connection, suggesting that stronger connections (having worked together on an article or during the PhD, in a mentor–mentee relationship) are potentially either better informed and/or more biased. The presence of a coauthor in the committee leads to a 11.7 percentage points increase

[4]There has been recent debate on the respective merits of logistic regression and linear probability models (Mood, 2010; Angrist J. D. and J. S. Pischke, 2008). Logistic regression provides a better functional form, especially near the 0 or 1 borders, but its constant variance may call into question the comparison of parameters from one regression to another. I also tested these relationships with logistic regression (see Appendix C), finding very similar results.

in the applicants' likelihood of success. Having the advisor or a colleague in the committee has a smaller but also significant positive effect (11.5 and 5.8 percentage points, respectively). I find no significant effect for the variables regarding the "weaker" connections: whether a member of the committee is affiliated to the PhD university or whether he is a candidate's advisor's coauthor.

In Columns 2 and 3, I present estimation results separately for exams for AP positions and for FP positions. The effect of connections is relatively stronger in exams for FP positions, suggesting that information asymmetries and/or biases are more important at higher stages of the academic career.

I estimate the model for the subsample of candidates who obtained their PhD in Italy (Column 4) in order to address the potential existence of unobserved connections by candidates who earned their doctoral degree abroad. The estimated coefficients show that for this group the most relevant connection is the one with the advisor and that the coefficient for colleague evaluators is not statistically different from zero.

In order to get a better understanding of the magnitude of the effects of the connections, I add to the model a number of observable individual characteristics. As proxy for candidates' quality I control for number of publications and number of citations. I also include controls for gender, age and years of experience in Italian academia. As it is possible to see from Column 5, there is no substantial change in the impact of connections.

Finally, in Column 6, in order to stress the experimental setting, I estimate the model for the subsample of "experimental exams". Thanks to the random component built into the committees' selection process, I can apply the classical experimental feature comparing the outcomes of two groups: (a) the treated group, i.e., the applicants whose personal contact has been randomly drawn; and (b) the control group, i.e., the applicants whose personal contact has not been randomly drawn. Thus I restrict the estimates to exams where I find both treated applicants (whose acquaintances have been drawn) and control applicants (whose acquaintances have not been drawn). Note that this specification only takes into account candidates who have at least one connection among the pool of eligible evaluators, so the sample size decreases to 6,680 observations. This empirical strategy is also less efficient since it does not take into account all the available information. Nevertheless, the estimated coefficients are qualitatively similar to the baseline estimations.

Since, as said, a few evaluators resigned and were randomly replaced in the final committees, following Bagues et al. (2015), I measure the presence of connections in these final committees. Hence, I estimate the following equation.

$$Y_{i,e} = \beta_0 + \mathbf{C}^{f}_{i,e}\beta_1 + \mathbf{E}_{i,e}\beta_2 + \epsilon_{i,e} \tag{3.2}$$

To account for the potential endogeneity of the replacement of these evaluators, I instrument the final composition of the committees ($\mathbf{C}^{f}_{i,e}$) using the initial composition as determined by the random draw. In Table 3.5 I present the results for the estimation of equation 3.2 using instrumental variables technique (IV)[5]. As it is possible to notice, the IV coefficients are slightly larger in absolute terms but their magnitude is statistically similar to the ITT estimates presented above.

Candidates' and committees' research quality

A question of interest is to know whether the network effects described above vary with the productivity profiles of both the candidates and the committee members. To answer these questions, first I examine how the impact of connections on success varies depending on researchers' observable research productivity as measured by their H-index. Table 3.6 reports results for both ITT and IV estimates.

It is very interesting to notice that, even if the magnitude of the effect is similar across the three groups, having the PhD advisor in the committee is significantly more beneficial for better-published researchers, in the medium and top productivity tercile. On the other hand, the magnitude of the effect of having a coauthor in the evaluating committee is 2.5–3 times bigger for researchers in the lower productivity tercile.

For what concerns the second question, I retrieved information about number of publications and number of citations of the jury members. For both measures of productivity (publications and citations), I split the sample in two: the subsample

[5] First-stage estimates can be found in Table C.2 in Appendix C

TABLE 3.5: Impact of connection on candidates' success (IV)

	(1) All	(2) AP	(3) FP	(4) PhD ITA	(5) All	(6) Experimental
Advisor	0.127**	0.107**	0.164**	0.111**	0.125**	0.131**
	(0.0347)	(0.0354)	(0.0625)	(0.0342)	(0.0346)	(0.0354)
Coauthor	0.129**	0.127**	0.147**	0.101**	0.125**	0.125**
	(0.0246)	(0.0270)	(0.0439)	(0.0276)	(0.0240)	(0.0242)
Colleague	0.0663**	0.0480+	0.0997**	0.0360+	0.0673**	0.0681**
	(0.0200)	(0.0261)	(0.0341)	(0.0216)	(0.0202)	(0.0208)
PhD Colleague	-0.00699	-0.00587	0.000856	0.00556	-0.00427	-0.0135
	(0.0211)	(0.0225)	(0.0343)	(0.0211)	(0.0214)	(0.0225)
Advisor's Coauth.	0.000708	-0.00431	-0.0153	0.0133	-0.0155	-0.0275
	(0.0317)	(0.0361)	(0.0556)	(0.0330)	(0.0318)	(0.0291)
Female					-0.00894	-0.00456
					(0.0143)	(0.0171)
Age					0.0398	0.0278
					(0.0249)	(0.0293)
Age^2					-0.000660*	-0.000514
					(0.000304)	(0.000362)
Experience					0.00389+	0.00291
					(0.00204)	(0.00251)
# Publications					0.00149*	0.00115
					(0.000755)	(0.000753)
# Citations					4.38e-05*	4.34e-05*
					(1.78e-05)	(1.76e-05)
Exp. ties	Yes	Yes	Yes	Yes	Yes	Yes
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,165	6,130	3,035	6,772	9,165	6,680
R-squared	0.056	0.060	0.061	0.056	0.074	0.077

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

TABLE 3.6: Impact of connection on candidates' success (by candidates' productivity)

	ITT			IV		
	Low prod. (1)	Medium prod. (2)	High prod. (3)	Low prod. (4)	Medium prod. (5)	High prod. (6)
Advisor	0.0924	0.106*	0.0932*	0.105	0.114*	0.104*
	(0.0683)	(0.0496)	(0.0390)	(0.0729)	(0.0534)	(0.0429)
Coauthor	0.175*	0.157**	0.0593*	0.180*	0.162**	0.0741**
	(0.0826)	(0.0389)	(0.0242)	(0.0815)	(0.0390)	(0.0276)
Colleague	0.0644*	0.0188	0.0672*	0.0737*	0.0203	0.0780*
	(0.0275)	(0.0290)	(0.0330)	(0.0323)	(0.0314)	(0.0379)
PhD Colleague	0.0132	-0.00404	0.00109	0.0134	-0.00342	9.74e-05
	(0.0368)	(0.0376)	(0.0251)	(0.0364)	(0.0371)	(0.0253)
Advisor's Coauth.	-0.0267	0.0519	-0.00843	-0.0271	0.0512	-0.0116
	(0.117)	(0.0596)	(0.0326)	(0.115)	(0.0591)	(0.0323)
Female	0.0208	-0.0167	0.0181	0.0205	-0.0171	0.0187
	(0.0229)	(0.0195)	(0.0165)	(0.0227)	(0.0193)	(0.0163)
Age	-0.0413	0.110**	0.0482	-0.0398	0.111**	0.0464
	(0.0415)	(0.0415)	(0.0391)	(0.0412)	(0.0411)	(0.0387)
age^2	0.000333	-0.00152**	-0.000691	0.000315	-0.00153**	-0.000671
	(0.000506)	(0.000499)	(0.000487)	(0.000503)	(0.000493)	(0.000481)
Experience	-0.000491	0.00311	0.00511*	-0.000502	0.00314	0.00520*
	(0.00437)	(0.00326)	(0.00228)	(0.00433)	(0.00322)	(0.00228)
#Publications	0.00635	0.00129	-8.92e-05	0.00638	0.00130	-8.92e-05
	(0.00520)	(0.00178)	(0.000680)	(0.00513)	(0.00176)	(0.000667)
#Citations	5.36e-05	3.45e-05	3.73e-05*	5.30e-05	3.41e-05	3.78e-05*
	(0.000178)	(4.48e-05)	(1.80e-05)	(4.67e-05)	(4.07e-05)	(1.20e-05)
Exp. ties	Yes	Yes	Yes	Yes	Yes	Yes
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,001	2,654	3,510	3,001	2,654	3,510
R-squared	0.058	0.095	0.041	0.120	0.071	0.050

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

TABLE 3.7: Impact of connection on candidates' success (by committees' productivity)

	HighPub (1)	HighCit (2)	LowPub (3)	LowCit (4)
Advisor	0.172**	0.149*	0.0794*	0.0994**
	(0.0535)	(0.0637)	(0.0372)	(0.0332)
Coauthor	0.0942**	0.106*	0.116**	0.109**
	(0.0316)	(0.0403)	(0.0285)	(0.0264)
Colleague	0.0649*	0.0726*	0.0447*	0.0548*
	(0.0247)	(0.0271)	(0.0212)	(0.0222)
PhD Colleague	-0.0215	-0.0241	0.00310	-0.00812
	(0.0379)	(0.0415)	(0.0252)	(0.0250)
Advisor's Coauth.	0.0423	0.0164	-0.0549	-0.0396
	(0.0354)	(0.0305)	(0.0449)	(0.0456)
Female	-0.0187	0.00480	-0.00495	-0.0128
	(0.0172)	(0.0143)	(0.0194)	(0.0202)
Age	0.0856*	0.0485	0.00735	0.0339
	(0.0349)	(0.0366)	(0.0319)	(0.0324)
Age^2	-0.00118*	-0.000717	-0.000298	-0.000611
	(0.000430)	(0.000455)	(0.000390)	(0.000394)
Experience	0.00187	0.00242	0.00563+	0.00446
	(0.00232)	(0.00324)	(0.00298)	(0.00283)
#Publications	0.00177*	0.00139*	0.00190+	0.00223+
	(0.000732)	(0.000593)	(0.00110)	(0.00122)
#Citations	5.01e-05*	5.23e-05**	3.63e-05	2.67e-05
	(2.04e-05)	(1.79e-05)	(1.80e-05)	(2.08e-05)
Exp. ties	Yes	Yes	Yes	Yes
Area FE	Yes	Yes	Yes	Yes
Observations	3,838	3,236	5,327	5,929
R-squared	0.094	0.083	0.069	0.077

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

of the exams with the most prolific committee members, and the one containing the least prolific ones[6]. Table 3.7 lists the ITT results for the four groups.

Results suggest that both high productivity and low productivity committees takes into account publication quantity, but only the first type of committees gives value also to publication quality. Furthermore, both types of committees are influenced by the network variables. However, for the more productive committees these variables are of higher magnitude and more significant. This contrast with the results of Durante et al. (2009) who show that the higher quality of the jury members (in terms of publications) the smaller the influence of networks.

[6]I take into account the average productivity of the committees inside each macro-scientific area.

TABLE 3.8: Future performance

	#Publications		#Citations		Promoted	
	(1)	(2)	(3)	(4)	(5)	(6)
Connection	1.269**		13.13**		0.0119	
	(0.350)		(4.174)		(0.0187)	
Advisor		1.136		19.46		0.0684+
		(0.879)		(13.70)		(0.0370)
Coauthor		0.977		6.488		-0.0456
		(0.370)		(10.86)		(0.0302)
Colleague		0.251		1.156		0.0190
		(0.430)		(4.876)		(0.0275)
PhD Colleague		1.543*		15.43*		0.0174
		(0.651)		(6.303)		(0.0197)
Advisor's Coauth		-0.658		0.0235		0.00716
		(1.192)		(16.61)		(0.0392)
Observations	5,748	5,748	5,748	5,748	5,748	5,748

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

3.5.2 Long-term efficiency of connections

Connected candidates are more likely to be habilitated, this could be due either to the fact that connected evaluators are better informed about acquainted candidates quality, or that they are biased in favor of them. In the first case, candidates promoted by an informed evaluator, unlike those promoted by a biased one, are expected to perform better in some characteristics which were unobservable to other evaluators at the time of the decision.

In order to disentangle this mechanism, I restrict the sample to promoted candidates only and proxy these unobservable characteristics using information about future performance. Following Zinovyeva and Bagues (2015), I thus estimate equation 3.3 using as dependent variable several measures of candidates' future performance.

$$p_{i,e} = \beta_0 + \mathbf{C}_{i,e}\beta_1 + \mathbf{E}_{i,e}\beta_2 + \epsilon_{i,e} \tag{3.3}$$

I measure future performance in two different ways. First, I look at indicators of research productivity (number of publications and number of citations) during the period following the habilitation. Second, I analyze candidates' future performance in peer evaluations. In particular I consider whether they manage to be promoted to an AP or FP position after having received the national scientific habilitation.

In Table 3.8 I both report the coefficients for every type of connected evaluator and for a dummy which takes value 1 if any type of connection is present in the committee. I omit the other covariates to make results more readable.

Conditional on their past research production, candidates habilitated by a connection turn out to be more productive in the future for number of publications and the number of citations received, even if taken singularly, most variables for acquaintances are not significant, with exception of colleagues at the time of the PhD. The future quality of the research production of candidates habilitated by a committee in which one of the evaluators was a PhD colleague is significantly higher relative to other promoted candidates. This result points towards a superior information about candidates' quality in the case of older acquaintances, who dates back to the PhD education. Results for promotion show that while coauthor and colleagues evaluators are not significant, having received the national habilitation from a scientific committee in which the advisor was an evaluator, increases the promotion probability by 6.8 percentage points.

This evidence is consistent with a situation where evaluators take advantage of their better information about the quality of connected candidates. Furthermore, being habilitated by an acquainted evaluator has a significant impact of the probability of being promoted at the university level in the case of the PhD advisor.

3.6 Conclusions

In this Chapter I analyze the effect of connections in the context of scientific committees in Italy. I focus on the first round of national scientific habilitation which took place in 2012.

In this context, more than one fourth of the candidates were evaluated by a commission which included an acquainted evaluator, such as their doctorate advisor, a coauthor or a colleague. The presence of these connections in the committee increases by 17 percent (in the cases of advisor and coauthor evaluator) and 10 percent (in the case of colleague evaluator) the chances of success of candidates. I also find that these candidates are more productive in the five-year period after the evaluation with respect to other promoted candidates, and that those promoted with the PhD advisor as a member of the committee are more likely to succeed in being promoted at the university level. I also explore the effect of weaker connections: being evaluated by an advisor's coauthor or by a colleague at the time of the PhD. Apparently these links have no impact on the selection process, whereas PhD colleagues seem to be better informed about unobservable quality of acquainted candidates which later translates in a higher productivity performance.

From this empirical evidence it seems that candidates with acquainted evaluators (in the case of advisors, coauthors and colleagues) do not receive a preferential treatment, but rather that the evaluators have a better knowledge of the unobservable characteristics and quality of connected candidates with respect to other evaluators who had no previous contact with candidates or just a weaker form of connection (the case of advisor's coauthors).

According to these findings, no further conflict of interests rules should be implemented in the habilitation process. Of course, results may vary depending on the institutional framework, but in the setting of Italian academia, the presence of acquaintances in evaluation committees tends to increase the research quality of promoted candidates.

This work also presents evidence that the chance of having a supporter in an evaluation committee is a key element for progressing in academia. Filling the gap in the literature concerning Italian academia, I identified former PhD advisors as among the strongest potential sponsors. This effect is similar to the one identified by Zinovyeva and Bagues (2015) for Spanish national recruitment and by Combes et al. (2008) for French *concours d'agregation*, so the results presented here may be extrapolated beyond the national case and contribute to the debate over European academic labor market (Aghion et al., 2010).

Finally, even if the results obtained in this work are established for academia (especially in other European countries, such as France and Spain), one may wonder whether they hold true also beyond this labor market. The particular recruitment structure of academia, where peers who have some chance of having worked with the candidate in the past are in charge of the evaluation and hiring decisions, maximizes the opportunities for strong ties influence. Future research efforts could measure the relative importance of these influence mechanisms also in the case of corporate firms.

Chapter 4

Gender Discrimination in Academic Promotions in Italy

4.1 Introduction

The aim if this Chapter is to understand if gender makes a difference in the path to promotion in Italian universities, in the context of the two most recent reforms of Italian academic labor market: Berlinguer's reform in 1998 and Gelmini's reform in 2010[1]. The first reform decentralized the academic recruitment, while the second one introduced a two-step procedure with a national habilitation and local concourses.

In the first case, in a "quasi-experimental" research framework, I study the differences between individual research productivity measures of researchers hired before and after the reform, focusing the analysis on gender differences. In the second case, I exploit data on the Italian habilitation concourse to analyze gender differences in the probability of success and in the willingness to compete in contests.

In Europe women account for 46% of PhD graduates, 37% of associate professors and only a mere 20% of full professors (European Commission, 2014). The US displays similar patterns and in Japan the gender imbalance is even larger (Council, 2010; Geuna and Shibayama, 2016).

The situation in the Italian academia is not that different. Women account for

[1] From the name of the two former University Ministries.

TABLE 4.1: Professors in Italian academia by gender and scientific field (MIUR)

	Assistant Prof.		Associate Prof.		Full Prof.	
Area	M	F	M	F	M	F
Science	51%	49%	60%	40%	78%	22%
Economics Statistics	52%	48%	61%	39%	78%	22%
Medicine Veterinary	55%	45%	71%	29%	85%	15%
Humanities Law	46%	54%	52%	48%	68%	32%
Architecture Engineering	70%	30%	77%	23%	88%	12%
Social Sciences	53%	47%	61%	39%	75%	25%
Tot	54%	46%	63%	37%	78%	22%

53% of PhD graduates, and are particularly represented in humanities and life sciences, 64% and 63% respectively; but are still the minority in hard science and engineering fields, 41% and 37% (AlmaLaurea 2015). As it is possible to see from Table 4.1, gender differences among Italian associate and full professors are in line with the rest of Europe and the situation is particularly critical in engineering and medicine&veterinary scientific fields.

There are several possible explanations for the persistent underrepresentation of women in top positions in academia. Gender differences in promotion rates might reflect differences in productivity, perhaps due to the existence of gendered roles at the household level or the lack of female mentors and role models (Blau et al., 2010). Furthermore, some authors have pointed out that women are less likely to apply for promotions (Bosquet et al., 2013; De Paola et al., 2015), maybe due to existence of gender differences in the preference for competitive environments (Niederle and Vesterlund, 2007) or in bargaining abilities in the labor market (Blackaby et al., 2005). Moreover, women seem to devote more time to tasks that are socially desirable but might not be taken into account in promotion decisions (Vesterlund et al., 2014).

These problems have reached policy makers in many countries, including Italy, which in fact applied profound modifications to the Italian academic recruitment system over the past two decades, with the objective of increasing transparency and meritocracy. The aim of this study is to explore the effect of gender on promotion through different institutional changes to determine whether these reforms enhance the quality of selected researchers, assuring a higher gender equality of the hiring process.

First, I focus on the impact of the 1998 decentralization reform of the Italian university system on research outcomes for candidates publishing in international journals, bearing in mind that all other aspects of the system remained unchanged over the last two decades (salary benefits, university funding mechanisms etc.). From a theoretical viewpoint, it is possible that the decentralization of recruitment mechanisms reduces the incentives for candidates to produce international research outcomes (conference papers, journal articles etc.) and/or to submit papers to higher-quality scientific journals (which usually implies longer publication times and lower acceptance rates). More local recruitment management could generate the expectation that less stringent requirements will be applied. This consideration would be most crucial for applicants to assistant professorships and for associate professors applying for full professorships rather than for newly appointed assistant professors.

Then I turn the attention to the 2010 reform which introduced in Italy a national habilitation for university professors, following the model of several European countries (eg. Spain and France). For this reform, I analyze the gender differences in the probability of taking part and being selected using evidence from the first qualification evaluations in Italy.

4.2 Institutional background

The academic labor market in Italy is quite different from other non-European countries, especially compared to the US. The main feature is that this labor market is highly regulated. Since the employees at Italian universities are civil servants, the wage, the contract length, the tasks (teaching load and others) are determined by law and cannot be bargained at the local level. In order to understand the factors which determine success (the probability of becoming tenured) in this market it is crucial to understand these regulations.

The Italian academic system is composed of 97 universities (30 private and 67 public) and 9 higher education institutions. The latter usually dispense only masters and PhD courses, being more research oriented than most of the other universities. Three out of the 67 public universities are polytechnics. Eleven out of the 30 private institutions are distance-learning universities. The university system is divided into 372 sectors of discipline (*settore scientifico disciplinare* – SSD), grouped into fourteen research areas, as designated by the Italian National University Council (CUN). Sectors of discipline are categorized for homogeneity within each research area, and the selection of research candidates is conducted by recruitment commissions within each academic discipline in both national and local recruitment systems.

The Italian academic system has three main positions called "Ricercatore universitario" (assistant professor), "Professore associato" (associate professor) and "Professore ordinario" (full professor). Each professor working at an Italian university is then categorized by a level of arrangement (full professor, associate professor and assistant professor) and by one out of 372 sectors of discipline. Each vacancy is coded in a standardized format, and each filled position becomes tenured after a review conducted three years after hiring. Salaries in public universities are set by law and vary only by level of arrangement and seniority. Schools and departments are prevented from differentiating wages among professors, linking payment to research productivity and/or teaching loads. As a consequence, in addition to celebrity and funds attraction, the strongest incentive to scientific productivity for individuals working in academia derives from expected promotion (being hired as assistant professor, being promoted associated or full professor). Given the public nature of the employment contracts, university professors can only be hired through public competitions that should grant publicity of the vacancy, selection of the selecting committee based on objective criteria, transparency of the selection process.

Since 1979, standardized competitions were held to hire assistants, associate and full professors, and until 1998, almost all academic recruitment was substantially centralized. Despite the legislative prescription of one concourse every two years, a three to four years interval occurred. National commissions of five members were chosen by lot within a pool of elected professors (from a pool of 15[2]) belonging to the same discipline. Commissioners declared which of the candidates had the qualifications to be promoted to associate/full professorship. Eligibility was given to a number of candidates greater than the available positions (usually 20% higher) for each discipline. Universities with opening positions drew by multilateral bargaining

[2]See DL 31/1979 and DPR 382/1980.

between them from the list of eligible applicants to fulfill their vacancies. Starting in 1999[3], recruitment procedures became entirely local, and each university could hold its own selection procedure (both for assistants, associates and full professors). Local commissions were comprised of five members: one belonging to the institution itself and the four others elected by the full set of Italian professors of that discipline. After 2005, a new reform act[4] established that the commission's members had to be drawn by lot in a pool of professors of three times the size of the local commission, elected by popular vote amongst the discipline's affiliates. The commissions initially declared three qualified candidates for each concourse, but moved to two between 2007 until 2008, and only one thereafter. In the following years, universities with open vacancies could hire any candidate who had obtained a qualification. The most recent reform[5] (in 2010) introduced a two-step procedure with a national habilitation and local concourses to recruit professors. In the first stage, candidates to Associate Professor (AP) and Full Professor (FP) positions are required to qualify in a national-level exam. Evaluations are conducted separately in 184 scientific fields designed by the Ministry of Education. A positive evaluation is valid for four years while a negative one implies the exclusion to participate in further national evaluations during the following two years. Qualified candidates can participate in the second stage, which is managed locally by each university. This may consist in either an open competition or the assessment of an internal candidate.

I do not have data on concourses (nor on commissions, candidates, winners or "idonei") in period I am considering, but I can observe promotions (from assistant professor to associate and from associate to full professor) and new entrances. Since concourses results do not translate immediately in new hirings, but the universities have to call the winners or "idonei" to fill their vacancies, I can interpret promotions and new entrances as the results of concourses. Figure 4.1 shows the number of newly promoted associate and full professors and the share of women promoted in Italian academia by year.

[3]"Berlinguer reform", DPR 390/1998.
[4]"Moratti reform", DL 230/2005.
[5]"Gelmini reform", L 240/2010.

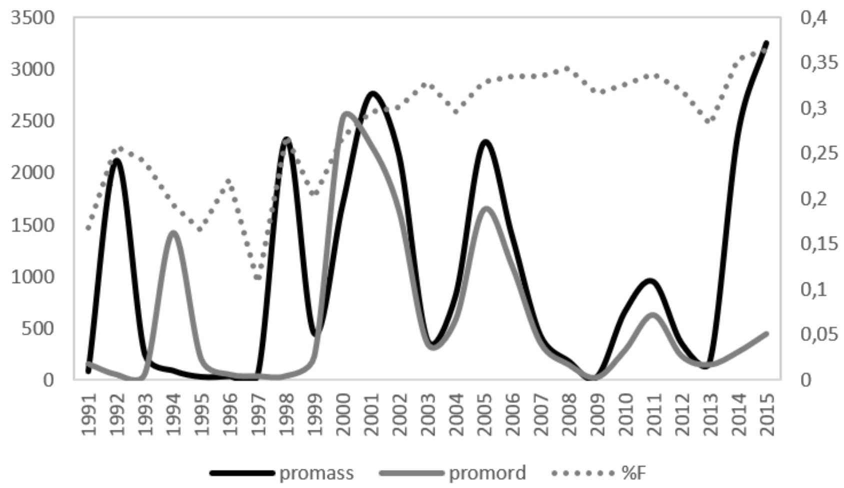

FIGURE 4.1: Promoted associate and full professors and share of promoted women (MIUR)

4.3 Theoretical expectations

Literature shows that universities are "gendered organizations" (Acker, 1990). Gender discrimination along academic careers has been widely documented: despite academic systems and promotion processes greatly vary among countries, there are some common issues.

When focusing on career progression and tenure, scholars have shown that the criteria of excellence and meritocracy are gendered at the detriment of women (Deem, 2009; Brink et al., 2010). Perna (2001) analyzes US data and shows that women are less likely to be appointed as professors, even when controlling for productivity, at 4-years Colleges. Danell and Hjerm (2013), in their analysis of the careers of PhDs in Sweden between 1995 and 2010, show that women have 37% lower chance to become professors than men, and that gender differences in promotion rates have not decreased in time. Weisshaar (2017) conducted a study across a sample of academics in Computer Science, English and Sociology in the US, and shows that neither productivity, nor organizational characteristics (department size, ranking, number of women in faculty) explain the gender gap in tenure. Van den Brink and Benschop, 2014 try to explore the processes that hinder women to be promoted to professor in the Netherlands, showing the relevance of networks along recruitment processes. They show that men in senior positions act as gatekeepers and favor other men. Furthermore, when it comes to support junior academics, men tend to support other men: a phenomenon explained through the concept of homophily, i.e. the perceived similarity among the members of a network.

Turning to the research focused on gender differences in promotion processes in Italy, literature presents mixed findings. Abramo et al. (2016) analyze the 2008 competition for associate professors, finding no gender difference in relation to favoritism. A previous study (Abramo et al., 2015) shows that the gender of the president of the evaluating committee has an impact: women in committees are more likely to evaluate according to pure merit in comparison to men evaluators. De Paola et al. (2015) focus on chemistry and economics, and show that women are likely to be less promoted when the committee is composed exclusively by men. While a recent study by Bagues et al. (2017) on the first round of the ASN (2012–2014) shows that having women in the selection committee does not favor women's applicants. if anything, having women in the committee means that men evaluators become more severe towards women's applicants. They calculated that each additional woman evaluator in the committee decreases by 2% the chance of women applicants to be promoted at parity of scientific productivity.

Summing up, although a substantial literature has examined the promotion gap across genders, a clear explanation is still lacking. Women may be less likely to be promoted either because they apply for promotion less often than men or because of a lower probability of being promoted conditional on being a candidate. These differences may in turn have different possible groups of candidate explanations: discrimination, different attitudes in and towards the promotion process itself and differences across genders in the costs of or rewards from promotion. The different types of possible explanations for gender discrimination in academia are summarized hereafter.

Discrimination The first possible explanation for the promotion gap across genders is simply that those making the promotion decisions discriminate against women. It implies a negative impact of being female on the conditional probability of promotion.

Expected discrimination If women believe that they will be discriminated against, then they will be less likely to apply for promotion than men irrespective of whether they are actually discriminated.

Under-performance Female candidates may also be less likely to succeed if they tend to under-perform in contests, as shown by Gneezy et al. (2003).

Self-selection Women may differ in the costs of and rewards from promotion and this can make them less willing to apply for promotion, causing “self-selection” out of the promotion race. For example, the opportunity cost of time could be higher for women if they undertake a disproportionate amount of domestic work, while differences in intra-household bargaining power could make it harder for female than for male professors to impose the cost of moving on their families.

Unwillingness to compete Women may also choose not to apply for promotion because they are less inclined to compete in tournaments than men, as indicated by experimental evidence (see Niederle and Vesterlund, 2007; Datta Gupta et al., 2013.

Preferences Men and women may have different preferences over department prestige and income. Women have a lower marginal utility of income because, often, they are the second earner in the household and they could also have stronger preferences for department quality. Thus, female professors may be less willing to apply for promotion if they are in a top department than if they are not (McDowell et al., 2001).

4.4 Berlinguer’s reform (1998)

4.4.1 Data and methods

Data for this analysis were collected from two primary sources: MIUR (Italian “Ministry of Education, University and Research”) - and the “bibliographic” Web version of Elsevier Scopus database.

On the MIUR website, information regarding academic positions, disciplinary areas, university affiliation and personal information, such as birth year and gender, are available on-line from 2000 to 2015. I obtained data on academic careers before 2000 from Cineca, a MIUR agency which collects administrative data on personnel as well as on competition for professorship in Italy.

Scopus, published by Elsevier, indexes a greater number of journals (12.850, including 500 open access journals) within the medicine, technical and social sciences and offers systematic, quantifiable statistical information based on citation data. Using Scopus API, I retrieved Scopus Author-Id (a unique identifier for academics created by Elsevier) for all academics who published at least one article in international scientific journals, and then downloaded the articles. In this way I avoid the risk of downloading the scientific production of homonyms authors which are not part of the database[6].

I then employ a multi-step matching procedure to assign the corresponding author identifying codes to each research product in the bibliometric dataset.

The methodological approach I employ to evaluate the impact of different selection mechanisms caters to the specific research question I ask: "Is there a causal effect of local (vs. national) recruitment programs on the subsequent gender gap in selected academics?"

I focus on evaluating the effects of a shift to decentralized selection mechanisms in terms of gender and the subsequent research productivity average level of the outcome and its time trend. The treatment status can be considered as the exposure of an individual to local selections instead of national ones. The problem is that I can observe almost one of these states for each individual of interest. Indeed, individuals who are exposed to local selection programs are by definition (due to a specific time constraint: they were selected after 2000) different from those who are exposed to national recruitment programs. These differences may invalidate the causal comparison of the impact of decentralization on future research productivity outcomes.

Recent studies in the econometric literature of program evaluation (Imbens and Wooldridge, 2009) and classic methodological research on causal inference (Rosenbaum and Rubin, 1983) from observational studies (where investigators have no control over the treatment assignment) suggests the use of parametric methods, such

[6]For details see Chapter 1.

as propensity score, or non-parametric strategies to accommodate general heterogeneity between two groups of individuals in estimating the treatment effects and to increase precision of the estimates. The treated (in the empirical application all the researchers selected with local mechanisms) and control (the ones selected with national mechanisms) groups may have significant differences in their observed covariates (scientific discipline and research productivity outcomes) that could lead to biased estimates of the selected effect.

Thanks to the dataset I am able to perform a detailed matching among the individuals, in order to create the treated and the control samples. The pre-treatment variables that I will use are: age, disciplinary area, covariates regarding bibliometric indicators (impact factor and number of publications and citations), years of experience (since the first publication or the first appointment in Italian academia) and the dummy variable $PhDITA$ which takes value one if the academic obtained the doctoral degree from an Italian university. I obtained the latter variable from the National Library of Florence repository, which provides information on all the doctoral dissertations discussed in Italy since the first cycle (1986) to 2006 (see Chapter 1 for details).

Matching methods

Methodologically, I use a standard propensity score matching approach and a recently proposed non-parametric matching algorithm called Coarsened Exact Matching (CEM) (Iacus et al., 2012), which has also already been used in the economics of education literature (Verzillo, 2013). Balancing is set at the first stage and then the number of matching units is a consequence.

CEM is a matching method recently introduced (see Iacus et al., 2012) to improve causal inference controlling for the confounding influence of covariates in observational studies. The time-dependent nature of this study, and the flexibility of CEM in estimating non-parametrically two balanced distributions of treated (locally selected professors) and controls (centrally selected) units, allows us to obtain desired counterfactuals for estimating the decentralization effect. Balanced groups avoid having the researcher control for the heterogeneity while specifying the model, meaning that simple differences in means are good estimates of the causal effect. But usually

TABLE 4.2: Promotion by gender in the pre-reform waves

	1992		1998	
	M	F	M	F
Promoted to PA	1132	467	1329	494
RU previous year	6129	4833	8035	5933
% Promoted	0.185	0.097	0.165	0.083

finding a matching solution in empirical propensity applications does not guarantee good balance to all the selected covariates. Improving balance on most of them could leave the remainders imbalanced, often introducing more bias with respect to the initial distribution.

In addition to this, propensity score matching has the drawback of violating the congruence principle, which requires congruences between data and analysis spaces metrics (the own metric of the two spaces is different). Parametric methods usually force covariates of the original data from a multi-dimensional original space in a new space defined by the propensity itself. In comparison, CEM meets the principle of not reducing the original data space, operating in the multidimensional variable space itself.

4.4.2 Empirical strategy

Following the existing literature, I adopt a difference-in-differences (DD) approach. From Table 4.2 it is possible to see that, taking into account the two promotion waves in the pre-reform period (1992 and 1998, see Figure 4.1) there is a persistent gap in the percentage of promoted males and females academics.

A variation in this gap due to the 1998 reform can thus be explored causally in a DD framework. I suppose that the probability of an academic i being in a professorship rank in year t is:

$$P(Y,i,t) = f(X_{i,t}`\beta_j) \tag{4.1}$$

where Y is being in an associate or full professor position or not and f denotes a logistic density function. The term $X_{i,t}`\beta_j$ is:

$$X_{i,t}`\beta_j = \beta_{j,0} + \beta_{j,1}Female + \beta_{j,2}Reform + \beta_{j,3}Age + \beta_{j,4}Age^2 + \beta_{j,5}Experience \\ + \beta_{j,6}\#Publications + \beta_{j,7}AvgSJR + \beta_{j,8}PhDIta + \beta_{j,9}SciArea$$

implying that the probability of promotion is a function of age and its square, experience (calculated since the year of the first article published or the first appointment in Italian academia), the number of publications and the average quality of these publications denoted by the SCImago Journal Rank (SJR), whether or not the individual has obtained her doctoral degree from an Italian university and the broad scientific area. $Reform$ is a dummy which takes value 1 if the individual has been selected under the decentralized mechanism, with *Female* being a dummy for females, $\beta_{j,1}$ measures the differences in promotion probability for men and women with the same characteristics, while $\beta_{j,3}$ is the coefficient of the interaction of these two dummies in order to evaluate their simultaneous effect.

4.4.3 Results

For the analysis, in order to allow comparisons, I consider the academics in the fields of Architecture&Engineering, Humanities&Law and Science[7], dropping the homonyms for "surname+initial" in order to avoid problems in the attribution of the publication records.

For the first part of the empirical analysis, related to the Berlinguer's reform, in order to perform the matching among the two sample of academics, treated and control groups, I considered those who entered as assistant professors in the period 1993-1996 (1998 promotion wave) and 1997-2000 (2001 promotion wave). In total I consider 6784 academics, the 58% of whom has at least one Scopus publication in the time span considered, with huge differences across disciplines: only the 16% of the humanities&law academics has published an article on international journals (see Table 4.3).

Propensity score matching

With propensity score methods, it is important to check for the overlapping of the propensity score distributions between treated and controls. The univariate reduction imposed by propensity scores avoids the need to check for multidimensional

[7]Scientific fields considered are: 01–Scienze Matematiche e Informatiche; 02–Scienze Fisiche; 03–Scienze Chimiche; 04–Scienze della Terra; 05–Scienze Biologiche; 07–Scienze Agrarie; 08–Ingegneria Civile e Architettura; 09–Ingegneria Industriale e dell'Informazione; 10–Scienze dell'Antichita, Filologico-letterarie e Storico-artistiche; 11–Scienze Storiche, Filosofiche, Pedagogiche, Psicologiche; 12–Scienze Giuridiche.

TABLE 4.3: Professors with at least one Scopus publication by scientific field

	N	$\geq$ 1 Pub	%
Science	2954	2468	0.84
Arch&Eng	1701	1119	0.66
Hum&Law law	2129	349	0.16
Area 01	613	494	0.81
Area 02	337	252	0.75
Area 03	590	568	0.96
Area 04	196	168	0.86
Area 05	812	717	0.88
Area 07	406	269	0.66
Area 08	617	253	0.41
Area 09	1084	866	0.80
Area 10	684	80	0.12
Area 11	619	188	0.30
Area 12	826	81	0.10
TOT	6784	3936	0.81

TABLE 4.4: Population propensity scores descriptives statistics

Treated	Obs.	Mean	SD	Min	Max
0	3248	0.481	0.123	0.049	0.891
1	3536	0.558	0.138	0.019	0.991

balancing distribution between treated and controls.

Table 4.4 reports that mean propensity is larger in the post-reform group in comparison to the pre-reform group, meaning that the fitted standard logistic model is a quite good predictor of the treatment status, and generates an effective estimate of individuals propensity score. To check for the goodness of fit for the overlap, we propose (Figure 4.2) histograms of propensity score distributions for each of the two groups.

In Table 4.5 I fit the treatment effect on the balanced distribution, considering the whole sample and, in Column 5, only those cohorts of academics which are eligible to have obtained a PhD after 1986 (the number of individuals considered decreases of about 700 units).

The results show that the covariates of interest follow the theoretical expectations. Female researchers have a lower probability of obtaining the promotion, effect which seems to be mitigated by the impact of the reform, even if the effect is only slightly significant. Controlling for bibliometric measures and scientific field (Column 3), the latter effect seem no longer to play a role in promotion probability for

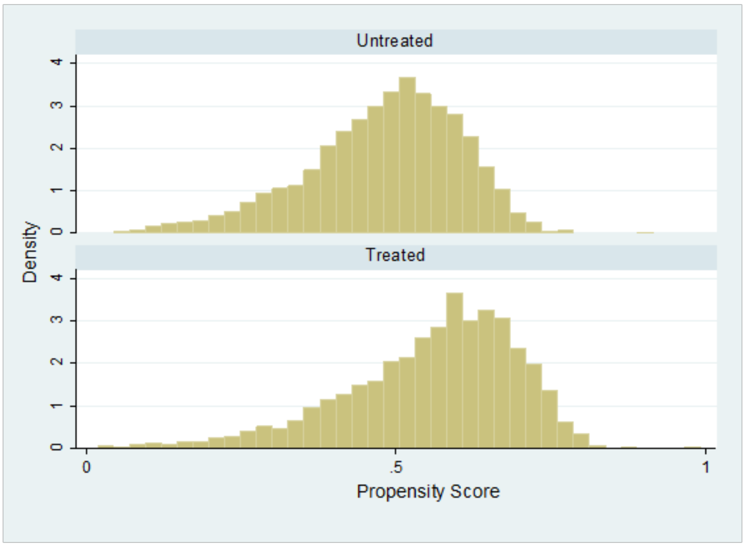

FIGURE 4.2: Propensity score histograms by treatment and control groups

TABLE 4.5: DD estimation - Marginal effects

	All (1)	All (2)	All (3)	All (4)	Restricted (5)
Female	-0.0636**	-0.0653**	-0.0504**	-0.0515**	-0.0678**
	(0.0129)	(0.0129)	(0.0128)	(0.0128)	(0.0145)
Reform	-0.00776	-0.00842	-0.0101	-0.0113	-0.0118
	(0.00912)	(0.00917)	(0.00913)	(0.00910)	(0.00956)
FemaleXReform	0.0292+	0.0300+	0.0255	0.0287+	0.0426*
	(0.0173)	(0.0173)	(0.0170)	(0.0170)	(0.0185)
Age		0.0170+	0.0159+	0.0169+	0.0328
		(0.00944)	(0.00944)	(0.00948)	(0.0247)
Age^2		-0.000203+	-0.000193	-0.000204+	-0.000438
		(0.000118)	(0.000118)	(0.000118)	(0.000338)
Experience		0.00370**	0.00674**	0.00746**	0.00826**
		(0.00108)	(0.00143)	(0.00134)	(0.00165)
#Publications			0.0210**		0.0205**
			(0.00389)		(0.00401)
#AvgSJR			-0.00361		-0.00506
			(0.00437)		(0.00458)
Productivity				0.0199**	
				(0.00332)	
PhdIta			0.0186*	0.0180*	0.0171+
			(0.00860)	(0.00859)	(0.00879)
Eng			0.104**	0.105**	0.103**
			(0.0105)	(0.0105)	(0.0109)
Hum			0.0978**	0.0992**	0.103**
			(0.0116)	(0.0114)	(0.0123)
Observations	6,496	6,496	6,496	6,496	5,853
Log Likelihood	-2217	-2208	-2112	-2111	-1883
Chi-square	35.92	55.12	245.4	248.4	244.5
Pseudo R2	0.00804	0.0123	0.0549	0.0556	0.0610

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

the whole sample, but it is significant at 10% level introducing a measure for productivity[8] which takes into account both quantity and quality of publications. Furthermore, the effect is strong and significant taking into account the reduced sample in Column 5. Seniority and experience play an important role, influencing positively the odds of being promoted, however, considering just the younger cohorts, the effect of age disappears. As expected, productivity and the dummy *PhDIta* have a positive and significant effect on promotion both for the whole and restricted samples.

Coarsened exact matching

A quick overview to the descriptive statistics for assistant professors at the year of selection in Table 4.6, evidences wide differences of bibliometric indicators within academic disciplines. Science professors in physics (Area 02) at the time of the selection had an average of almost 21 Scopus paper, with 1.68 average SJR, 26 average citation per paper and 565 cumulative citations over their careers. At the other end of the spectrum, Humanities&Law academics in classic history and literature (Area 10) have on average less than 4 papers on Scopus, with 0.84 SJR and 18 average citations per paper and 72 citations over their careers. I normalized research outcomes by sceientific areas. In order to maintain the maximum number of observations, I input zero values for all the individuals with missing values over these variables.

Applying CEM to this study means firstly to set variable-by-variable the non-overlapping intervals to coarsen original data about academics subject to the selection procedures (before and after the reform) at the year of their selection. Then I match one-to-one each stratum treated and controls units. Table 4.7 reports the description of the covariates which are set as coarsening rules.

Missing data are treated as "missing as zeros" due to the particular nature of the data. Indeed, missing values of bibliometric indicators (SCImago Journal Rank, number of Scopus paper, number of citations) reflect the absence of the author in the data and absence on Scopus is equal to 0 international papers published, with 0 citations received and 0 average SJR. A real drawback of missing replaced with

[8] *Prodictivity* consists of the sum of articles, adjusted for journal quality, through the SCImago Journal Rank, and the number of authors. The formula is the following: $\sum_{a=1}^{Ni,t} \left(\frac{SJR}{aut}\right)_a$.

TABLE 4.6: Descriptive statistics of bibliometric indicators by disciplinary area

	≥ 1 Pub	Avg Pub	Avg Cit	Avg SJR	Cit Stock
Science	2468	14.06	26.10	1.33	422.70
Arch&Eng	1119	8.42	22.38	0.99	196.72
Hum&Law	349	5.17	23.97	0.88	143.15
Area 01	494	6.20	15.46	1.06	110.82
Area 02	252	20.79	25.99	1.68	565.62
Area 03	568	18.33	27.82	1.21	582.52
Area 04	168	8.54	28.57	1.12	244.47
Area 05	717	17.70	32.00	1.72	590.07
Area 07	269	6.92	24.86	0.84	189.33
Area 08	253	5.30	21.83	1.07	118.96
Area 09	866	9.33	22.54	0.97	219.43
Area 10	80	3.83	18.11	0.84	72.15
Area 11	188	5.96	27.92	0.75	179.73
Area 12	81	4.68	20.58	1.20	128.40
Tot	3936	11.67	24.85	1.19	333.67

TABLE 4.7: Variables used for coarsening

Variable	**Description**
Birthyear	Academic's year of birth
#Publications	Number of articles published in Scopus journals
AvgCitations	Average number of citations per article
AvgSJR	Average Scimago journal rank per article
Citations Stock	Number of citation cumulated throughout the academic's career
Experience	Years of experience since first publication or first appointment in Italian academia
PhdIta	1: academic obtained the doctoral degree from an Italian university
SciArea	Academic's scientific field

TABLE 4.8: Treated and untreated units by CEM groups)

	Untreated	Treated
All	3248	3536
Matched	1449	1449
Unmatched	1799	2087

TABLE 4.9: $L1$ measure for matched and original population

$L1$ **Population**	$L1$ **Matched**
0.86261089	0.78191856

zeros could be represented by the equal 'treatment' of an author with few Scopus publications with zero SJR and 0 citations and an author without Scopus records. However, equal treatment of zero Scopus publications or few records with no SJR and no citations in the restricted sample could be considered, without a significant loss of information, acceptable. The desirable output of this procedure is a sample of balanced treated and controls. For this case, I found 1449 treated professors with one-to-one coarsened exactly matched controls over 3248 potentially possible 1:1 couples (Table 4.8).

The selected sample population is now composed of comparable sub-groups of individuals (selected before and after the reform) with similar levels of bibliometric indicators (according to the coarsened intervals settled as before) and operating their research effort in the same disciplinary areas. Balancing details for the performed CEM can be found in Appendix D.

Iacus et al. (2012) propose a measure of imbalance ($L1$) that is the semi-sum of the absolute differences between relative frequencies of treated and controls for each identified strata in this case. $L1$ for the entire population is close to 0.9 (highly unbalanced distribution of treated and controls). This means that a substantial number of cells in the multidimensional matrix have zero controls (or treated). Comparing the $L1$ of the matched population with the previous one provides evidence of the unbalanced reduction due to CEM. $L1$ is equal to 0.78 after CEM, this means higher rate of balancing between treated and controls (Table 4.9).

Table 4.10 shows the estimated results for the DD model of equation 4.1. The variables of interest maintain their significance throughout all five model specifications. Females have a lower probability of being promoted but coefficient β_3, which represents the DD estimator, shows that the effect of the reform on the probability of

TABLE 4.10: DD estimates – Marginal effects

	All (1)	All (2)	All (3)	All (4)	Restricted (5)
Female	-0.0591**	-0.0615**	-0.0495*	-0.0486*	-0.0546*
	(0.0201)	(0.0202)	(0.0199)	(0.0199)	(0.0225)
Reform	0.0102	0.00699	0.0109	0.0105	0.00561
	(0.0134)	(0.0140)	(0.0141)	(0.0141)	(0.0162)
Female*Reform	0.0682**	0.0708**	0.0667**	0.0658**	0.0742*
	(0.0251)	(0.0251)	(0.0247)	(0.0246)	(0.0296)
Age		0.00403	-0.00229	-0.00140	0.211*
		(0.0138)	(0.0135)	(0.0136)	(0.0958)
Age^2		-3.97e-05	2.59e-05	1.76e-05	-0.00314*
		(0.000174)	(0.000169)	(0.000169)	(0.00140)
Experience		0.00454*	0.00634*	0.00893**	0.00391
		(0.00196)	(0.00270)	(0.00246)	(0.00383)
#Publications			0.0498**		0.0598**
			(0.0102)		(0.0115)
AvgSJR			-0.0227*		-0.0239+
			(0.0113)		(0.0122)
z_prod				0.0316**	
				(0.00734)	
PhdIta			0.00674	0.00422	-0.00340
			(0.0121)	(0.0121)	(0.0139)
ArcEng			0.0834**	0.0857**	0.0670**
			(0.0163)	(0.0162)	(0.0175)
HumLaw			0.0871**	0.0911**	0.0717**
			(0.0179)	(0.0174)	(0.0202)
Observations	2,866	2,866	2,866	2,866	1,870
Log Likelihood	-903.8	-900.7	-862.7	-865.8	-515
Chi-square	17.75	23.91	99.91	93.74	86.06
Pseudo R2	0.00972	0.0131	0.0547	0.0514	0.0771

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

women being promoted is significant at the 1 percent level with the reform having a positive effect.

Adding controls from specifications 2 to 3, it is possible to see that, as expected, experience and number of publications are valuable characteristics and have a positive impact on the probability of being promoted, while the average quality of the publications, measured by the SJR, seems to have a negative impact. In Column 4 I estimate the model using a productivity measure (see Footnote 8), that takes into account both quantity and quality of the publications, which has a positive impact on the odds of being promoted. Surprisingly, the dummy $PhDIta$ is positive, even if only slightly, and thus follows the theoretical expectations, but it is not significant. As a robustness check, I repeated the estimation of equation 4.1 only on those academic cohorts which are eligible to have obtained a doctoral degree from 1986.

Column 5 reports the results, which are similar to those already discussed above. The dummy $PhDITA$ is now negative but still not statistically significant.

4.5 Gelmini's reform (2010)

4.5.1 Data and methods

For the evaluation of the more recent reform, the sample consists of the entire population of Italian academics provided by the MIUR for the years 1990 to 2015 in the fields of Science, Architecture&Engineering and Humanities&Law. For each individual I have information on age, rank, publication stock gathered from Scopus (see above) and affiliation. I use information from the first round of the Italian national scientific habilitation which, for all the 184 scientific fields, include information on applicants, evaluators and the final outcome of the evaluation process.

Although all academics in the rank of assistant or associate professor prior the habilitation are eligible for habilitation, I introduce an age limit to define potential candidate, in order to avoid having in the pool a large number of candidates unlikely to apply. I hence consider as potential candidates for habilitation for associate professorship academics aged between 27 and 50, and as potential candidates for habilitation for full professorship academics aged between 28 and 60. The lower bounds are given by the age of the youngest candidate, while the upper bounds imply that I consider 88% and 80% of the candidates for habilitation for associate and full professorship, respectively.

I use these data first to evaluate the probability of habilitation, taking into account age, publications and gender, in order to measure the difference in promotion probability for men and women with the same characteristics; in addition to that, I consider also the probability that an individual applies for habilitation, in order to see to what extent the difference in the probability of being habilitated is due to women being less likely to enter the habilitation concourse.

4.5.2 Empirical strategy

In this case I am first able to consider the probability that an academic i applies for habilitation, which is:

TABLE 4.11: Academics by rank and gender in 2015

	Tot	%	Female	%	Male	%
			Architecture&Engineering			
Assistant prof.	3088	0.357	920	0.483	2168	0.321
Associate prof.	3400	0.393	747	0.392	2653	0.393
Full prof.	2158	0.249	235	0.123	1923	0.285
			Science			
Assistant prof.	5488	0.400	2808	0.507	2680	0.328
Associate prof.	5219	0.380	2064	0.372	3155	0.386
Full prof.	2990	0.218	657	0.118	2333	0.285
			Humanities&Law			
Assistant prof.	2337	0.364	1098	0.465	1239	0.306
Associate prof.	2261	0.352	875	0.370	1386	0.342
Full prof.	1808	0.282	386	0.163	1422	0.351

$$P(C,i,t) = f(X_{i,t}`C_j) \tag{4.2}$$

where the two states C are being a candidate for habilitation or not and, similarly as above

$$X_{i,t}`C_j = \gamma_{j,0} + \gamma_{j,1}Female + \gamma_{j,2}Age + \gamma_{j,3}Age^2 + \gamma_{j,4}\#Publications + \gamma_{j,5}\#Citations$$

Lastly, the probability of success in the habilitation exam, conditional on being a candidate is given by

$$P(A,i,t|C=1) = f(X_{i,t}`A_j) \tag{4.3}$$

where the two states A are succeeding of failing in being habilitated, C is a dummy which takes value one if the academic applied for habilitation and

$$X_{i,t}`A_j = \alpha_{j,0} + \alpha_{j,1}Female + \alpha_{j,2}Age + \alpha_{j,3}Age^2 + \alpha_{j,4}\#Publications + \alpha_{j,5}\#Citations$$

I can then estimate these two models to see to what extent the difference in the probability of being habilitated is due to women being less likely to enter the concourse or to them having a lower success rate in the concourse than men. That is, I am interested in whether $\alpha_{j,1}$ and $\gamma_{j,1}$ are significantly different from zero.

4.5.3 Descriptive statistics

Table 4.11 gives the decomposition in terms of gender and rank for the most recent year in the data, 2015. There were 28,749 academics in the considered scientific

TABLE 4.12: Descriptives of potential candidates

	Males				Females			
	Min	Max	Mean	SD	Min	Max	Mean	SD
				Architecture&Engineering				
Prob. Candidate	0	1	0.506	0.500	0	1	0.566	0.496
Prob. Promotion	0	1	0.337	0.473	0	1	0.278	0.448
Age2015	36	58	45	5.412	36	58	46	5.400
Publisher	0	1	0.797	0.402	0	1	0.703	0.457
#Publications	0	119	10.816	12.714	0	103	7.507	10.154
#Citations	0	4508	236.552	391.398	0	2833	165.859	283.754
				Science				
Prob. Candidate	0	1	0.648	0.478	0	1	0.536	0.499
prob. Promotion	0	1	0.305	0.460	0	1	0.213	0.410
Age2015	36	58	46	5.615	36	58	46	5.645
Publisher	0	1	0.830	0.375	0	1	0.876	0.330
#Publications	0	227	24.381	23.212	0	149	24.944	20.905
#Citations	0	7764	578.219	795.117	0	6508	587.887	700.572
				Humanities&Law				
Prob. Candidate	0	1	0.687	0.464	0	1	0.618	0.486
prob. Promotion	0	1	0.299	0.458	0	1	0.229	0.420
Age2015	36	58	45	5.470	36	58	44	5.145
éublisher	0	1	0.634	0.482	0	1	0.635	0.482
#Publications	0	51	2.370	4.312689	0	67	2.000	4.380387
#Citations	0	2244	30.500	107.3666	0	2466	31.098	128.2516

fields in Italy that year. Women account for 34% of observations, and they are over-represented amongst university professors in Science (40%) and under-represented amongst academics in Architecture&Engineering, where they account for only 22% of the population.

Slightly under one fourth of the population hold a full professor position, with the fraction being lower for academics in the Science fields and higher for those in Humanities&Law. Note that the data do not seem to indicate that women choose a career path that offers higher average promotion rates, which would counterbalance negative discrimination. Feminization, in fact, is higher for Science fields, which have a lower promotion to full professorship rate, indicating that there is no selection of this type taking place. The gender promotion gap is large, 17.2 percentage points on average, and is larger for Humanities&Law (18.8%) than for Architecture&Engineering and Science (16.2% and 16.7% respectively).

Table 4.12 reports some descriptive statistics for the sample of potential candidates for promotion. I hence consider all researchers in the scientific fields of Architecture&Engineering, Science and Humanities&Law at the assistant professor level in 2012 aged between 33 and 55, for a total of 11,035 academics.

The probability of being a candidate for the habilitation concourse is higher for men in Science and Humanities&Law fields but lower for this gender group in Architecture&Engineering. The unconditional probability of being promoted is low and is on average 31.7% for men and 22.9% for women, depending on the scientific field. On average, women in the fields of Science and Humanities&Law are slightly more likely to publish in international journals, while the opposite is true for Architecture&Engineering. Using the number of citations as a proxy for the quality of publications, I observe that the latter is higher for men in Architecture&Engineering but lower in the other two scientific areas, Science and Humanities&Law.

4.5.4 Results

Gender difference in success in habilitation concourse

The limitation of the analysis conducted so far, shared by part of the literature which examines gender differences in promotion in academia, is that it uses only outcomes and hence it is not possible to disentangle whether lower observed promotions are the result of a lower likelihood to apply for promotion or lower success in obtaining the promotions. For the Gelmini's reform, data allow to examine the two separate steps. I start by considering what determines success in the habilitation contest conditional on being a candidate, and then move to the determinants of the decision to enter the contest. The determinants of success in the habilitation contests are assumed to be gender and the variables measuring research productivity. I also add to equation 4.3 the dummy *Academic* which captures whether the individual was, at the time of application, holding a position in Italian academia (i.e. was not an external candidate, e.g. coming from others research entities, from abroad, etc.). In Columns 5 and 6 I consider respectively only the subsamples of researchers who do not hold a position in Italian academia in 2012 and those who do. For the latter group I have more detailed information and thus can add variables about age and research experience.

The results reported in Table 4.13 indicate that research output is a key determinant of the probability of success in the habilitation concourse since it is positive and significant through all the specifications and subsamples. Both quantity and

TABLE 4.13: Likelihood to be habilitated conditional on applying

	All (1)	All (2)	All (3)	All (4)	Others (5)	Academics (6)
Female	-0.0253**	-0.0176**	-0.0197**	-0.0119*	-0.0233**	-0.00349
	(0.00586)	(0.00578)	(0.00560)	(0.00550)	(0.00743)	(0.00819)
#Pulications		0.00133**	0.00120**		0.000811**	0.00201**
		(0.000133)	(0.000138)		(0.000155)	(0.000248)
#Citations		0.000595**	0.000638**		0.000564**	0.000598**
		(6.23e-05)	(5.97e-05)		(6.55e-05)	(0.000113)
Academic			0.254**	0.253**		
			(0.00459)	(0.00454)		
H-index				0.0179**		
				(0.000437)		
ArcEng			-0.0283**	-0.0583**	-0.117**	0.0576**
			(0.00792)	(0.00768)	(0.0119)	(0.0111)
Sci			0.0206**	-0.0679**	0.0526**	-0.00471
			(0.00761)	(0.00773)	(0.0101)	(0.0116)
Age						-0.00742
						(0.00620)
Age^2						-7.58e-05
						(6.59e-05)
Experience						0.00156+
						(0.000931)
Observations	31,759	31,759	31,759	31,759	16,369	15,390
Log Likelihood	-21801	-21329	-20141	-19665	-9695	-9985
Chi-square	18.70	962.1	3339	4291	1137	1021
Pseudo R2	0.000429	0.0221	0.0765	0.0984	0.0554	0.0486

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

quality (measured by number of citations) and the H-index (introduced in specification 4) have a positive and significant coefficient. The coefficients for productivity indicators are higher, even if only slightly in the case of number of citations, for the subsample of academics (see specifications 5 and 6). This is consistent with the fact that there is a stronger selection for those who are not following the standard university track. In line with this argument, holding a position in Italian academia (measured by the dummy *Academic*) at the time of the concourse has in fact a strong positive impact on the probability of passing the habilitation.

Turning to the coefficient of interest, the impact of being a female on the probability of habilitation conditional of being a candidate, there is a negative and significant difference across genders, with the exception of the academic subsample (see Column 7), for which the coefficient is not significant. To further explore this result, in Table 4.14 I split the academic subsample in two subgroups according to the rank of the habilitation for which the candidates were applying (associate professor or full professor). It is interesting to notice that the female dummy is negative and significant just at the associate professor level, while for the full professor habilitation exams, the most important variables in determine the probability of being habilitated are age and research experience of the candidates. The variables related to the research output, both quantity and quality and the H-index of the candidates, are not significant for the associate professor level habilitation exams. I can then interpret the insignificant coefficient of *Female* as a lack of evidence of ex post discrimination against women in the habilitation process for full professorship only.

Likelihood to enter the habilitation concourse

I turn now to the determinants of the likelihood to enter the concourse. I estimate equation 4.2 considering all researchers holding a position as assistant or associate professor, active in Italian academia in 2012 and aged between 33 and 55, adding to equation 4.2 the dummy variable *TopUni*, which takes value 1 if the individual is in a top university, according with the Academic Ranking of World Universities (ARWU). The results are reported in Table 4.15.

The unconditional probability of applying is lower for women than for men (Column 1) and the negative impact of being female persists once I include individual

TABLE 4.14: Likelihood to be habilitated conditional on applying (by academic rank)

	Full Prof.				Associate Prof.					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	-0.0185	-0.0102	-0.00450	-0.000993	0.00256	-0.0468**	-0.0335**	-0.0300**	-0.0176+	-0.0181+
	(0.0140)	(0.0139)	(0.0138)	(0.0139)	(0.0137)	(0.0104)	(0.0102)	(0.0101)	(0.0101)	(0.00997)
Age		0.0442**	0.0389**	0.0371**	0.0337**		0.00761	-0.00278	-0.00157	0.00213
		(0.0115)	(0.0114)	(0.0114)	(0.0113)		(0.00936)	(0.00915)	(0.00911)	(0.00903)
Age^2		-0.000587**	-0.000527**	-0.000508**	-0.000456**		-0.000259*	-0.000131	-0.000140	-0.000172+
		(0.000115)	(0.000114)	(0.000114)	(0.000113)		(0.000105)	(0.000103)	(0.000102)	(0.000101)
Experience		0.00601**	0.00552**	0.00554**	0.00598**		0.00217	0.000546	0.00159	0.00130
		(0.00149)	(0.00149)	(0.00149)	(0.00146)		(0.00133)	(0.00132)	(0.00132)	(0.00130)
#Publications			0.00118**	0.00113**				0.00298**	0.00428**	
			(0.000271)	(0.000305)				(0.000329)	(0.000417)	
#Citations			0.000343*	0.000388**				0.00129**	0.00133**	
			(0.000138)	(0.000140)				(0.000194)	(0.000196)	
H-Index					0.0167**					0.0268**
					(0.00108)					(0.00118)
ArcEng				0.0373*	0.00776				0.0482**	0.0283*
				(0.0175)	(0.0170)				(0.0142)	(0.0139)
Sci				0.00669	-0.0870**				-0.0733**	-0.131**
				(0.0182)	(0.0185)				(0.0156)	(0.0152)
Observation	6,327	6,327	6,327	6,327	6,327	9,063	9,063	9,063	9,063	9,063
LLikelihood	-4382	-4303	-4248	-4245	-4159	-6022	-5867	-5658	-5617	-5541
Chi-square	1.741	160.1	269.7	275.5	447.9	20.15	329.6	748.4	829.2	982.2
Pseudo R2	0.000199	0.0183	0.0308	0.0314	0.0511	0.00167	0.0273	0.0620	0.0687	0.0814

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

TABLE 4.15: Likelihood to apply for habilitation (Marginal effects)

	All (1)	All (2)	All (3)	All (4)	Assistant prof. (5)	Associate prof. (5)
Female	-0.0611**	-0.0638**	-0.0519**	-0.0534**	-0.0458**	-0.0660**
	(0.00876)	(0.00871)	(0.00853)	(0.00855)	(0.0129)	(0.0191)
Age		0.123**	0.112**	0.108**	0.118**	-0.0110
		(0.0106)	(0.0103)	(0.0103)	(0.0130)	(0.0276)
Age^2		-0.00140**	-0.00123**	-0.00118**	-0.00129**	9.57e-05
		(0.000120)	(0.000118)	(0.000117)	(0.000151)	(0.000294)
Experience		-0.00323**	-0.00717**	-0.00654**	-0.000635	-0.00928**
		(0.00101)	(0.00100)	(0.000997)	(0.00135)	(0.00179)
#Publications			0.000704	0.00304**	0.00353**	0.00317**
			(0.000591)	(0.000623)	(0.000862)	(0.000895)
#Citations			0.000240**	0.000233**	0.000272**	0.000186**
			(2.29e-05)	(2.30e-05)	(3.22e-05)	(3.19e-05)
TopUni				0.0514**	0.0642**	0.0410*
				(0.00853)	(0.0136)	(0.0173)
Female*TopUni					-0.0190	-0.00516
					(0.0215)	(0.0306)
ArchEng				-0.108**	-0.124**	-0.0864**
				(0.0106)	(0.0130)	(0.0181)
Sci				-0.169**	-0.193**	-0.162**
				(0.0117)	(0.0146)	(0.0203)
Observations	12,773	12,773	12,773	12,773	8,323	4,449
Log Likelihood	-8392	-8316	-7981	-7864	-5077	-2737
Chi-square	47.87	199.9	871.1	1104	744.5	453.3
Pseudo R2	0.00284	0.0119	0.0518	0.0656	0.0683	0.0765

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

characteristics, the marginal effect being -0.053. Research output has a significant but not strong effect on the likelihood of being a candidate in the concourse, while being affiliated to a top university significantly tends to increase the probability of being a candidate. This may be due to the fact that being in a top university has a positive effect of the probability of success; potential candidates may anticipate this and hence be more likely to apply if they are in those universities. An additional effect may come from peer pressure to pass the concourse. Removing this variable from the analysis has no impact on the other estimated coefficients.

In order to try to understand what lies behind the gender gap in seeking promotion, I examine whether the effect of gender differs between the two types of academic positions: assistant and associate professor. Furthermore, if women have a stronger relative preference for university prestige, they may choose not to apply for habilitation in order to stay in a top university. Since promotion often implies mobility for university professors, if females cared more about university quality they would be less willing to move, and hence to candidate for habilitation, when they hold an associate professor position.

I hence run again the regressions for the likelihood to apply for habilitation and include an interaction between being a woman and being in a top university. If the trade-off between rank and university quality differs across genders, this would be captured by a negative coefficient on being in a top university for female associate professors but not for female assistant professors.

The marginal effects obtained by these regressions are reported in Column 5 and 6. Being in a top university does not affect the likelihood of female academics to apply for habilitation. For male university professors, the university of affiliation seem to matter a great deal. This indicates that the effect of being in a top university obtained earlier is driven by the effect of men in top universities.

Networks effects

Another possible reason for differences in the likelihood of seeking a promotion is that there is some individual variable that I have so far ignored and which has an impact on the actual (or perceived) probability of success. Possible explanations are research networks or groups of coauthors. The idea that networks are important in obtaining jobs and achieving promotions is widespread in the literature, and the issue has been studied in the case of promotions in academia (Combes et al., 2008; Zinovyeva and Bagues, 2015; McDowell et al., 2001).

Coauthors networks have been shown to differ across genders, with females having fewer coauthors and a lower fraction of male coauthors (see Boschini and Sjögren, 2007). If women have smaller or less efficient networks, then this may affect the expected outcome and hence the payoff from entering the competition for promotion. Potential reasons for this effect are that a candidate who has a member of her network in a promotion committee has a higher likelihood of success, but also that with a larger network the candidate's work may be better known and cited, or that this could provide extra information about how to best prepare for a concourse.

As a result, the gap that I find between men and women could be due to differences in coauthors and networks, and in this case including these variables would have an effect on the coefficient on gender.

In order to test this hypothesis, I estimate 4.2 substituting to bibliometric indicators the dummy *Publisher*, which takes value 1 if the individual has published in Scopus-classed journals. I then construct for each individual two measures of networks, based on co-authorship. The first measure is the size of an individual's network (*NetwSize*), defined as the total number of different coauthors in Italian academia the researcher has had over her publishing lifetime; the second is the fraction of network members that are men active in Italian academia (*ShareNetwMale*).

Table 4.16 reports the regressions for the likelihood of entering the habilitation contest to which I have added network variables. Both variables have positive coefficients although the proportion of men is significant at the 1% level only considering the associate professors in the sample. The coefficients on gender barely changes as compared to earlier specifications, that is, gender differences in networks do not seem to explain the lower propensity of women to seek promotion. The reason is probably that differences in network size across genders are minor and not statistically significant.

TABLE 4.16: Network effects: likelihood to apply for habilitation (Marginal effects)

	Whole sample	Assistant prof.	Associate prof.
	(1)	(2)	(3)
Female	-0.0625**	-0.0596**	-0.0750**
	(0.00880)	(0.0107)	(0.0155)
Age	0.117**	0.128**	-0.0160
	(0.0105)	(0.0133)	(0.0285)
Age^2	-0.00130**	-0.00143**	0.000125
	(0.000119)	(0.000154)	(0.000303)
Experience	-0.00549**	-0.000560	-0.00966**
	(0.00104)	(0.00140)	(0.00187)
Publisher	0.0360*	0.0572**	-0.000495
	(0.0153)	(0.0188)	(0.0264)
TopUni	0.0570**	0.0619**	0.0459**
	(0.00872)	(0.0108)	(0.0147)
NetwSize	0.00777**	0.00877**	0.00694**
	(0.000690)	(0.000931)	(0.00103)
ShareNetwMale	0.0126	-0.0216	0.0770**
	(0.0161)	(0.0198)	(0.0274)
ArcEng	-0.0723**	-0.0833**	-0.0522**
	(0.0114)	(0.0141)	(0.0197)
Sci	-0.113**	-0.128**	-0.0977**
	(0.0127)	(0.0159)	(0.0219)
Observations	12,773	8,323	4,449
Log Likelihood	-7851	-5050	-2738
Chi-square	1130	799.1	450.8
Pseudo R2	0.0672	0.0733	0.0760

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

TABLE 4.17: The competition environment: likelihood to apply for habilitation (Marginal effects)

	Whole sample (1)	Assistant prof. (2)	Associate prof. (3)
Female	0.0142	0.0371	-0.0352
	(0.0247)	(0.0302)	(0.0426)
Age	0.108**	0.117**	-0.00653
	(0.0103)	(0.0130)	(0.0277)
Age^2	-0.00118**	-0.00129**	0.00005
	(0.000117)	(0.000151)	(0.000294)
Experience	-0.00628**	-0.000231	-0.00929**
	(0.00101)	(0.00136)	(0.00180)
#Publications	0.00380**	0.00484**	0.00346**
	(0.000679)	(0.000939)	(0.000977)
#Citations	0.000233**	0.000261**	0.000196**
	(0.00002)	(0.00003)	(0.00004)
TopUni	0.0512**	0.0566**	0.0402**
	(0.00853)	(0.0106)	(0.0143)
shareWomenFP	-0.000965	0.158	-0.263
	(0.0989)	(0.124)	(0.163)
Female*ShareWomenFP	-0.393**	-0.536**	-0.168
	(0.137)	(0.168)	(0.234)
ArcEng	-0.110**	-0.116**	-0.105**
	(0.0139)	(0.0171)	(0.0240)
Sci	-0.189**	-0.211**	-0.189**
	(0.0126)	(0.0157)	(0.0218)
Observations	12,773	8,323	4,449
Log Likelihood	-7865	-5073	-2739
Chi-square	1104	752.8	448.8
Pseudo R2	0.0656	0.0691	0.0757

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

The competition environment

The final test consists of examining whether the environment under which the competition takes place affects men and women differently. I hence consider, for each scientific sector, the proportion of women amongst full professors in the year of the competition, which can act as a proxy for the "perceived discrimination", and the interaction with the female dummy.

Table 4.17 presents the estimated regressions. Interestingly, the fraction of women amongst full professors has a negative and significant effect on the propensity of women to apply for habilitation. The effect is slightly larger and more precisely estimated considering the assistant professor sample rather that associate professor one (Column 2 and 3).

4.6 Conclusions

This work has used data on promotions among academics in Italy to look at the attitude and outcome of women in the context of two different academic recruitment reforms. The first one (Berlinguer's reform) in 1998 moved the recruiting system towards a decentralized mechanism, while the latter (Gelmini's reform) in 2010 introduced a two-step procedure with a national habilitation and local concourses to recruit professors.

In the first case, I just observe the outcome of promotion, meaning that I do not have any information on who was actually a candidate. Thus, in a quasi-experimental research framework, I apply propensity score and CEM techniques to select two balanced groups of academics: one including researchers who faced the pre-reform institutional setting and the second with researchers who faced the decentralized mechanism. Then, within a DD framework, I measure the average impact of the policy change on women promotion chances.

Results of the DD estimator show that the probability of women being promoted after the reform took place is higher than in the previous recruitment mechanism, with the reform thus having a positive impact.

In the case of the most recent academic reform, on the other hand, I exploit information on the first round of the Italian national scientific habilitation which took place in 2012. Thus I have information both on who was a candidate for habilitation (with, of course, the outcome of the evaluation procedure) and who could have been a candidate. This allows to test different hypothesis about the causes of women's lower promotion rates.

First, I consider what determines success in the habilitation concourse conditional on being a candidate. I find evidence that there is a stronger selection for those who are not following the conventional university track (i.e. apply to habilitation holding a position in academia). Furthermore, for this subgroup, there is a negative and significant difference across genders in the probability of habilitation. Splitting the academic subsample in two subgroups according to the rank of the habilitation exam (habilitation for associate or full professorship) I find a lack of evidence of discrimination against women in the habilitation process for full professorship only.

Examining the determinants of the likelihood of entering the habilitation concourse, I also find that females are less likely than men to enter the contest. This could be due to the fact that women believe that they will be discriminated against and hence decide not to enter the competition. Data, however, point against this hypothesis, with women being less likely to apply the more females are amongst full professor academics within each scientific field. Another possible explanation is that women are less willing than men to enter the contest, especially if they already work in a prestigious university. Results show that being in a top university does not affect the likelihood of female academics to apply, while it seems to matter a great deal for male university professors. Finally, I also test for network differences among men and women being the cause for the lower propensity of the latter to enter the concourse, but, although statistically significant, these differences in size and genders are probably minor and do not have an impact on the magnitude of the other gender coefficients.

Gender segregation is expected to play a more relevant role in evaluation decisions which are held at a more aggregate level (Bagues et al., 2017). However, as shown in Chapter 3, an important factor of bias is the existence of strong connections between candidates and committees which are in charge of evaluation or hiring decisions. These connections are likely to be more relevant at the university level or, for instance, in laboratories, which often are organized as family businesses by their directors (Freeman et al., 2001). In these contexts, PhDs play a central role and so does their relationship with the advisor, relationship which could be influenced by the gender pairing between advisor and advisee (Pezzoni et al., 2016). More empirical evidence is needed to understand the impact of gender quotas in these contexts.

Concluding Remarks

In the past 20 years, the number of empirical contributions to the economics of science has grown considerably (Stephan, 2012). Very few studies, however, have examined explicitly the issue of academic careers, the main exceptions being limited to the US case (see for example Ehrenberg and Zhang, 2005 or contributions from economists who are interested in the analysis of recruitment examinations in their own discipline, such as Ginther and Kahn, 2004). This thesis contributes to the existing literature on career progression of academics focusing on the Italian labor market. Exploiting a novel source of data, I addressed several issues in academic careers starting from the very early stage, i.e. the granting of the doctoral degree. Furthermore, it was possible to explore the role of social capital in researchers' success and to deepen our understanding about the persistent gender gap in academic promotion. Hereafter I summarize contents and results of each chapter.

In Chapter 1 I presented a novel dataset which covers the academic career of Italian doctorate holders from their PhD for a period of two decades, tracing their publication performance. This database exploits an unused source of information: the repository of doctoral dissertations at the national library of Florence. This is a valuable source which provides the identity and affiliation of dissertations' authors, advisors and in some cases committee members of the doctoral thesis discussed in Italian universities from the first cycle (in 1986) to 2006. This information was matched through a record linkage algorithm to ministerial data on yearly position held in Italian academia by university professors in the period 1990-2015. The database created allows to shed light on several issues related to the Italian academic labor market, such as gender, inbreeding, mobility, hiring and promotion patterns. Furthermore, thanks to the detailed information at the doctoral career stage, the database allows to explore the role of social capital in career progression, in particular with respect to the mentor–mentee relationship, as I showed in particular in Chapter 2 and 3.

Chapter 2, in fact, has explored the process of transition to tenure in academia to address the issues related both to time to first appointment and tenure at Italian universities. The most relevant variables related to time to first appointment and time to tenure refer to the early stage of the academic career. Results in terms of the role of past academic performance provide support for the claims of cumulative effect on careers, considering publication during the doctorate and the age at PhD. Of particular interest are the results unveiling the role of the PhD supervisor on mentees' career. As far as co-authorship is considered, in the first appointment stage it seems detrimental, since the more prestigious author might take greater credit for the performed work, while a longer mentor–mentee scientific collaboration accelerates the transition to permanent positions. Furthermore, the advisor's influence seems to matter also in more indirect ways. At first appointment stage, researchers who "inherited" the affiliation from the advisor's network have a higher hazard ratio of getting an academic position. At the tenure stage, on the other hand, advisor's prestige plays a positive role both in terms of scientific network size and his/her centrality within said network. The results associated with mobility are mixed. On the one hand, international mobility is negative related to a fast career in Italian academia, but this negative effect is declining with time. In terms of hazard rate of getting tenure, after 8 years of duration researchers may eventually benefit from international mobility experience. On the other hand, having worked in several Italian universities is a valuable characteristic. It seems that career advancement in the Italian academia, while grounded in a merit-based system, partly suffers from the absence of an open academic job market, and from the existence of mechanisms for accessing the profession that could be shaped by particularistic dynamics. From a policy point of view, it is difficult to reconcile the emphasis placed on the desirability of international mobility as a way of assuring knowledge circulation, with the evidence that it is partly negatively associated to the duration of transition to an academic position.

Chapter 3 aimed at exploring the role of social capital in faculty recruitment, with particular emphasis on the role played by ties established during candidates' PhD education. The effect of social capital is often overestimated because it can be a consequence of success rather than its cause. The Chapter exploits data from a randomized natural experiment which allows to assess the real causal effect of social capital: the first round of Italian scientific habilitation concourse in 2012. It exploits the fact that the commissions which produce the final evaluation of applicants are composed of academics drawn at random. This Chapter furthers the debate on the efficiency of social capital and contacts in getting a job, as well as the role played by ties of different strengths. The analysis shows that social capital clearly does matter and also that strong ties matter more. Main results of this Chapter are that when a strong acquaintance, such as the PhD advisor or a coauthor, is randomly drawn the chances of receive a positive evaluation by the committee increase by 17 percentage points (10 p.p. in the case of colleague evaluators). There could be several mechanisms through which these applicants' ties have a purely causal effect on the evaluation committees' decision. First, advisor–advisee and coauthors share similar scientific preferences, which are at the basis of their collaboration, either for a PhD or a scientific article. So it could be the case that that evaluators' support is not toward a person but rather toward a scientific approach which they share. A second mechanism in term of information and evaluation costs could lead to similar results: strong ties are aware of what it is interesting about their acquainted candidates' work, so that it is less costly to find and promotes the key points of these applicants, and they have a better understanding of their intrinsic value and unobserved ability with respect to other evaluators. Results, in fact, show that candidates selected by acquainted evaluators do not receive a preferential treatment, on the contrary they turn out to be more productive and successful with respect to other positive evaluated applicants. Academic systems (should) share the ideal of meritocracy and indifference to personal characteristics. Bias in recruitment are prejudicial both to the quality of the academic system and to its equity. On the other hand, academia consists of communities where people know one another well. Excluding all persons with a potential bias toward an applicant could lead to a decrease in the overall efficiency of the system itself.

Finally, Chapter 4 explores another source of bias in the academic labor market: gender discrimination. It analyses gender differences in promotion in Italian universities in the context of the two most recent reforms of the academic labour market, in order to see whether they helped or harmed the path toward gender equality. The first reform under analysis took place in 1998 and had the effect of decentralizing the recruitment system at the university level. Through propensity score matching and a non-parametric matching algorithm (CEM), I identified two groups of comparable scientists who were selected one before and the other after the introduction of the reform. In a differences-in-differences framework, results show that show that the probability of women being promoted after the reform took place is higher, with the reform thus having a positive impact on gender neutrality. In this case, however just the outcome of promotion is observed, meaning that no information was available on who actually was a candidate in the concourses. For the second reform, on the other hand, I exploited information both on candidates and those eligible to be candidates in order to disentangle different possible explanations for women's lower promotion rate in academia. The reform took place in 2010 and introduced a two-step procedure, in line with other European academic systems, with a national scientific habilitation at which the researchers have to qualify before taking part to local concourses. Results show that gender has a significant negative effect on candidates' habilitation rate only in the case of associate professor positions and considering those researchers who do not follow the conventional academic track (i.e. apply for habilitation from outside Italian academia). I also find that women are less likely to seek promotion than men. After eliminating a number of possible explanations for the gap in seeking promotion, the possible interpretation is that men are more willing that women to participate in contests, in line with experimental evidence that highlights male overconfidence in certain situations. These results provide empirical evidence to the question of what type of policy intervention can help in increasing female promotion rates.

Appendix A

Appendix to Chapter 1

Matching Validation

Comparing the matching result of the record linkage operation between the two databases is not an easy task. There are no official statistics or data about the percentage of Italian doctorates who hold a permanent position in Italian academia, thus we have to rely on partial studies carried on by some universities or research institutions on the employment outcome of (some cohorts of) the doctorate holders. The interest here is to have some information about the Italian doctorates who work within the Italian academic system. Unfortunately, often these studies just investigate whether the doctorates of a given cycle are employed or not, without distinguishing between different activities, and even if they do, they usually do not disaggregate per country of residence. The following Table A.1 summarizes in our knowledge all the most recent studies conducted on this topic. As is possible to see, our findings are in line with these studies on the employment outcome of Italian PhD holders: roughly one third of the doctorates from Italian universities holds a position in Italian academia.

In what follows we describe how this procedure applied to academics in physics and chemistry from university and polytechnic of Turin, which can be considered an exemplary case.

Italian academics in chemistry and physics from the two Turin universities who are active in 2015 in the dataset are 231; among these, the researchers who were eligible for obtaining the PhD degree in the period 1986-2006 are 173.

The narrow matching exercise resulted in 127 matches and the broad matching

XML structure	Spreadsheet structure
<rec>	
<cf t="001">BNI0000005</cf>	001: Record id
<df t="200" i1="1" i2=" ">	
<sf c="a">^La %malattia di Gaucher</sf>	200-a: Title
<sf c="e">efficacia della terapia sostitutiva enzimatica in due pazienti affetti dalla forma di tipo 1.</sf> <sf c="e">tesi di dottorato di ricerca: scienze pediatriche</sf>	200-e: Subtitle(s)
<sf c="f">Francesco Nigro</sf>	
<sf c="g">tutore: A. Fiumara</sf> <sf c="g">coordinatore: Giustiniano Reitano</sf> <sf c="g">Università degli studi di Catania</sf>	200-g: Responsibility
</df>	
<df t="210" i1=" " i2=" "> <sf c="d">2003.</sf> </df>	210: Publication year
<df t="300" i1=" " i2=" "> <sf c="a">8. ciclo</sf> </df> <df t="300" i1=" " i2=" "> <sf c="a">A. a. 1995-1996</sf> </df>	300: Notes
<df t="676" i1=" " i2=" "> <sf c="a">618.92</sf> <sf c="v">21</sf> <sf c="9">PEDIATRIA</sf> </df>	676: Dewey classification
<df t="686" i1=" " i2=" "> <sf c="a">MED/38</sf> <sf c="9">Pediatria generale e specialistica</sf> </df>	686: Miur classification
<df t="700" i1=" " i2="1"> <sf c="a">Nigro</sf> <sf c="b">, Francesco</sf> <sf c="3">BNIV000008</sf> </df>	700: Author name
<df t="702" i1=" " i2="1"> <sf c="a">Fiumara</sf> <sf c="b">, Agata</sf> <sf c="3">BNIV000007</sf> </df> <df t="702" i1=" " i2="1"> <sf c="a">Reitano</sf> <sf c="b">, Giustiniano</sf> <sf c="f"><1931- ></sf> <sf c="3">SBLV148695</sf> </df>	702: Other author(s)
<df t="712" i1="0" i2="2"> <sf c="a">Università degli studi</sf> <sf c="c"><Catania></sf> <sf c="3">CFIV008644</sf> </df>	712: University affiliation
</rec>	

FIGURE A.1: XML tags of a thesis stored in the BNCF OPAC

TABLE A.1: Summay of the studies

STELLA	Year: 2009 University(/ies): Bergamo, Brescia, Milano, Milano-Bicocca, Palermo, Pisa, Pisa-Sant'Anna. Cohort(s): 2005, 2006, 2007. Methodology: census, CAWI, 1,758 respondents (49.5%).	85.5% is employed, among them the 81.6% works in Italy. 40% is employed within a public or private university, but country and type of contract are not specified.
STELLA	Year: 2010 University(/ies): Bergamo, Brescia, Milano, Milano-Bicocca, Palermo, Pisa, Pisa-Sant'Anna. Cohort(s): 2007, 2008. Methodology: census, CAWI, 1,579 respondents (53.4%).	75.1% is employed, among them the 82.4% works in Italy. 33% is employed within a public or private university, but country and type of contract are not specified.
STELLA	Year: 2011 University(/ies): Bergamo, Brescia, Milano, Milano-Bicocca, Palermo, Pavia, Pisa, Pisa-Sant'Anna. Cohort(s): 2008, 2009. Methodology: census, CAWI, 1,637 respondents (50.7%).	77.8% is employed, among them the 86.8% works in Italy. 34.6% is employed within a public or private university, but country and type of contract are not specified.
STELLA	Year: 2012 University(/ies): Bergamo, Brescia, Milano, Milano-Bicocca, Palermo, Pavia, Pisa, Pisa-Sant'Anna. Cohort(s): 2009, 2010. Methodology: census, CAWI, 1,769 respondents (51.1%).	81.9% is employed, among them the 90.3% works in Italy. 32.4% is employed within a public or private university, but country and type of contract are not specified.
STELLA	Year: 2013 University(/ies): Bergamo, Brescia, Milano, Milano-Bicocca, Palermo, Pavia, Pisa. Cohort(s): 2010, 2011. Methodology: census, CAWI+CATI, 1,938 respondents (60.1%).	87.5% is employed, among them the 91.8% works in Italy. 29.3% is employed within a public or private university, but country and type of contract are not specified.
STELLA	Year: 2014 University(/ies): Bergamo, Brescia, Milano, Palermo, Pisa, Pavia. Cohort(s): 2011, 2012. Methodology: census, CAWI+CATI, 1,537 respondents (55%).	80.3% is employed, among them the 92.8% works in Italy. 24.8% is employed within a public or private university, but country and type of contract are not specified.
UniPR	Year: 2012 University(/ies): Parma. Cohort(s): 2006, 2007, 2008, 2009. Methodology: census, CAWI, 378 respondents (48.2%).	97.3% is employed. 45.5% is employed within a public or private university. Among them, the 19.2% is either RU, AP or FP and the 7.6% is employed in a foreign university.
UniTN	Year: 2006 University(/ies): Milano, Milano-Bicocca, Trento. Cohort(s): 1998-2005. Methodology: census, CATI, 1,179 respondents (66%).	97.1% is employed. 58.5% is employed within a public or private university, where the 21.1% is either RU, AP or FP while the 37.4% holds a non-permanent position in academia.
LUISS	Year: 2013 University(/ies): LUISS. Cohort(s): 2007, 2009. Methodology: census, CATI, 58 respondents (83%).	100% is employed. 31% is employed within a public or private Italian university, where the 20.7% holds a non-permanent position in academia.
ISTAT	Year: 2009 University(/ies): all. Cohort(s): 2004, 2006. Methodology: census, CAWI+CATI.	93.5% is employed. Among these, the 17.3% has either a post-doc scholarship or a temporary position in academia. 6.9% is resident in a foreign country.
UniTN	Year: 2014 University(/ies): all. Cohort(s): 2008, 2010. Methodology: census, CAWI+CATI.	92.4% is employed. Among these, the 36.5% has either a post-doc scholarship or a temporary position in academia. 12.9% is resident in a foreign country.
ISFOL	Year: 2012 University(/ies): all. Cohort(s): 2006. Methodology: survey, CAWI+CATI, 4,879 respondents (48.7%).	92.5% is employed. Among these, the 18% has either a post-doc scholarship or a temporary position in academia. 7.9% is resident in a foreign country.

resulted in 12 additional academic-doctorate pairs. The filtering out exercise allowed to have 139 unique (i.e. one-academic-to-one-doctorate) matches (see Figure A.2 for an example).

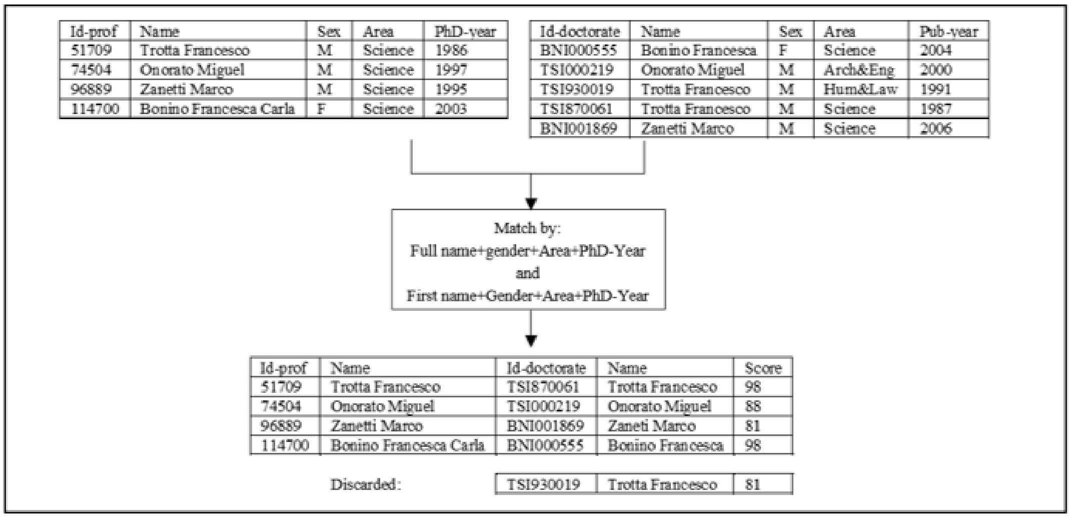

FIGURE A.2: Example of the record linkage

we manually checked the CVs of the remaining academics:

- 17 do not have a PhD;
- 7 hold a PhD from a foreign university;
- 4 hold a PhD from Italian institutions that are not legally obliged to deposit the doctoral dissertations in the BNCF repository;
- 4 hold a doctorate from an Italian university but their thesis does not appear in the BNCF repository[1];
- 2 of them were not matched by the algorithm.

Thus the algorithm allows us to correctly identify 95% of the academics who earned a doctorate from an Italian university. Figure A.3 shows graphically the matching results obtained.

[1] This is due to the small unbalance between the thesis discussed and those deposited in the BNCF, which has been explained in the previous section.

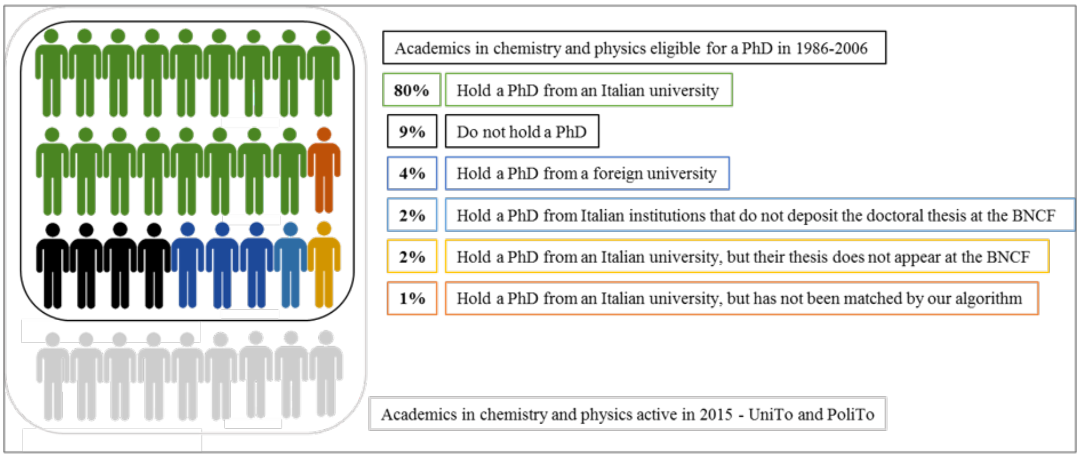

FIGURE A.3: Graphical representation of the record linkage

Summary Statistics

Hereafter we provide summary statistics of the resulting database of doctorates from Italian universities who pursued an academic career in Italian academia. First, in Table A.2 and Table A.3 we provide the detail by year and discipline of the number of, respectively, thesis linked and not linked to an Italian academic by the algorithm. Table A.4 and Figure A.4 provide instead the share of the thesis-academic links found over the total number of thesis.

We observe that the share of doctorates who held a position in Italian universities decreases over time because, intuitively, these researchers had less time to try enter in the academia and recent reforms of the higher education sector introduced new non-permanent positions which expanded the time window of the post-doc period. Furthermore, as expected, doctorates in Medicine&Veterinary field are those less interested in an academic jobs[2], while Economics&Statistics and Social Sciences are the fields where doctorates are most keen to pursue this type of career.

Similarly, in Table A.5 and Table A.6 we give, respectively, the number of academics linked to thesis by disciplinary area and expected year of doctorate, and those who were not linked by the algorithm. In Table A.7, Figure A.5 and Figure A.6 we provide the share of academic-thesis links found over the total number of academics.

[2]A common practice in this field, in fact, is to give PhD scholarships to young physicians in order to retain them at the hospital while they wait to be permanently hired. In this period of time they do not involve in any scientific research, but rather they perform their everyday tasks as doctors.

From this point of view, it is interesting to notice that, with the exclusion of Medicine&Veterinary for the reasons already explained (see footnote 8), Economics&Statistics is the scientific field in which the share of professor with a PhD degree from an Italian university is lowest (57%). Disaggregating by scientific sector, one can notice that this low percentage is also shared by mathematics and letters professors (58%). On the other hand, Architecture&Engineering exhibits the highest share of professors with an Italian PhD degree (68%) while, among the scientific sectors, chemistry stands out (72%).

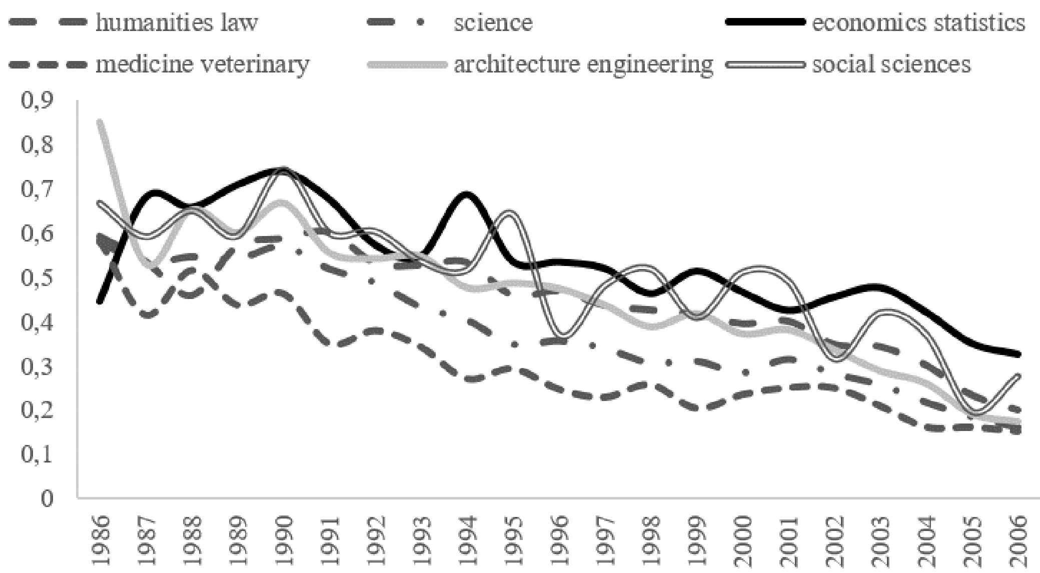

FIGURE A.4: Thesis linked to an academic (%)

TABLE A.2: Links “academic-thesis” identified per broad scientific area and year (1986-2006)

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Tot
Hum&Law	16	126	94	138	174	118	208	204	206	200	398	460	449	329	393	419	190	404	501	576	482	6,085
Science	30	144	165	184	334	136	365	407	399	322	434	518	526	368	406	476	236	468	506	593	486	7,503
Econ&Stat	4	38	23	46	45	25	71	70	94	77	87	136	140	138	163	182	85	146	190	211	215	2,186
Med&Vet	7	36	74	65	109	43	100	141	89	98	136	162	156	99	135	173	98	165	189	239	207	2,521
Arch&Eng	34	79	58	96	122	56	191	189	221	223	322	361	407	328	339	348	171	302	379	329	266	4,821
SociSci	2	26	13	16	23	9	35	30	32	43	30	50	55	36	47	56	12	73	85	65	93	831
Tot	93	449	427	545	807	387	970	1,041	1,041	963	1,407	1,687	1,733	1,298	1,483	1,654	792	1,558	1,850	2,013	1,749	23,947

TABLE A.3: Thesis not linked per broad scientific area and year (1986-2006)

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Tot
Hum&Law	11	109	111	104	122	78	183	183	180	236	451	595	606	464	604	628	357	776	1,167	1,904	1,946	10,815
Science	22	127	138	157	254	127	385	537	594	605	788	1,013	1,196	821	1,029	1,041	605	1,356	1,827	2,628	2,538	17,788
Econ&Stat	5	18	12	19	16	12	53	58	43	67	76	126	163	131	187	247	102	161	262	393	446	2,597
Med & Vet	5	51	70	84	127	80	164	271	240	236	414	545	450	385	439	517	294	621	975	1,234	1,152	8,354
Arch&Eng	6	70	31	64	61	45	162	157	244	236	358	466	645	459	575	567	339	749	1,083	1,402	1,272	8,991
SocSci	1	18	7	11	8	6	23	26	30	24	51	54	51	52	45	58	26	101	144	265	243	1,244
Tot	50	393	369	439	588	348	970	1,232	1,331	1,404	2,138	2,799	3,111	2,312	2,879	3,058	1,723	3,764	5,458	7,826	7,597	49,789

TABLE A.4: Percentage of thesis linked to an academic per broad scientific area and year (1986-2006)

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Tot
Hum&Law	0.59	0.54	0.46	0.57	0.59	0.60	0.53	0.53	0.53	0.46	0.47	0.44	0.43	0.41	0.39	0.40	0.35	0.34	0.30	0.23	0.20	0.36
Science	0.58	0.53	0.54	0.54	0.57	0.52	0.49	0.43	0.40	0.35	0.36	0.34	0.31	0.31	0.28	0.31	0.28	0.26	0.22	0.18	0.16	0.30
Econ&Stat	0.44	0.68	0.66	0.71	0.74	0.68	0.57	0.55	0.69	0.53	0.53	0.52	0.46	0.51	0.47	0.42	0.45	0.48	0.42	0.35	0.33	0.46
Med&Vet	0.58	0.41	0.51	0.44	0.46	0.35	0.38	0.34	0.27	0.29	0.25	0.23	0.26	0.20	0.24	0.25	0.25	0.21	0.16	0.16	0.15	0.23
Arch&Eng	0.85	0.53	0.65	0.60	0.67	0.55	0.54	0.55	0.48	0.49	0.47	0.44	0.39	0.42	0.37	0.38	0.34	0.29	0.26	0.19	0.17	0.35
SocSci	0.67	0.59	0.65	0.59	0.74	0.60	0.60	0.54	0.52	0.64	0.37	0.48	0.52	0.41	0.51	0.49	0.32	0.42	0.37	0.20	0.28	0.40
Tot	0.65	0.53	0.54	0.55	0.58	0.53	0.50	0.46	0.44	0.41	0.40	0.38	0.36	0.36	0.34	0.35	0.31	0.29	0.25	0.20	0.19	0.32

TABLE A.5: Academics linked to thesis per disciplinary area and expected PhD year (1986-2006)

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Tot
Hum&Law	198	202	227	229	249	334	343	343	363	344	343	337	333	308	326	309	299	220	184	104	67	5,662
Science	225	314	340	351	389	392	427	435	414	430	392	365	334	383	379	363	348	265	223	146	68	6,983
Econ&Stat	38	39	42	73	60	107	108	116	139	113	158	127	145	149	109	130	120	111	96	76	36	2,092
Med&Vet	138	90	120	115	105	121	127	136	134	125	111	97	90	108	79	82	56	46	34	27	15	1,956
Arch&Eng	136	135	161	197	247	232	281	284	238	255	242	223	267	248	280	263	209	179	136	97	56	4,366
SocSci	22	13	32	30	40	42	53	42	47	55	35	40	47	41	48	40	45	27	24	14	10	747
Ius	54	55	57	56	77	104	113	136	145	137	152	143	143	122	141	148	140	113	85	48	32	2,201
Chim	40	58	61	59	65	61	81	88	69	79	95	93	85	86	95	88	76	65	44	31	13	1,432
Secs-	38	39	42	73	60	107	108	116	139	113	158	127	145	149	109	130	120	111	96	76	36	2,092
Med	117	75	97	98	82	92	95	113	112	103	99	77	74	81	55	61	35	33	23	20	11	1,553
Mat	25	38	40	49	63	71	72	73	75	103	74	62	68	65	71	69	77	66	49	43	23	1,276
L-	86	93	90	95	85	131	118	119	120	106	94	92	91	76	84	72	70	53	39	28	19	1,761
Ing-	73	67	78	91	117	131	161	177	148	162	156	161	186	174	197	199	155	128	96	84	49	2,790
Agr&Vet	53	64	79	75	82	78	83	73	87	75	54	67	46	68	56	61	74	36	52	24	9	1,296
Icar	63	68	82	106	130	101	119	107	90	93	86	62	81	74	83	64	54	51	40	13	7	1,574
M-	58	54	80	78	87	99	112	88	98	101	97	102	99	110	101	89	89	54	60	28	16	1,700
Sps	22	13	32	30	40	42	53	41	47	55	35	40	47	41	48	40	45	27	24	14	10	746
Bio	75	87	106	103	112	123	126	133	113	117	119	93	98	124	106	97	80	71	42	27	8	1,960
Fis	21	47	54	43	47	54	51	62	67	50	41	47	34	46	50	44	45	34	38	21	16	912
Geo	30	35	22	39	42	34	46	29	25	28	21	23	19	21	25	25	17	6	9	7	3	506
Tot	757	793	922	995	1,090	1,228	1,339	1,356	1,335	1,322	1,281	1,189	1,216	1,237	1,221	1,187	1,077	848	697	464	252	21,806

TABLE A.6: Academics not linked to a thesis per disciplinary area and expected PhD year (1986-2006)

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Tot
Hum&Law	233	235	222	219	232	202	218	195	190	191	187	162	161	148	160	139	165	141	129	158	133	3,820
Science	301	318	315	272	244	237	207	202	166	138	157	138	124	133	144	159	126	124	144	151	162	3,962
Econ&Stat	90	100	83	91	89	94	94	72	76	70	64	83	54	64	48	61	69	64	85	84	70	1,605
Med&Vet	328	284	247	211	199	219	183	164	135	115	123	114	119	147	166	165	153	119	114	112	87	3,504
Arch&Eng	93	136	131	133	134	106	112	78	81	81	67	72	77	86	84	74	85	80	92	99	107	2,008
SocSci	19	23	14	12	16	23	19	16	22	24	15	24	13	24	21	17	18	24	20	13	15	392
Ius	80	72	75	76	83	67	91	74	60	59	78	61	59	60	55	48	67	48	51	60	51	1,375
Chim	37	33	40	43	40	36	34	39	17	16	21	18	21	16	13	21	21	26	17	24	31	564
Secs-	90	100	83	91	89	94	93	72	76	70	64	83	54	64	48	61	69	64	85	84	70	1,604
Med	309	268	226	197	191	202	175	150	126	107	116	102	107	138	150	160	144	109	107	101	79	3,264
Mat	50	69	49	66	47	60	52	54	44	43	40	40	29	39	47	42	30	26	41	30	33	931
L-	97	89	89	73	89	71	82	70	74	69	44	51	46	45	45	47	41	46	40	51	35	1,294
Ing-	57	83	79	93	82	61	67	45	36	45	42	41	46	54	56	52	57	52	61	72	85	1,266
Agr&Vet	79	81	81	50	51	41	35	37	29	28	28	30	27	22	29	24	21	22	36	32	22	805
Icar	36	53	52	40	52	45	45	33	45	36	25	31	31	32	28	22	28	28	31	27	22	742
M-	56	74	58	70	60	64	45	51	56	63	65	50	56	43	60	44	57	47	38	47	47	1,151
Sps	19	23	14	12	16	23	19	16	21	24	15	24	13	24	21	17	18	24	20	13	15	391
Bio	95	89	102	74	64	70	50	52	50	35	38	39	33	42	45	40	42	39	36	46	49	1,130
Fis	38	48	49	44	32	33	30	23	28	12	27	19	21	19	17	31	13	18	15	22	24	563
Geo	20	14	14	9	18	14	14	11	6	12	10	4	5	4	9	6	8	3	6	8	11	206
Tot	1,064	1,096	1,012	938	914	881	833	727	670	619	613	593	548	602	623	615	616	552	584	617	574	15,291

TABLE A.7: Share of academics linked to a thesis per disciplinary area and expected PhD year (1986-2006)

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Tot
Hum&Law	0.46	0.46	0.51	0.51	0.52	0.62	0.61	0.64	0.66	0.64	0.65	0.68	0.67	0.68	0.67	0.69	0.64	0.61	0.59	0.40	0.34	0.60
Science	0.43	0.50	0.52	0.56	0.61	0.62	0.67	0.68	0.71	0.76	0.71	0.73	0.73	0.74	0.72	0.70	0.73	0.68	0.61	0.49	0.30	0.64
Econ&Stat	0.30	0.28	0.34	0.45	0.40	0.53	0.53	0.62	0.65	0.62	0.71	0.60	0.73	0.70	0.69	0.68	0.63	0.63	0.53	0.48	0.34	0.57
Med&Vet	0.30	0.24	0.33	0.35	0.35	0.36	0.41	0.45	0.50	0.52	0.47	0.46	0.43	0.42	0.32	0.33	0.27	0.28	0.23	0.19	0.15	0.36
Arch&Eng	0.59	0.50	0.55	0.60	0.65	0.69	0.72	0.78	0.75	0.76	0.78	0.76	0.78	0.74	0.77	0.78	0.71	0.69	0.60	0.49	0.34	0.68
SocSci	0.54	0.36	0.70	0.71	0.71	0.65	0.74	0.72	0.68	0.70	0.70	0.63	0.78	0.63	0.70	0.70	0.71	0.53	0.55	0.52	0.40	0.66
Ius	0.40	0.43	0.43	0.42	0.48	0.61	0.55	0.65	0.71	0.70	0.66	0.70	0.71	0.67	0.72	0.76	0.68	0.70	0.63	0.44	0.39	0.62
Chim	0.52	0.64	0.60	0.58	0.62	0.63	0.70	0.69	0.80	0.83	0.82	0.84	0.80	0.84	0.88	0.81	0.78	0.71	0.72	0.56	0.30	0.72
Secs-	0.30	0.28	0.34	0.45	0.40	0.53	0.54	0.62	0.65	0.62	0.71	0.60	0.73	0.70	0.69	0.68	0.63	0.63	0.53	0.48	0.34	0.57
Med	0.27	0.22	0.30	0.33	0.30	0.31	0.35	0.43	0.47	0.49	0.46	0.43	0.41	0.37	0.27	0.28	0.20	0.23	0.18	0.17	0.12	0.32
Mat	0.33	0.36	0.45	0.43	0.57	0.54	0.58	0.57	0.63	0.71	0.65	0.61	0.70	0.63	0.60	0.62	0.72	0.72	0.54	0.59	0.41	0.58
L-	0.47	0.51	0.50	0.57	0.49	0.65	0.59	0.63	0.62	0.61	0.68	0.64	0.66	0.63	0.65	0.61	0.63	0.54	0.49	0.35	0.35	0.58
Ing-	0.56	0.45	0.50	0.49	0.59	0.68	0.71	0.80	0.80	0.78	0.79	0.80	0.80	0.76	0.78	0.79	0.73	0.71	0.61	0.54	0.37	0.69
Agr&Vet	0.40	0.44	0.49	0.60	0.62	0.66	0.70	0.66	0.75	0.73	0.66	0.69	0.63	0.76	0.66	0.72	0.78	0.62	0.59	0.43	0.29	0.62
Icar	0.64	0.56	0.61	0.73	0.71	0.69	0.73	0.76	0.67	0.72	0.77	0.67	0.72	0.70	0.75	0.74	0.66	0.65	0.56	0.33	0.24	0.68
M-	0.51	0.42	0.58	0.53	0.59	0.61	0.71	0.63	0.64	0.62	0.60	0.67	0.64	0.72	0.63	0.67	0.61	0.53	0.61	0.37	0.25	0.60
Sps	0.54	0.36	0.70	0.71	0.71	0.65	0.74	0.72	0.69	0.70	0.70	0.63	0.78	0.63	0.70	0.70	0.71	0.53	0.55	0.52	0.40	0.66
Bio	0.44	0.49	0.51	0.58	0.64	0.64	0.72	0.72	0.69	0.77	0.76	0.70	0.75	0.75	0.70	0.71	0.66	0.65	0.54	0.37	0.14	0.63
Fis	0.36	0.49	0.52	0.49	0.59	0.62	0.63	0.73	0.71	0.81	0.60	0.71	0.62	0.71	0.75	0.59	0.78	0.65	0.72	0.49	0.40	0.62
Geo	0.60	0.71	0.61	0.81	0.70	0.71	0.77	0.73	0.81	0.70	0.68	0.85	0.79	0.84	0.74	0.81	0.68	0.67	0.60	0.47	0.21	0.71
Tot	0.42	0.42	0.48	0.51	0.54	0.58	0.62	0.65	0.67	0.68	0.68	0.67	0.69	0.67	0.66	0.66	0.64	0.61	0.54	0.43	0.31	0.59

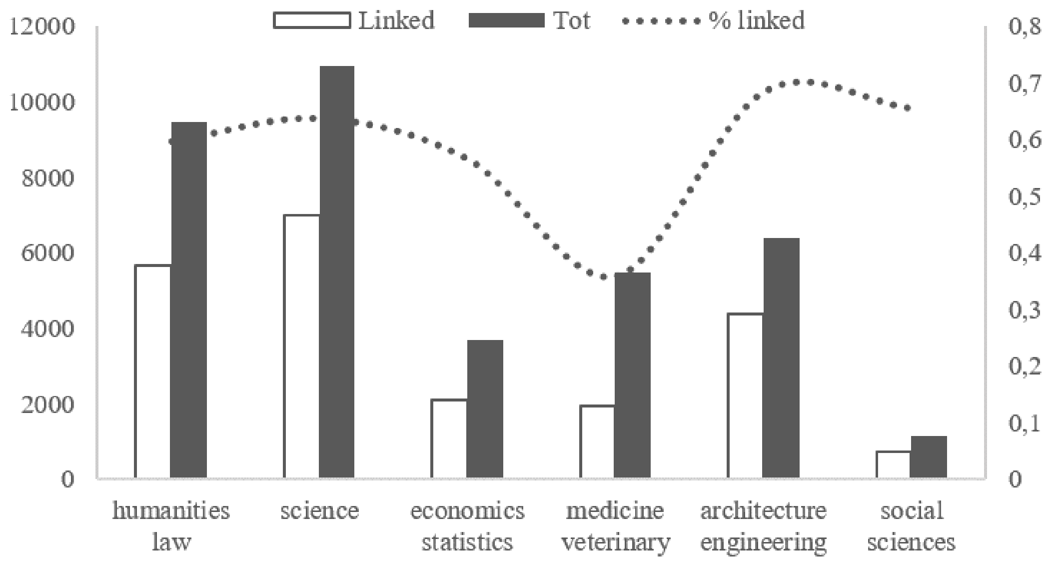

FIGURE A.5: Academics linked to thesis by macro-area (86-06)

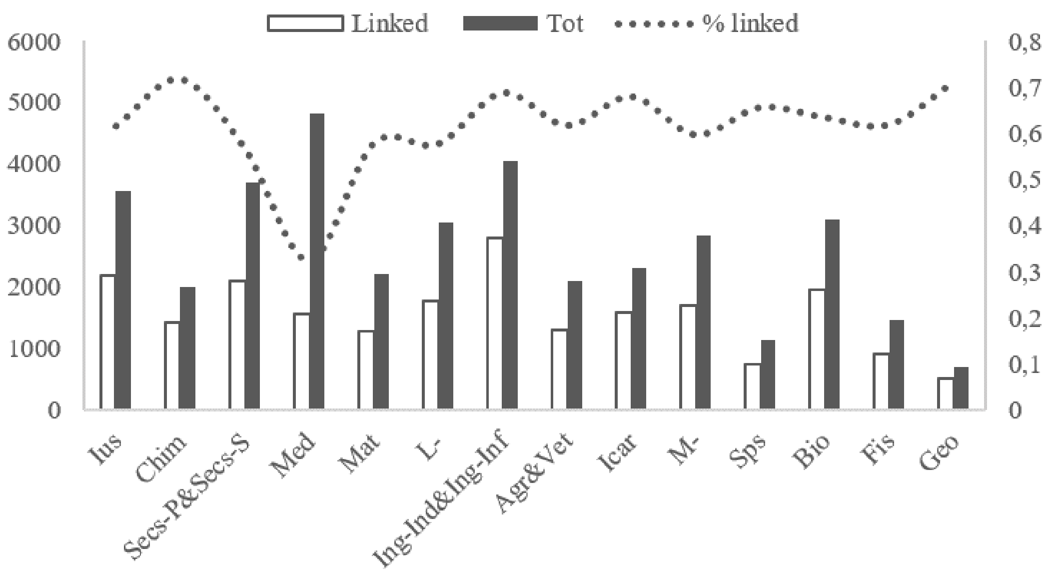

FIGURE A.6: Academics linked to thesis by SSD (86-06)

Appendix B

Appendix to Chapter 2

Summary Statistics

TABLE B.1: Summary statistics (time-to-first appointment)

	N	Mean	SD	Min	Max
AgePhd	1,590	30.40	2.490	26	36
TopPhdUni	1,590	0.429	0.495	0	1
Precocity	1,590	0.128	0.334	0	1
ExternalAdvisor	1,590	0.406	0.491	0	1
NetworkAdvisor	1,590	4.636	7.081	1	81
Inherited	1,590	0.271	0.445	0	1
External	1,590	0.346	0.476	0	1
Inbred	1,590	0.383	0.486	0	1
Female	1,590	0.427	0.495	0	1
SharePubAdvisor	1,590	0.0458	0.211	0	1
ResearchAbroad	1,590	0.0598	0.237	0	1
Econ	1,590	0.315	0.465	0	1
Busi	1,590	0.413	0.492	0	1
Stat	1,590	0.220	0.415	0	1
Hist	1,590	0.0462	0.210	0	1
Cohort9397	1,590	0.182	0.386	0	1
Cohort9802	1,590	0.381	0.486	0	1
Cohort0306	1,590	0.438	0.496	0	1
Productivity	1,590	0.0589	0.172	0	2.010
#Publications	1,590	0.128	0.266	0	2.333
AvgCitations	1,590	2.121	9.631	0	205

TABLE B.2: Impact of connection on candidates' success – Logit (Marginal effects)

	N	Mean	SD	Min	Max
AgePhd	1,672	30.53	2.575	26	38
TopPhdUni	1,672	0.418	0.493	0	1
Precocity	1,672	0.131	0.337	0	1
ExternAladvisor	1,672	0.422	0.494	0	1
NetworkAdvisor	1,672	4.457	6.704	1	81
Inherited	1,672	0.280	0.449	0	1
Inbred	1,672	0.380	0.485	0	1
#ChangeAff	1,672	1.900	0.768	1	7
Female	1,672	0.418	0.493	0	1
SharePubAdvisor	1,672	0.0961	0.329	0	1
ResearchAbroad	1,672	0.166	0.372	0	1
Econ	1,672	0.281	0.450	0	1
Busi	1,672	0.434	0.496	0	1
Stat	1,672	0.239	0.426	0	1
Hist	1,672	0.0390	0.194	0	1
Cohort9397	1,672	0.199	0.399	0	1
Cohort9802	1,672	0.363	0.481	0	1
Cohort0306	1,672	0.438	0.496	0	1
Productivity	1,672	0.118	0.226	0	3.361
#Publications	1,672	0.261	0.352	0	3.400
AvgCitations	1,672	3.013	10.92	0	205

Appendix C

Appendix to Chapter 3

TABLE C.1: Impact of connection on candidates' success – Logit (Marginal effects)

	All (1)	AP (2)	FP (3)	PhD ITA (4)	All (5)	Experimental (6)
Advisor	0.150**	0.143**	0.165*	0.134**	0.143**	0.145**
	(0.0412)	(0.0450)	(0.0701)	(0.0401)	(0.0412)	(0.0408)
coauthor	0.129**	0.126**	0.151**	0.0973**	0.128**	0.125**
	(0.0306)	(0.0375)	(0.0501)	(0.0327)	(0.0297)	(0.0291)
Colleague	0.0615**	0.0421+	0.0942**	0.0338+	0.0634**	0.0633**
	(0.0182)	(0.0239)	(0.0314)	(0.0200)	(0.0182)	(0.0185)
PhD Colleague	-0.00566	-0.00213	0.000348	0.00748	-0.00199	-0.0117
	(0.0223)	(0.0235)	(0.0358)	(0.0219)	(0.0225)	(0.0234)
Advisor's Coauth.	0.0227	0.0349	-0.0218	0.0337	0.00913	-0.00659
	(0.0435)	(0.0561)	(0.0674)	(0.0438)	(0.0433)	(0.0386)
Female					-0.00811	-0.00402
					(0.0141)	(0.0167)
Age					0.0321	0.0185
					(0.0255)	(0.0299)
Age^2					-0.000556+	-0.000389
					(0.000310)	(0.000367)
Experience					0.00331+	0.00239
					(0.00199)	(0.00241)
#Publications					0.00135	0.000984
					(0.000947)	(0.000963)
#Citations					6.59e-05**	6.76e-05**
					(2.46e-05)	(2.62e-05)
Observations	9,165	6,130	3,035	6,772	9,165	6,680
LogLikelihood	-5754	-3720	-1974	-4204	-5661	-3991
Chi-square	251	192.3	160.4	169.9	367.1	279.1
Pseudo R2	0.0494	0.0576	0.0482	0.0491	0.0647	0.0693

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

TABLE C.2: First-stage estimates

	Advisor^f (1)	*Coauthor*^{f} (2)	*Colleague*^{f} (3)
Advisor	.9418901**	-.0105695+	-.0419671*
	(.028357)	(.0060701)	(.0193345)
Coauthor	-.0014728	.9671191**	-.1244072**
	(.0017109)	(.0132444)	(.0186054)
Collegaue	.0028798	-.0010439	.8788412**
	(.0022858)	(.0019937)	(.0230088)
F-statistic	1308.86	1892.64	852.67
SW F-statistics	4269.26	4943.19	2000.63

SE in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Appendix D

Appendix to Chapter 4

CEM Balancing

Table D.1 shows statistics for seven of the selected variables used to perform the CEM. Frequencies and descriptive statistics reported underline the differences between the two sub-populations (most of the difference is plausibly due to the different time horizon at which the two populations refer to).
Table D.2 provides first evidence (univariate absolute difference in means) of balancing between CEM selected treatment and control groups in the overall sample. Mean and standard deviations of the two, equal-size, samples of units are relatively close from one to the next. Before and after reform professors have an average number of 4 papers published on Scopus journals, with an average number of citations per paper close to 8 and an average Scimago journal rank higher than 0.4.

TABLE D.1: Descriptive statistics by treated and controls

	N	Min	Mean	Max	SD
			Untreated		
#Publications	3248	0	6.372229	146	10.57647
AvgCitations	3248	0	13.35514	638.4375	30.63182
AvgSJR	3248	0	0.640323	10.7512	0.938639
CItation Stock	3248	0	167.2127	10759	518.2206
Birthyear	3248	1937	1961.44	1972	4.366033
Experience	3248	1	5.179495	22	2.991797
PhdIta	3248	0	0.482451	1	0.499692
			Treated		
#Publications	3536	0	7.136312	193	12.22742
AvgCitations	3536	0	15.39642	572.4	26.90748
AvgSJR	3536	0	0.740328	19.794	1.092535
Citations Stock	3536	0	217.8213	18939	644.4442
Birthyear	3536	1930	1963.997	1975	4.931687
Experience	3536	1	4.719174	25	3.758706
PhdIta	3536	0	0.583993	1	0.492895

TABLE D.2: Descriptive statistics of matched units by treated and controls

	N	Min	Mean	Max	SD
			CEM–Untreated		
#Publications	1449	0	3.989648	62	7.19591
AvgCItations	1449	0	7.762366	124	13.13088
AvgSJR	1449	0	0.414479	4.439	0.614284
Citations Stock	1449	0	72.89855	1417	161.1979
Birthyear	1449	1943	1962.524	1972	4.176025
Experience	1449	1	4.242926	17	2.626836
PhdIta	1449	0	0.507246	1	0.499947
			CEM–Treated		
#Publications	1449	0	4.162871	69	7.265713
AvgCitations	1449	0	9.036043	142.2222	13.90037
AvgSJR	1449	0	0.452367	3.978545	0.612438
Citations Stock	1449	0	85.77226	1344	182.6217
Birthyear	1449	1942	1962.835	1972	4.335608
Experience	1449	1	4.138716	17	2.824413
PhdIta	1449	0	0.507246	1	0.499947

Bibliography

Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Francesco Rosati (2015). "Selection committees for academic recruitment: Does gender matter?" In: *Research Evaluation* 24.4, pp. 392–404.

— (2016). "Gender bias in academic recruitment". In: *Scientometrics* 106.1, pp. 119–141.

Acker, Joan (1990). "Hierarchies, jobs, bodies: A Theory of Gendered Organizations". In: *Gender & Society* 4.2, pp. 139–158.

Aghion, Philippe, Mathias Dewatripont, Caroline Hoxby, Andreu Mas-Colell, and André Sapir (2010). "The governance and performance of universities: Evidence from Europe and the US". In: *Economic Policy* 25.61, pp. 7–59.

Allison, Paul D . and John A . Stewart (1974). "Productivity Differences Among Scientists: Evidence for Accumulative Advantage". In: *American Sociological Review* 39.4, pp. 596–606.

Allison, Paul D. and J. Scott Long (1990). "Departmental Effects on Scientific Productivity". In: *American Sociological Review* 55.4, pp. 469–478.

Ancaiani, Alessio et al. (2015). "Evaluating scientific research in Italy: The 2004-10 research evaluation exercise". In: *Research Evaluation* 24.3, pp. 242–255.

Angrist J. D. and J. S. Pischke (2008). "Mostly Harmless Econometrics : An Empiricist ' s Companion". In: *Massachusettts I nstitute of Technology and The London school of Economics* March, p. 290.

ANVUR (2013). "Rapporto sullo stato del sistema universitario e della ricerca". In:

Argentin, Gianluca, Gabriele Ballarino, and Sabrina Colombo. "Investire in formazione dopo la laurea: il dottorato di ricerca in Italia". In: *AlmaLaurea Working Papers* 60.

Auriol, Laudeline, Max Misu, and Rebecca Ann Freeman (2013). "Careers of Doctorate Holders: Analysis of Labour Market and Mobility Indicators". In: *OECD Science, Technology and Industry Working Papers*, p. 62.

Avveduto, S. and M. C. Brandi (2004). "Le migrazioni qualificate in Italia". In: *Studi Emigrazione* 41, pp. 797–829.

Bagues, Manuel, Mauro Sylos-Labini, and Natalia Zinovyeva (2015). "Connections in Scientific Committees and Applicants' Self-Selection: Evidence from a Natural Randomized Experiment". In: pp. 1–45.

— (2017). "Does the gender composition of scientific committees matter?" In: *American Economic Review* 107.4, pp. 1207–1238.

Bäker, Agnes (2015). "Non-tenured post-doctoral researchers' job mobility and research output: An analysis of the role of research discipline, department size, and coauthors". In: *Research Policy* 44.3, pp. 634–650.

Ballarino, Gabriele and Sabrina Colombo (2010). "Occupational outcomes of PhD graduates in Northern Italy". In: *Italian Journal of Sociology of Education*, pp. 149–171.

Baruffaldi, Stefano, Fabiana Visentin, and Annamaria Conti (2016). "The productivity of science & engineering PhD students hired from supervisors' networks". In: *Research Policy* 45.4, pp. 785–796.

Bianchetti, Marco and Francesca Della Ratta (2003). "Progetto CNVSU per la ricognizione, raccolta e analisi dei dati esistenti sul dottorato di ricerca e per l'indagine sull'inserimento professionale dei dottori di ricerca". In: pp. 1–53.

Blackaby, David, Alison L. Booth, and Jeff Frank (2005). "Outside offers and the gender pay gap: Empirical evidence from the UK academic labour market". In: *Economic Journal* 115.501.

Blau, Francine D., Janet M. Currie, Rachel T. A. Croson, and Donna K. Ginther (2010). "Can Mentoring Help Female Assistant Professors? Interim Results from a Randomized Trial". In: *American Economic Review* 100.2, pp. 348–52.

Boschini, Anne and Anna Sjögren (2007). "Is Team Formation Gender Neutral? Evidence from Coauthorship Patterns". In: *Journal of Labor Economics* 25.2, pp. 325–365.

Bosquet, Clément, Pierre-Philippe Combes, and Cecilia Garcia-Penalosa (2013). "Gender and competition: evidence from academic promotions in France". In: *CEPR Discussion Paper* DP9711.

Box-Steffensmeier, Janet M. and Bradford S. Jones (2004). *Event history modeling: A guide for social scientists*, p. 234.

Bozeman, Barry, James S. Dietz, and Monica Gaughan (2001). "Scientific and technical human capital: an alternative model for research evaluation". In: *International Journal of Technology Management* 22.7/8, p. 716.

Brait, F, C De Vitiis, R Petrillo, and M Russo (2009). "L' indagine sui dottori di ricerca: un'esperienza pilota". In: *ISTAT Working Papers*.

Brink, Marieke van den and Yvonne Benschop (2012). "Gender practices in the construction of academic excellence: Sheep with five legs". In: *Organization* 19.4, pp. 507–524.

Brink, Marieke van den, Yvonne Benschop, and Willy Jansen (2010). "Transparency in academic recruitment: A problematic tool for gender equality?" In: *Organization Studies* 31.11, pp. 1459–1483.

Cameron, A. Colin and Pavin K. Trivedi (2005). *Microeconometrics: methods and applications. Cambridge university press.*

Cameron, Susan W and Robert T Blackburn (1981). "Sponsorship and Academic Career Success". In: *The Journal of Higher Education* 52.4, pp. 369–377.

Checchi, Daniele (1999). "Tenure. An Appraisal of a National Selection Process for Associate Professorship". In: *Giornale degli Economisti* 58.2, pp. 137–181.

Cipollone, P. E. and S. Avveduto (1998). *La mobilità delle intelligenze in Europa: internazionalizzazione della formazione e dottorato di ricerca*. Franco Angeli.

Clauset, Aaron, Samuel Arbesman, and Daniel B Larremore (2015). "Systematic inequality and hierarchy in faculty hiring networks". In: *Science Advances* 1.1.

Clemente, Frank (1973). "Early Career Determinants of Research Productivity". In: *American Journal of Sociology* 79.2, pp. 409–419.

Clotfelter, Charles T (1999). "The familiar but curious economics of higher education: Introduction to a symposium". In: *The Journal of Economic Perspectives* 13.1, pp. 3–12.

Combes, Pierre Philippe, Laurent Linnemer, and Michael Visser (2008). "Publish or peer-rich? The role of skills and networks in hiring economics professors". In: *Labour Economics* 15.3, pp. 456–474.

Council, National Research (2010). *Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty*. Washington, DC: The National Academies Press.

Cox, D R (1972). "Regression models and life tables". In: *Journal of the Royal Statistical Society. Series B:* 34.2, pp. 187–220.

Crane, Diana (1965). "Scientists at Major and Minor Universities: A Study of Productivity and Recognition". In: *American Sociological Review* 30.5, pp. 699–714.

— (1970). "The Academic Marketplace Revisited: A Study of Faculty Mobility Using the Cartter Ratings". In: *American Journal of Sociology* 75.6, pp. 953–964.

Cruz-Castro, Laura and Luis Sanz-Menéndez (2010). "Mobility versus job stability: Assessing tenure and productivity outcomes". In: *Research Policy* 39.1, pp. 27–38.

Danell, Rickard and Mikael Hjerm (2013). "Career prospects for female university researchers have not improved". In: *Scientometrics* 94.3, pp. 999–1006.

Datta Gupta, Nabanita, Anders Poulsen, and Marie Claire Villeval (2013). "Gender matching and competitiveness: Experimental evidence". In: *Economic Inquiry* 51.1, pp. 816–835.

De Paola, Maria, Michela Ponzo, and Vincenzo Scoppa (2015). “Gender Differences in Attitudes Towards Competition: Evidence from the Italian Scientific Qualification”. In: *IZA Discussion Paper No. 8859* 8859.

Deem, Rosemary (2009). “Leading and managing contemporary UK Universities: Do excellence and meritocracy still prevail over diversity?” In: *Higher Education Policy*. Vol. 22. 1, pp. 3–17.

Dietz, James S., Ivan Chompalov, Barry Bozeman, Eliesh O’Neil Lane, and Jongwon Park (2000). “Using the curriculum vita to study the career paths of scientists and engineers: an exploratory assessment”. In: *Scientometrics* 49.3, pp. 419–442.

Durante, Ruben, Giovanna Labartino, Guido Tabellini, and Roberto Perotti (2009). “Academic Dynasties”.

Ehrenberg, Ronald G. and Liang Zhang (2005). “Do tenured and tenure-track faculty matter?” English (US). In: *Journal of Human Resources* 40.3, pp. 647–659.

Ehrenberg, Ronald G., Paul J. Pieper, and Rachel a. Willis (1998). “Do Economics Departments with lower Tenure Probabilities pay Higher Faculty Salaries?” In: *Review of Economics and Statistics* 80.4, pp. 503–512.

European Commission (2014). *Innovation Union Competitiveness Report 2013*, p. 336.

Fitzenberger, Bernd and Ute Schulze (2014). “Up or out: Research incentives and career prospects of postdocs in Germany”. In: *German Economic Review* 15.2, pp. 287–328.

Freeman, R., E. Weinstein, E. Marincola, J. Rosenbaum, and F. Solomon (2001). *Competition and careers in biosciences*.

Geuna, A and S Shibayama (2016). “Gender Gap in Science in Japan”. In: *University of Tokyo Discussion Paper Series*.

Giambalvo Ornella Montalbetti, Claudia, Nello Scarabottolo, Marco Trezzi, and Francesca Bonini (2014). *Consorzio STELLA - Indagine sui dottori di ricerca 2011-2012 a un anno dal titolo.*

Giambalvo, Ornella and Claudia Montalbetti (2013). *Consorzio STELLA - Indagine sui dottori di ricerca 2010-2011 a un anno dal titolo.*

Ginther, Donna K and Shulamit Kahn (2004). "Women in Economics: Moving Up or Falling Off the Academic Career Ladder?" In: *Journal of Economic Perspectives* 18.3, pp. 193–214.

Gneezy, U., M. Niederle, and A. Rustichini (2003). "Performance in Competitive Environments: Gender Differences". In: *The Quarterly Journal of Economics* 118.3, pp. 1049–1074.

Godechot, Olivier (2016). "The chance of influence: A natural experiment on the role of social capital in faculty recruitment". In: *Social Networks* 46, pp. 60–75.

Godechot, Olivier and Alexandra Louvet (2008). *Academic Inbreeding: An Evaluation*.

Grambsch, Patricia M. and Terry M. Therneau (1994). "Proportional hazards tests and diagnostics based on weighted residuals". In: *Biometrika* 81.3, pp. 515–526.

Granovetter, Mark S. (1973). "The Strength of Weak Ties". In: *American Journal of Sociology* 78.6, pp. 1360–1380.

Granovetter, Mark S (1996). *Getting a Job: A Study of Contacts and Careers*. Vol. 25. 3, pp. 391–392.

Hargens, Lowell L. and Grant M. Farr (1973). "An Examination of Recent Hypotheses About Institutional Inbreeding". In: *American Journal of Sociology* 78.6, pp. 1381–1402.

Hargens, Lowell L. and Warren O. Hagstrom (1967). "Sponsored and Contest Mobility of American Academic Scientists". In: *Sociology of Education* 40.1, pp. 24–38.

Heining, Jorg, Jurgen Jerger, and Jorg Lingens (2007). "Success in the Academic Labour Market for Economists – The German Experience Success in the Academic Labour Market for Economists – The German Experience ". In: *Management Information Systems*.

Horta, Hugo (2013). "Deepening our understanding of academic inbreeding effects on research information exchange and scientific output: New insights for academic based research". In: *Higher Education* 65.4, pp. 487–510.

Horta, Hugo, Francisco M. Veloso, and Rócio Grediaga (2010). "Navel Gazing: Academic Inbreeding and Scientific Productivity". In: *Management Science* 56.3, pp. 414–429.

Iacus, Stefano M., Gary King, and Giuseppe Porro (2012). "Causal inference without balance checking: Coarsened exact matching". In: *Political Analysis* 20.1, pp. 1–24.

Imbens, Guido W and Jeffrey M Wooldridge (2009). "Recent Developments in the Econometrics of Program Evaluation". In: *Journal of Economic Literature* 47.1, pp. 5–86.

Ioannides, Yannis M and Linda Datcher Loury (2004). "Job Information Networks, Neighborhood Effects, and Inequality". In: *Journal of Economic Literature* 42.4, pp. 1056–1093.

ISFOL (2014). *Non sempre mobili. I risultati dell'indagine ISFOL sulla mobilità geografica dei dottori di ricerca*.

Jonkers, Koen (2011). "Mobility, productivity, gender and career development of Argentinean life scientists". In: *Research Evaluation* 20.5, pp. 411–421.

Jungbauer-Gans, M and C Gross (2013). "Determinants of Success in University Careers: Findings from the German Academic Labor Market". In: *Zeitschrift Fur Soziologie* 42.1, pp. 74–92.

Kahn, Shulamit (1993). "Gender Differences in Academic Career Paths of Economists". In: *The American Economic Review* 83.2, pp. 52–56.

Kalbfleisch, John D and Ross L Prentice (2002). "The Statistical Analysis of Failure Time Data". In: *John Wiley and Sons New York*, p. 462.

Kaplan, E. L. and Paul Meier (1958). "Nonparametric Estimation from Incomplete Observations". In: *Journal of the American Statistical Association* 53.282, pp. 457–481.

Leahey, Erin (2007). "Not by productivity alone: How visibility and specialization contribute to academic earnings". In: *American Sociological Review* 72.4, pp. 533–561.

Levin, Sharon G. and Paula E. Stephan (1998). "Gender Differences in the Rewards to Publishing in Academe: Science in the 1970s". In: *Sex Roles* 38.11, pp. 1049–1064.

Long, J. S. and R. McGinnis (1985). "The effects of the mentor on the academic career". In: *Scientometrics* 7.3-6, pp. 255–280.

Long, J. Scott and Mary Frank Fox (1995). "Scientific Careers: Universalism and Particularism". In: *Annual Review of Sociology* 21.1, pp. 45–71.

Long, J. Scott and Robert McGinnis (1981). "Organizational Context and Scientific Productivity". In: *American Sociological Review* 46.4, p. 422.

Long, J. Scott, Paul D. Allison, and Robert McGinnis (1979). "Entrance into the Academic Career". In: *American Sociological Review* 44.5, p. 816.

— (1993). "Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity". In: *American Sociological Review* 58.5, pp. 703–722.

Lutter, Mark and Martin Schröder (2016). "Who becomes a tenured professor, and why? Panel data evidence from German sociology, 1980-2013". In: *Research Policy* 45.5, pp. 999–1013.

Luzzatto, Giunio (2009). "Gli esiti occupazionali dei Dottori di ricerca: una indagine a Bologna 1". In: pp. 1–22.

Marsden, Peter V and Elizabeth H Gorman (2001). "Social Networks, Job Changes and Recruitment". In: *Sourcebook of Labor Markets Evolving Structures and Processes*, pp. 467–502.

McDowell, John M., Larry D. Singell, and James P. Ziliak (2001). "Gender and promotion in the economics profession". In: *Industrial and Labor Relations Review* 54.2, pp. 224–244.

McPherson, Michael S and Morton Owen Schapiro (1999). "Tenure Issues in Higher Education". In: *Journal of Economic Perspectives* 13.1, pp. 85–98.

Melin, Göran (2005). "The dark side of mobility: negative experiences of doing a postdoc period abroad". In: *Research Evaluation* 14.3, pp. 229–237.

Merton, Robert K. (1976). "The Sociology of Science: Theoretical and Empirical Investigations". In: *Contemporary Sociology* 5, p. 555.

Modena, M.G., M. Lalla, and R. Molinari (1999). "Determinants of career structure and advancement among Italian cardiologists. An example of segregation and discrimination against women?" In: *European Heart Journal* 20.17, pp. 1276–1284.

Mood, Carina (2010). "Logistic regression: Why we cannot do what We think we can do, and what we can do about it". In: *European Sociological Review* 26.1, pp. 67–82.

Mouw, Ted (2006). "Estimating the Causal Effect of Social Capital: A Review of Recent Research". In: *Annual Review of Sociology* 32.1, pp. 79–102.

Niederle, M. and L. Vesterlund (2007). "Do Women Shy Away From Competition? Do Men Compete Too Much?" In: *The Quarterly Journal of Economics* 122.3, pp. 1067–1101.

Perna, Laura W. (2001). "Sex and race differences in faculty tenure and promotion". In: *Research in Higher Education* 42.5, pp. 541–567.

Pezzoni, Michele, Valerio Sterzi, and Francesco Lissoni (2012). "Career progress in centralized academic systems: Social capital and institutions in France and Italy". In: *Research Policy* 41.4, pp. 704–719.

Pezzoni, Michele, Jacques Mairesse, Paula Stephan, and Julia Lane (2016). "Gender and the publication output of graduate students: A case study". In: *PLoS ONE* 11.1.

Reskin, Barbara F. (1979). "Academic Sponsorship and Scientists' Careers". In: *Sociology of Education* 52.3, p. 129.

Romano, Maria Francesca and Nello Scarabottolo (2012). *Consorzio STELLA - Indagine sui dottori di ricerca 2009-2010 a un anno dal titolo.*

Rosenbaum, P R and D B Rubin (1983). "The central role of the propensity score in observational studies for causal effects". In: *Biometrika* 70.1, pp. 41–55.

Sansone, Alessandro (2016). "l'offerta formativa e la domanda di formazione terziaria". In: pp. 2–14.

Sanz-Menéndez, Luis, Laura Cruz-Castro, and Kenedy Alva (2013). "Time to Tenure in Spanish Universities: An Event History Analysis". In: *PLoS ONE* 8.10, pp. 1–18.

Schizzerotto, Antonio (2006). "Gli esiti occupazionali dei dottori di ricerca degli Atenei di Milano , Milano-Bicocca e Trento". In: pp. 1–43.

Stephan, Paula (2012). *How Economics Shapes Science*.

Stephan, Paula E. (1996). "The economics of science". In: *Journal of Economic Literature* 34.3, pp. 1199–1235.

Van den Berg, Gerard J. (2001). *Duration models: specification, identification and multiple durations*.

Van den Brink, Marieke and Yvonne Benschop (2014). "Gender in academic networking: The role of gatekeepers in professorial recruitment". In: *Journal of Management Studies* 51.3, pp. 460–492.

Verzillo, S (2013). "Decentralized academic selection mechanisms in Italy: opportunity or parochialism?" In: *University of Milan Working Paper Series*.

Vesterlund, Lise, Linda Babcock, and Laurie Weingart (2014). "Breaking the Glass Ceiling with "No": Gender Differences in De-clining Requests for Non Promotable Tasks". In: *Carnegie Mellon Working Paper*.

Weisshaar, Katherine (2017). "Publish and perish? An assessment of gender gaps in promotion to tenure in academia". In: *Social Forces* 96.2, pp. 529–560.

Zinovyeva, Natalia and Manuel Bagues (2015). "The role of connections in academic promotions". In: *American Economic Journal: Applied Economics* 7.2, pp. 264–292.

Zoghi, Cindy (2003). "Why have public university professors done so badly?" In: *Economics of Education Review* 22.1, pp. 45–57.